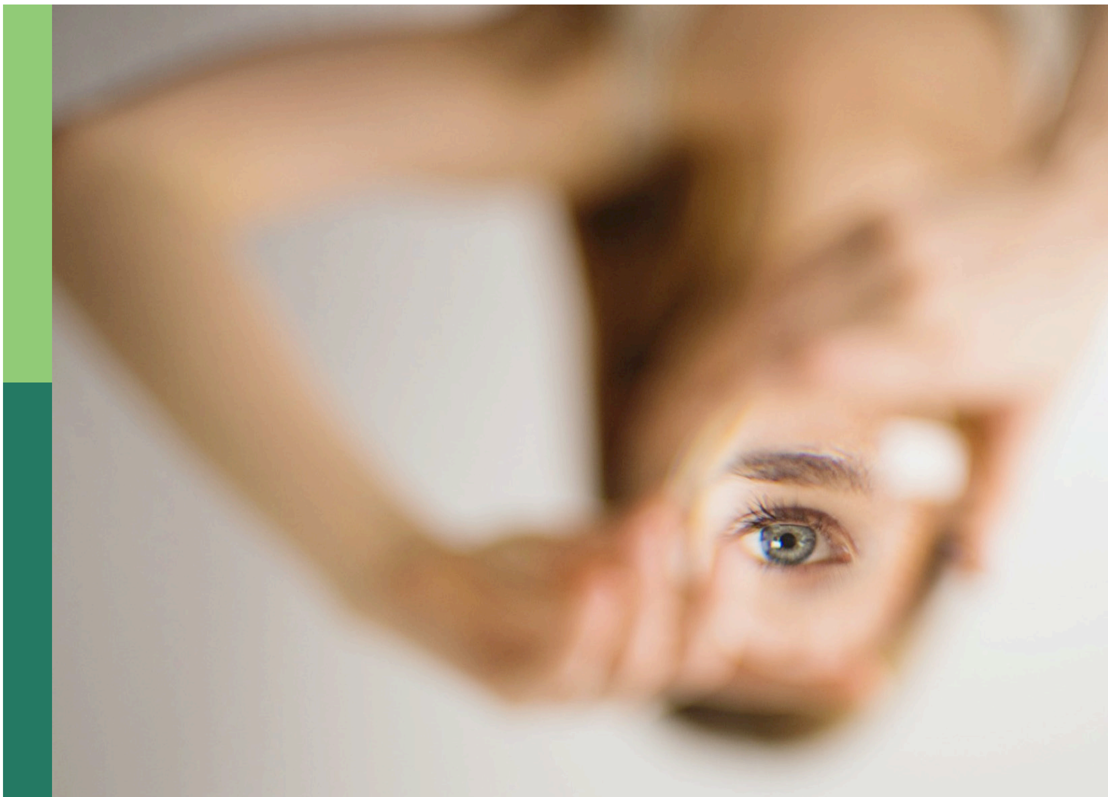
frontiers | Research Topics

Cognitive factors in bilingual language processing

Edited by
Yan Jing Wu, Koji Miwa and Haoyun Zhang

Published in
Frontiers in Psychology



ISSN 1664-8714
ISBN 978-2-83250-620-2
DOI 10.3389/978-2-83250-620-2

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Cognitive factors in bilingual language processing

Topic editors
Yan Jing Wu — Ningbo University, China
Koji Miwa — Nagoya University, Japan
Haoyun Zhang — The Pennsylvania State University (PSU), United States

Citation
Wu, Y. J., Miwa, K., Zhang, H., eds. (2023). *Cognitive factors in bilingual language processing*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83250-620-2

Table of contents

frontiers | Frontiers in Psychology

TYPE Editorial
PUBLISHED 13 October 2022
DOI 10.3389/fpsyg.2022.1055759

OPEN ACCESS

EDITED AND REVIEWED BY
Xiaolin Zhou,
Peking University, China

*CORRESPONDENCE
Yan Jing Wu
wuyanjing@nbu.edu.cn

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 28 September 2022
ACCEPTED 04 October 2022
PUBLISHED 13 October 2022

CITATION
Wu YJ, Miwa K and Zhang H (2022) Editorial: Cognitive factors in bilingual language processing. *Front. Psychol.* 13:1055759. doi: 10.3389/fpsyg.2022.1055759

Editorial: Cognitive factors in bilingual language processing

Yan Jing Wu[1]*, Koji Miwa[2] and Haoyun Zhang[3]

[1]Faculty of Foreign Languages, Ningbo University, Ningbo, China, [2]Graduate School of Humanities, Nagoya University, Nagoya, Japan, [3]Centre for Cognitive and Brain Sciences, University of Macau, Macau, Macau SAR, China

KEYWORDS
bilingualism, cognition, psycholinguistic, neuroscience, translation

Editorial on the Research Topic
Cognitive factors in bilingual language processing

In this Research Topic, we received a wide range of submissions concerning several aspects of language processing in bilinguals. It is our greatest honor to review such a fascinating collection of articles, but space limits allowed publishing only a small fraction of them. Here, we are delighted to present eleven articles, ten original research based on empirical studies and one review. Their main findings and perspectives are summarized below.

At the center of bilingual language processing is the question of how bilinguals access (and control) lexical-semantic representations of the two languages. Experimental psychology and neuroscience have made the case that when bilingual individuals speak and read in one language, the other language is simultaneously activated, a phenomenon sometimes referred to as non-selective activation or cross-language interactions. Zeng et al. examine the effects of task demands and L2 proficiency on cross-language interactions by comparing the performance of Chinese-English speakers with high and low proficiency in English on a semantic and a lexical task. Their findings add to the existing literature of cross-language interactions and shed new light on classic psycholinguistic models.

A classic paradigm in the study of cross-language interactions involves the use of cognates, words that share the same or similar semantic contents and lexical (phonological and orthographic) forms between bilinguals' two languages. Typically, cognates are processed faster as compared to non-cognate controls by bilinguals (i.e., the cognate facilitation effect). However, Frances et al. show that Spanish-English cognates with orthographic similarities lead to greater response time and less accuracy, indicating an unexpected cognate inhibitory effect. Interestingly, the effect which is found in an auditory task is affected by the speaker's accent. The accent of the bilingual's native language (but foreign to that of the L2) reduces this inhibitory effect. While the exact mechanism is still under exploration, Frances et al.'s results bring new insights into mental operations underlying the cognate effect, an observation that is almost as old as research in bilingualism itself. Unlike cognates in alphabetic languages, those in two languages with different script systems have completely unrelated orthography while sharing semantics and phonology in common. Using event-related brain potentials, Chen et al. investigate how word concreteness affects the processing of cross-script cognates in a lexical decision task. A masked translation priming paradigm involving both forward and backward translation directions is applied to prevent spurious effects due to participants' awareness of the cognate status of the critical stimuli. Results of Chen et al. highlight an interaction between semantic (i.e., word concreteness) and lexical (i.e., phonology) processing of cross-script cognates.

Language acquisition context is another important factor that affects language processing in bilinguals and the interaction between L1 and L2. Using a lexical decision task, Hevia-Tuero et al. compare Spanish-English bilingual children from a monolingual school to those from a bilingual school with a focus on cross-language (phonological) interference. The lexical decision task involves real words and pseudohomophones in both Spanish and English. Hevia-Tuero et al.'s results show the effects of both instructional language (i.e., language acquisition context) and level of education on the ability to control L1-to-L2 interference on phonological access and also at the level of grapheme-phoneme correspondence regularities. Short-term language training is an effective way to examine language (L2) acquisition process in an experimental context. Deng et al. train a group of Chinese learners of English with subject-verb agreement in English and test the same group of participants with a different set of stimuli that involve the same syntactic structure. A syntactic transfer effect is found as the processing of grammatically incorrect sentences induces a larger P600 effect, classic ERP index of syntactic violation, as compared to correct sentences, suggesting that even for late L2 learners, syntactic knowledge can be developed with a relatively short period of training.

Embodiment is another perspective by which language processing is studied and compared between bilinguals and monolinguals. It has been shown that the sensorimotor system is involved (i.e., language embodiment) when advanced bilinguals process words in L2, but less is known regarding L2 beginners. Bai and He examine how less advanced bilinguals process spatially associated words in L2 and show that the degree of embodiment as indexed by automatic activation of sensorimotor response is dependent on the level of task demands.

Spoken word segmentation, a process in which listeners spontaneously "cut" continuous utterances into meaning parts during oral communication, presents a serious problem for less advanced L2 learners. Yang et al. study the cognitive mechanisms of spoken word segmentation by characterizing the interaction between spoken word segmentation efficiency on one hand and cognitive inhibition, cognitive flexibility, and L2 listening proficiency on the other hand. Yang et al.'s findings support an interactive model as they show that both the bottom-up and top-down processes determine spoken word segmentation performance in bilinguals. In a similar context, Guan et al. investigate to what extent bilingual listeners can take advantage of pitch accents as a memory cue when recalling contents of spoken discourse that was presented in L2. Pitch accents are detailed auditory information that can be used as a processing cue to facilitate speech comprehension and recalls. In Guan et al., both L2 proficiency and working memory are considered as cognitive factors in an auditory recognition task, where signal detection theory is applied in the analysis of the data.

Interpretation and translation are unique processing contexts that are often studied as an independent subject of bilingualism. Zhao takes a novel approach to interpretation by examining how L2 proficiency, working memory, and anxiety levels affect the fluency when interpreting speeches from L2 to L1, effectively taking considerations of linguistic, cognitive, and emotional factors in the same functioning context of bilinguals. In addition to its contribution to the growing literature of interpretation, the findings of Zhao have real-life implications for practitioners. Similar to Zhao, but in a visual translation context, Yuan and Tu study the affective valence of visual imagery expressions when English-Chinese bilinguals read a classic Chinese poem. While poetry comprehension in L2 is an underdeveloped subject in bilingualism research, Yuan and Tu's study reveals how translation strategies and cultural factors affect emotional responses to words that are intended, in the native language of the original poem, to stimulate visual imagination and emotional reactions associated with the imagination.

Knowledge of cross-language variances is the foundation of research in bilingualism. In the review article by Li, semantic fusion, which is the realization and integration of multiple semantic roles in one syntactic element, is investigated and compared between Chinese and English using corpus analysis. Since merging semantic roles is a critical step in event alignment during sentence comprehension, variances in the semantic fusion process between languages could be a potential factor affecting language processing in bilinguals. Interestingly, however, Li shows highly comparable patterns of corpus data when short sentences with two verbs designating a double semantic role (i.e., patient and agent) in one noun are analyzed between Chinese and English, suggesting universality in syntactic processing across languages.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This work was funded by the National Social Science Fund of China (21AYY014).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

ORIGINAL RESEARCH
published: 03 January 2022
doi: 10.3389/fpsyg.2021.777225

An Event-Related Potentials Study on the Syntactic Transfer Effect of Late Language Learners

Taiping Deng[1], *Dongping Deng*[2] *and Qing Feng*[3]*

[1] School of Marxism Studies, Zhejiang University of Finance and Economics, Hangzhou, China, [2] No. 3 Middle School of Gan County, Ganzhou, China, [3] College of Education, Hebei Normal University, Shijiazhuang, China

This study explored the syntactic transfer effect of the non-local subject-verb agreement structure with plural head noun after two intensive phases of input training with event-related potentials (ERP). The non-local subject-verb agreement stimuli with the plural head nouns, which never appeared in training phases, were used for the stimuli. A total of 26 late L1-Chinese L2-English learners, who began to learn English after a critical period and participated in our previous experiments, were asked back to take part in this syntactic transfer experiment. Results indicated that a significant ERP component P600 occurred in the key region (the verb) of the sentences with syntactic violations in the experimental group, but none occurred in the control group. This demonstrated that there was a significant transfer effect of the input training. The possible theoretical explanation was provided and also the malleability of the late L2 learners was discussed.

Keywords: transfer effect, input training, syntactic representation, P600, malleability

OPEN ACCESS

Edited by:
Yan Jing Wu,
Ningbo University, China

Reviewed by:
Ali Idrissi,
Qatar University, Qatar
Bin Shao,
Zhejiang University, China

***Correspondence:**
Qing Feng
msfengq@163.com

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *15 September 2021*
Accepted: *29 November 2021*
Published: *03 January 2022*

Citation:
Deng T, Deng D and Feng Q (2022) An Event-Related Potentials Study on the Syntactic Transfer Effect of Late Language Learners. Front. Psychol. 12:777225. doi: 10.3389/fpsyg.2021.777225

INTRODUCTION

The Input Factor and L1 Representation Entrenchment

In the L1 field of syntactic acquisition, cognitive grammar theory such as the usage-based models advocates that input is an important factor in representation entrenchment. Each time the structure (input) is encountered, the representation could be entrenched deeper (Barlow and Kemmer, 2000; Langacker, 2000; Bybee, 2006).

Studies on the differences in L1 speakers provide evidence for the role of input in representation entrenchment (Dabrowska, 1997, 2008a,b; Kayne, 2000; Chipere, 2001; Luka and Barsalou, 2005; Brooks and Sekerina, 2006; Street and Dąbrowska, 2010; Luka and Choi, 2012). For example, Chipere (2001) tested two groups of L1 English adult students from the same school: low academic attainment students (LAA in abbreviation) with a score "D" or below in English curricula and a high academic attainment group (HAA in abbreviation) with a score "A" (Chipere, 2001). The students were tested on comprehension and recall of complex NP sentences. The results indicated that great differences existed in the comprehension performances with the LAA group performing much worse. Then, the LAA group was divided into two subgroups. One subgroup received memory training and the other group received comprehension training which involved explicit instruction and also the practice session. The results of the new complex NP sentences comprehension showed that memory training led to only improvement in recall task and comprehension training led to improvement on both recall task and comprehension task. The results from this research suggested the important role of input or experience especially on particular grammatical structures.

Similarly, Street and Dąbrowska (2010) conducted two experiments to explore the differences in passive structures and quantifiers between HAA participants and low LAA participants. Experiment

1 demonstrated considerable individual differences in these adult native speakers which were strongly connected with their educational attainments. While the HAA participants performed at ceiling in both conditions, the LAA performed worse. Experiment 2 was conducted with another group of participants. The results showed that training led to significant improvement only on structure trained. This demonstrated that the factor experience or input played an important role in L1 syntactic entrenchment. That is to say, as the HAA group might receive much more input or experience with the syntactic structures than the LAA group, its corresponding representations might be entrenched deeper than those of the LAA group. Thus, the HAA performed better than the LAA in comprehension test.

According to the usage-based models of language (Barlow and Kemmer, 2000; Langacker, 2000; Bybee, 2006), structure emerges from use. In other words, linguistic knowledge or underlying syntax is shaped by usage factors such as input. As input that the language learner is exposed to contains recurrent patterns or means of the repeated structures, then the syntax with the special structures becomes entrenched through repeated use or input. Input together with cognitive factors forms the representation. These models predict frequency effects and individual differences which might be attributable to input factors. In accordance with these models and the studies above, repeated input with specific structures entrenches the representation deeper, and the entrenched syntactic representation brings better performance in language comprehension, etc. Here a meaningful question is raised whether the repeated input or use refers to the identical input or whether it can be extended to similar structures but in different expressions. Can it can be extended to the similar structures not trained in the training phase, that is, can the transfer effect be predicted? What mechanism might facilitate its occurrence? Is there a possibility that it originates from the entrenched representation? It could be interesting to explore the possible connection between the transfer effect and the usage-based theory.

The Input Factor and L2 Syntactic Processing

In recent years, the body of research dealing with syntactic acquisition of a non-native language has grown greatly, especially the acquisition of late L2 learners. The late L2 learners are those who began L2 learning after a critical period (roughly adolescence) and are thought to encounter great difficulty in L2 acquisition (Foucart and Frenck-Mestre, 2012). Many researchers have shown great interest in the affects of non-native syntactic acquisition (Juffs and Harrington, 1995; Weber-Fox and Neville, 1996; Hahne, 2001; Jiang, 2004, 2007; Cunnings and Clahsen, 2007; Felser and Roberts, 2007; Morgan-Short et al., 2010, 2012a,b; Van Hell and Tokowicz, 2010; Foote, 2011; Roberts and Felser, 2011; White et al., 2012; Bowden et al., 2013; Hopp, 2013a,b, 2015; Deng et al., 2015, 2017; Boxell and Felser, 2017; Cunnings, 2017; Felser and Drummer, 2017). Many factors are thought to be critical in L2 syntactic acquisition such as the age of acquisition (DeKeyser, 2012, 2013), the L1 background (Chen et al., 2007; Sabourin and Stowe, 2008), the language proficiency (Ojima et al., 2005; Steinhauer et al., 2009; White et al., 2012), the input factor and its training method (Ellis, 2002; Dussias and Sagarra, 2007; Ellis and Collins, 2009; Morgan-Short et al., 2012a,b; Unsworth, 2013; Deng et al., 2015, 2017; Deng and Chen, 2019), grammatical integration ability (Hopp, 2013a, 2015), etc. For example, DeKeyser (2012) claimed that the age of acquisition was the decisive factor in affecting the syntactic acquisition. Steinhauer et al. (2009) proposed the syntactic processing performances of the non-native language relied on proficiency (Steinhauer et al., 2009).

Among all the factors mentioned above, the input factor, which has been investigated recently, was also thought to be an important factor that affected L2 syntactic acquisition (Morgan-Short et al., 2012a,b; Montrul et al., 2013; Deng et al., 2015). Morgan-Short et al. (2012b) adopted an artificial language training paradigm to explore whether the type of language training crucially impacts syntactic acquisition. The participants were randomly divided into implicit training group and explicit training group. The results showed that, at the high proficiency phase after input training, the implicit training group showed native-like processing paradigm with an anterior positivity followed by P600 component, which indicated the important role of the input and input type in syntactic acquisition.

Deng et al. (2015) provided the electrophysiological evidence that input was a very important factor. A pretest-training-posttest paradigm was adopted to investigate the effect of input training. Participants with relatively low proficiency were recruited. After two intensive and specific input training phases with non-local subject-verb agreement structure as "The price of the car was very high," the P600 component absent in the pretest was elicited in the posttest of the experimental group. For the control group, as they received other structures training, the P600 component was absent from the pretest to posttest when the syntactic structures were violated compared with the correct versions. This may be due to the entrenched representation, which is in accordance with the cognitive grammar theory in L1 which emphasizes the input factor in syntactic representation entrenchment. As the experimental group received the specific non-local subject-verb agreement structure with singular head noun, an interesting and new issue turned up. Can the experimental group who had been trained with the non-local subject-verb agreement structure show sensitivity to the similar forms of subject-verb agreement structure violations? Was the entrenched representation limited to the subject-verb agreement structure with singular head noun trained, or could it be extended to the similar structure in different expression? What would happen if the experimental group encountered the grammatically violated sentence such as "The girls of the family was very beautiful and polite"? The present study aimed to explore this question about the syntactic transfer within L2 context.

Transfer is thought of as a ubiquitous, continuous, systematic use of selected parts of the immense body of prior knowledge, and it means the use of previously acquired knowledge or skills in new learning or problem-solving situations (Steiner, 2001). As the materials of the present study were quite similar to those of Deng et al. (2015), the only differences were the singular or plural head noun and their corresponding predicate.

Then, according to the similarity theory provided by Thorndike and Woodworth (1901), the transfer could happen in cases where common elements were shared between the source and the target. Transfer increased proportionally with the number of such overlapping associations in the learning and the test tasks (Thorndike, 1913). This means that the similarity between previous and actual learning content and processes may play a crucial role. This transfer effect, called linguistic transfer, has also been widely studied in the language field (Kellerman, 1979; Sharwood Smith, 1986; Bertram and Kuperman, 2019). Linguistic transfer mainly focused on the cross-language transfer, that is, from L2 to L1 or L1 to L2 on syntactic, lexical, and metalinguistic levels (Kellerman, 1979; Sharwood Smith, 1986; Pavlenko and Jarvis, 2002; Brown and Gullberg, 2008). Syntactic transfer here refers to the ability of using the entrenched representation in processing the new syntactic structure that is not trained. Yet, little attention is paid to the intralinguistic transfer effect within the L2 context which is of great importance. The present study set to explore this syntactic transfer effect by focusing on the subject-verb agreement structure. As the materials trained and investigated highly overlap, the transfer effect might take place.

The Current Study

The current study aimed to investigate transfer effect of linguistic input training in late L1-Chinese L2-English learners within the L2 context. The participants who took part in the study of Deng et al. (2015) were invited back to participate in this independent transfer effect test.

As introduced above, in Deng and colleagues' studies, the materials for the training phases were non-local subject-verb agreement structure with singular head noun, while the similar structure with plural head noun was not provided in that study. Though the previous study provides evidence for the relationship between input and representation entrenchment, it is only limited to a specific structure, that is, the structure with a singular head noun. As the routine paradigm of the training-related studies, transfer effect thus comes into consideration. The present study aimed to explore the transfer effect for the similar structure with a plural head noun. By exploring this syntactic transfer effect, we tried to answer the research questions as follows: Could the relationship between the input and syntactic representation be one-to-one correspondence? Or could it be extended to the broader syntactic category? Could these late L2 learners still show plasticity in L2 syntactic acquisition?

As the Chinese language is well known for its impoverished system of grammatical morphology (Li et al., 1993, 2004) and Chinese syntax does not require any subject-verb agreement (Li et al., 1993; Jiang, 2004; Chen et al., 2007) and any nominal subject (Chen et al., 2007), it was difficult for Chinese learners to successfully acquire this structure. The subject-verb agreement structure in English includes several expressions as local "The boy usually goes to school by bus," non-local as "The price of the cars was very high," etc. Though the experimental group in the study of Deng et al. (2015) showed sensitivity to the syntactic violations after two sessions of training, it is still unknown whether this specific input training with singular head noun could be transferred to the same subject-verb agreement structure but in different expressions. Therefore, the present study aimed to take a small step in exploring transfer effect in late English learners. Specifically, we are concerned with whether participants showed sensitivity to the subject-verb agreement violations with plural head nouns which were not trained.

In this transfer effect test, the experimental materials were sentences with violations of non-local subject-verb agreement with plural head nouns and their grammatical counterparts. These sentences were newly constructed for the transfer effect experiment.

To sum up, the present study used the ERP technique to investigate the transfer effect of linguistic input training in late Chinese-English learners, using subject-verb agreement structures with plural head nouns as the stimuli. The focus will be on the P600 component as the indicator.

MATERIALS AND METHODS

Materials

The materials in this transfer experiment were all newly constructed. The experiment consisted of the non-local subject-verb agreement structure with plural head nouns, but it differs from the materials in the study of Deng et al. (2015).

The materials in this transfer experiment were non-local subject-verb agreement with plural head nouns, including the grammatical correct sentences and the grammatical incorrect counterpart. After the plausibility ratings, 80 pairs of sentences (160 in total) were chosen for the formal transfer experiment: half grammatical and half ungrammatical. To balance the materials, two lists were finally adopted with each list containing 80 experimental sentences (40 correct and 40 incorrect) and 160 fillers in other syntactic structures except subject-verb agreement were included.

The examples of the experimental stimuli are presented in **Table 1**.

For the abbreviations in the table, the first capital letter "P" represents "plural head noun." The second capital letter "S" represents "singular local noun." The third capital letter "P" or "S" represents "plural verb (were)" or "singular verb (was)" respectively. "G" or "UG" means "grammatical" or "ungrammatical." It should be noted that the sentence marked as PSS-UG is the ungrammatical version of the sentence PSP-G.

Eighteen Chinese college students from the same background as the participants were asked to rate the plausibility of syntactically correct experimental sentences on a 5-point

TABLE 1 | Sentence examples for the transfer experiment.

Group	Sentence type	Examples
EG	PSP_G	The girls of the family were very beautiful.
	PSS_UG	The girls of the family was very beautiful.
	Fillers	She used to eat apples after supper.
CG		This part was the same as the EG.

EG, experimental group; CG, control group.

scale where 1 signified *definitely implausible* and 5 signified *perfectly plausible*.

These participants never participated in our formal experiments. Sentences with a mean semantic plausibility above 4 were selected. In total, 80 plausible sentences were chosen for the final transfer experiment (M_{PSP} = 4.29, SD = 0.21) on the plausibility rating.

Participants

The participants, who took part in the experiment of Deng et al. (2015), were invited to come back to participate in this independent transfer effect experiment. Because some participants graduated from school, eventually 26 college students participated in this experiment: 13 participants in the EG (8 females, average age = 23.71 years old, average age of classroom exposure = 12.07 years) and 13 participants in the CG (6 females, average age = 21.21 years old, average age of classroom exposure = 12.21 years). In short, 26 participants participated in this transfer effect test.

All participants were late L1-Chinese L2-English learners. They received classroom teaching of English in China and none had experience of living in English-speaking countries. They had passed CET (College English Test) band 4 but not band 6. CET is a large-scale standardized exam conducted to evaluate the college students' English proficiency, with 710 as full marks and 425 as the passing line. Band 6 represents a higher proficiency level than Band 4. The test consists of five parts including listening comprehension, reading comprehension, vocabulary knowledge, grammar knowledge, and writings. Only those who have successfully passed the Band 4 were qualified to take the Band 6 exam.

These participants completed the Oxford Placement Test (OPT) and a self-rating questionnaire. The OPT is a standardized objective test for university English foreign language classes, which is proved to be an effective instrument and a reliable means of grading students at all levels (Moll, 1999). It includes 25 multiple-choice questions and a cloze test, and the full mark is 50. The higher the score is, the higher the proficiency is. The 5-point self-rating questionnaire where "1" signifies *quite poor* and "5" *highly proficient* is used to evaluate the participants' listening, speaking, reading, and writing skills. It is a subjective indicator of English proficiency.

All participants reported being right-handed and having normal or corrected-to-normal vision. A written informed consent form was signed before the formal experiment, and money compensation was provided for the participation.

Experimental Procedure

The participants who finished the sessions of the study of Deng et al. (2015) were asked to take part in this independent study on transfer effect.

The test sentences were those that were non-local subject-verb agreement structure with plural head nouns. Two counterbalanced lists were created. Participants were randomly distributed to one of the two lists. One hundred and sixty fillers were constructed with half of the fillers being syntactically incorrect (i.e., containing violations in verb subcategorizations or reflexive pronouns, etc., for example, the violation in reflexive pronoun of the sentence "The boy quickly adopt herself to new circumstances") and the other half being simple grammatical sentences with various sentence structures.

Participants' ERPs were recorded when they were reading the test sentences. In accordance with the previous study (Havas et al., 2012), each trial began with an asterisk fixation (500 ms) in the center of the screen, followed by test sentences that were presented word-by-word in the center of the screen (500 ms per word with an inter-stimulus interval of 500 ms). The last word of each sentence was followed by two asterisks indicating the end of the sentence. Half of the sentences were followed by a comprehension question to make sure the participants read the sentences attentively. The comprehension questions were relatively easy in order to not cause much pressure for participants. For example, the question for the example (1) listed above was "Were the girls very beautiful?" Participants were tested individually in a quiet room and were asked to minimize their blinks and body movements. The ERP recording session began with 10 practice trials to make sure the participants were familiar with the procedure. They could take a short break every 40 trials. Each ERP recording session lasted about 2 h including preparations.

Event-Related Potentials Data Analysis

EEG signals were recorded at a 1,000 Hz sampling rate from a 64-channel Quik-cap with Ag/AgCI electrodes. EEG electrodes were placed according to the extended 10–20 system. All electrodes were referenced to the left mastoid during recording and off-line referenced to linked mastoids. Impedances were kept below 5 KΩ. Eye movements were measured using vertical EOG with two electrodes placed above and below the left eyes, and the horizontal EOG with two electrodes placed to the outer canthi of the two eyes. EEG data analysis was performed using Scan 4.3. The electrophysiological signals were filtered with a bandpass of 0.05–100 Hz (half-amplitude cutoffs) at a sampling rate of 500 Hz. In the off-line analysis, the EEG was filtered with a 0.12–40 Hz band-pass filter. Trials with voltage exceeding $\pm$ 90 μV and trials with eye movements were excluded from ERP averages, resulting in an exclusion of about 12.6% and 10.4% of the trials in the EG and the CG respectively. As the task used in the present study was sentence comprehension, both trials answered correctly and incorrectly were included in the EEG data analysis. ERPs time-locked to the onset of the violation word (i.e., "was") or matched control word (i.e., "were") were averaged for each participant for all the electrodes from −200 to 1,000 ms.

The main ERP component of interest was P600, which is maximal at centro-parietal electrodes (Bowden et al., 2013). Therefore, the electrodes selected for data analysis after visual inspection were: left region (C5, CP5, P5, PO5), central region (Cz, CPz, Pz, POz), right region (C6, CP6, P6, PO6). Time window of 500–1,000 ms was selected to analyze P600 in accordance with the previous research (Chen et al., 2007).

Mean amplitudes for each time window were analyzed using a global ANOVA with the between-subject factor Group (experimental, control), and the within-subject factors Grammatical condition (grammatical, ungrammatical), and

Laterality (left, central, right). For convenience, we used the abbreviation Gro instead of Group, Gra as the abbreviation of Grammaticality, and L as the abbreviation of Laterality in the following content. Significance levels of the F ratios were adjusted with Greenhouse-Geisser correction. As we only cared about whether participants in the experimental group could elicit the P600 component, any global ANOVA that yielded any significant ($p < 0.05$) interaction involving the factor Grammatical condition and Group was followed up with the step-down ANOVAs to clarify the nature of the interaction.

RESULTS

Oxford Placement Test Results and Self-Rating Results

Mean age of first English classroom exposure, English self-rating scores, and OPT scores are presented in **Table 2**. The t-test results showed that there were no significant differences between the two groups in any of the proficiency measures ($ps > 0.05$), indicating that they were well matched.

Behavioral Results

The behavioral accuracy of the two groups in this transfer effect test is presented in **Table 3**.

Paired-samples t-test showed that there was no significant difference between the two groups ($p > 0.1$), indicating that they did not differ in sentence comprehension.

Event-Related Potentials Results

The global ANOVAs including the between-subject factor Group (experimental, control), the within-subjects factors Grammatical condition (grammatical, ungrammatical), and Laterality (left, central, right) was conducted, respectively, in the 300–500 ms and 500–1,000 ms time window. Results are listed in **Table 4**.

TABLE 2 | Means age of first English classroom exposure (years), English self-ratings and OPT scores (standard deviations in parentheses).

Group	EG	CG
Age of first English classroom exposure	12.07 (0.83)	12.21 (1.71)
Listening	3.07 (1.14)	2.85 (1.09)
Speaking	3.07 (0.62)	2.85 (0.66)
Reading	2.93 (1.07)	2.57 (0.65)
Writing	2.61 (1.56)	2.57 (1.08)
OPT	38.04 (3.53)	38.00 (2.89)

EG, experimental group; CG, control group; OPT, Oxford Placement Test.

TABLE 3 | Mean accuracy (%) and standard deviations (SD) for sentence comprehension.

Group	Accuracy	SD
Experimental group	83.78	3.97
Control group	83.70	2.91

TABLE 4 | Summary of global ANOVA for the two groups.

Source	Df	*F* value	
		300–500 ms	500–1,000 ms
Gra	1.24	4.23	3.65
Gra × Gro	1.24	3.88	**5.00***
Gra × Gro × L	2.48	1.12	0.49
Gra × L	2.48	1.43	0.04

Gra, grammaticality; Gro, group; L, laterality.
$^*p < 0.05$.

300–500 Time Window

Results in this time window showed that the main effects and the interactions were not significant ($ps > 0.05$).

500–1,000 Time Window

In the time window of 500–1,000 ms, results of the global ANOVA showed that the interaction between Grammatical condition and Group [$F(1, 24) = 5.00$, $p = 0.035$, $\eta_p^2 = 0.17$] was significant. Other main effects and interactions were not significant ($ps > 0.05$). In order to clarify the nature of the interaction between Group and Grammatical condition, further analysis by Group was conducted.

For the experimental group (EG), the results of the paired-samples t-test between the grammatical condition and ungrammatical condition showed that difference between these two conditions was significant [$t(12) = 2.814$, $p = 0.016$, $d = 0.85$], that was the ungrammatical condition elicited a more positive component than that of the grammatical condition (0.46 μV vs. −0.66 μV). According to its distribution, it should be termed as P600 (see **Figure 1**), indicating obvious transfer effect. However, for the control group (CG), the results of the paired-samples t-test revealed no significant difference between the grammatical condition and the ungrammatical condition ($ps > 0.05$) (see **Figure 2**).

DISCUSSION

The present study investigated the transfer effect in late L2 learners to explore whether the relationship between input and the entrenched representation was extended to the broader category of the subject-verb agreement structure, and whether these late L2 learners still show plasticity in syntactic acquisition. Results showed that syntactic violations which were evident in the subject-verb agreement structure with plural head nouns elicited a significant P600 in the experimental group (EG) but not in the control group (CG). The findings indicated a significant syntactic transfer effect, which indicated the important role of input in L2 syntactic acquisition of the late L2 learners.

The Role of Input, L2 Syntactic Transfer, and Late L2 Learners

The experimental group indicated a significant P600 to the violations of the subject-verb agreement structure with plural nouns, compared with the grammatical counterpart,

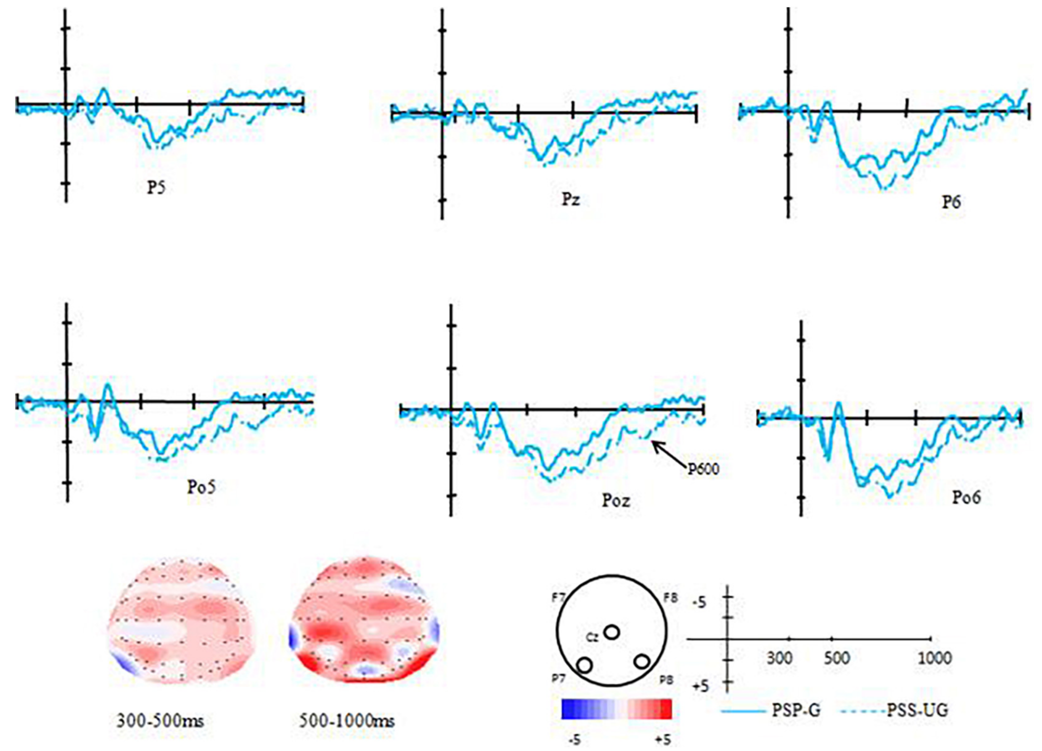


FIGURE 1 | Voltage maps and the grand average ERPs of the EG for the difference between correct and violated sentences in the transfer-effect test.

while the control group did not. The P600, as the index of grammaticalization (White et al., 2012) and indicator of entrenched representation (Deng et al., 2015), can be regarded as the instantiation of grammatical knowledge into the learners' online language processing system (Osterhout et al., 2008) and increased accessibility originated from the entrenched representation (Bybee, 2006). Participants from the EG benefited from the specific trainings with subject-verb agreement structures with singular head nouns and had their corresponding representation entrenched, which might give rise to the elicitation of P600. The results were in accordance with the related studies on input in L2 field (Morgan-Short et al., 2012a,b; Montrul et al., 2013; Deng et al., 2015). Additionally, the neural processes underlying L2 (morpho) syntax especially for late L2 learners are thought to be complicated. Some studies showed the biphasic components of LAN and P600 (Deutsch and Bentin, 2001; Friederici et al., 2002; Hahne et al., 2006; Steinhauer et al., 2009; Zawiszewski and Friederici, 2009), while others showed only the P600 (Bowden et al., 2007). The only P600 component, without the LANs component, might indicate that the EG participants were not proficient enough with the subject-verb agreement structure to reach the automatic processing, according to White et al. (2012).

Due to insufficient input with the specific subject-verb agreement structure, participants of both the EG and CG showed no sensibility to violations of this structure in the pretest of Deng et al. (2015). Two intensive training sessions with specific subject-verb agreement structures made participants in the experimental group relatively proficient with this specific structure (White et al., 2012) and thus made its corresponding representations entrenched (Bybee, 2006). Though the participants were not provided with the subject-verb agreement structures with plural head nouns in the training sessions (Deng et al., 2015), the EG still showed sensitivity to these violations, that is, syntactic transfer effect, which might be due to the entrenched representation. The P600 component of this syntactic transfer effect might be the indication of the role of input not only in entrenching the corresponding the one-to-one corresponding representation (Deng et al., 2015; Deng and Chen, 2019) but also in entrenching the broader category of the syntactic representation.

Questions remains in the L2 field regarding what might affect the L2 syntactic acquisition: the AoA (age of acquisition, thus the distinction the early learners and late learners), the level of proficiency, input and other variables as learning strategy, etc. (Lenneberg, 1967; De Haene et al., 1997; Chee et al., 2001; Reiterer et al., 2005, 2009; White et al., 2012). Some researchers suggested that early learners outperformed late learners due to the age of onset (Lenneberg, 1967; De Haene et al., 1997). Some insisted that it might be the proficiency level that decided the acquisitions (Steinhauer et al., 2009). The fact that the participants of the present study who were later learners with relatively low general proficiency showed a significant transfer effect cannot be simply attributed to the factor of AOA or the general proficiency level. The relationship between input and transfer effect of the present study might partly indicate that input might be an important

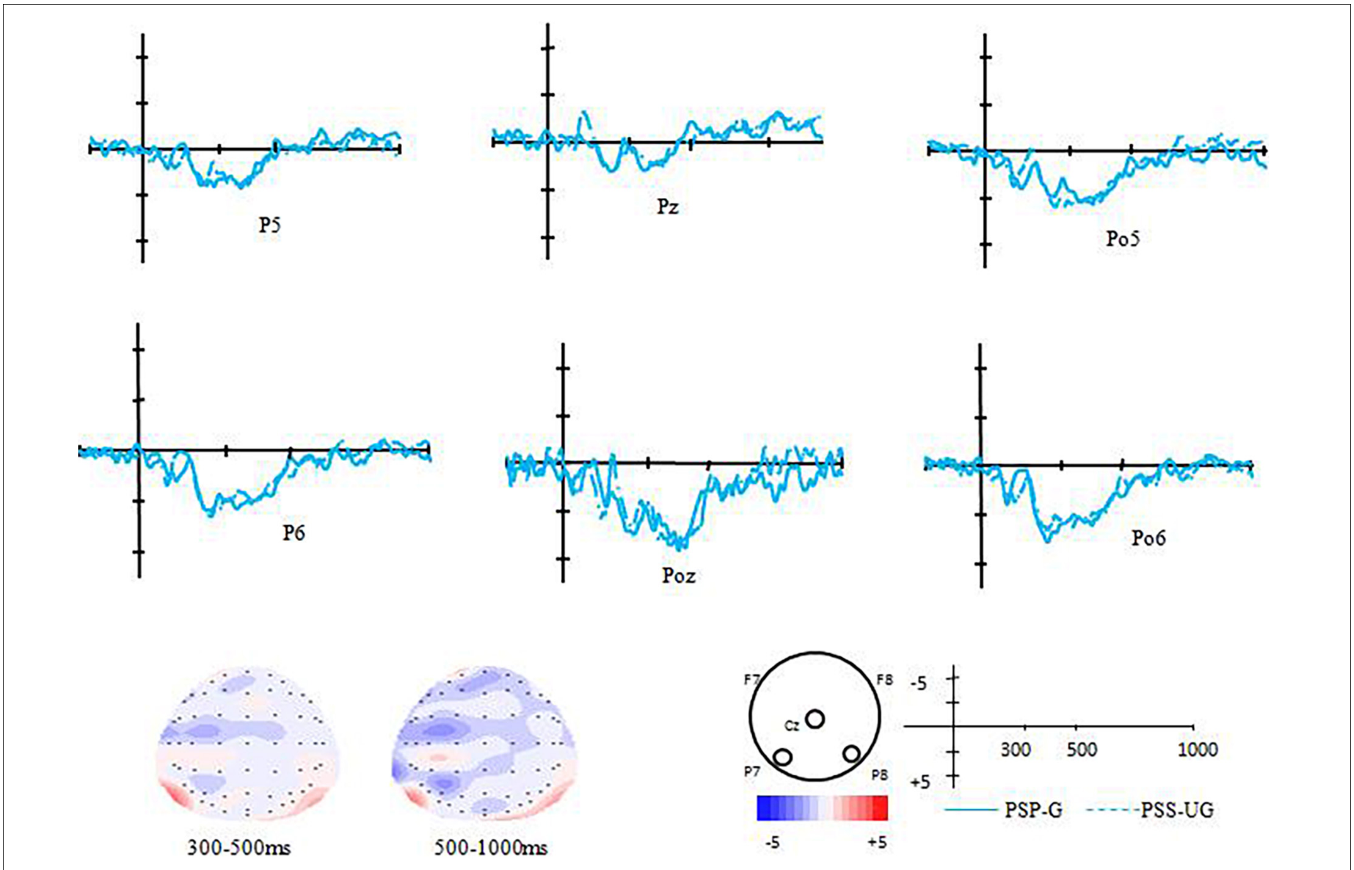


FIGURE 2 | Voltage maps and the grand average ERPs of the CG for the difference between correct and violated sentences in the transfer-effect test.

factor behind AOA and proficiency. Distinction between early and late leaners lies not only in age but also the possible input or exposure. Similarly, proficiency is the outcome variable, as claimed by Reiterer et al. (2009), and it is a complex variable that functions as an umbrella term and subsumes many of the other factors such as input training (Reiterer et al., 2009). Input training might improve proficiency level with the specific structure trained, which is thought to be the possible important factor affecting the processing performance (White et al., 2012; Deng et al., 2015). Though the present study, together with other studies (Morgan-Short et al., 2012a,b; Montrul et al., 2013; Deng et al., 2015), explored the role of input in L2, it still calls for more efforts in this L2 field.

Previous studies exploring the transfer effect mainly focused on the inter-language influence, that is, from L1 to L2 or L2 to L1 to explore the role of background language or the directionality (Kellerman, 1979; Sharwood Smith, 1986; Pavlenko and Jarvis, 2002; Brown and Gullberg, 2008), while very few studies focused on the intralinguistic syntactic transfer effect. The elicitation of P600 in EG, as the indicator of the syntactic transfer effect, not only indicated the role of input in intralinguistic syntactic transfer effect but also made a relatively small step in extending the studies on the role of input in L2.

In short, these late L2 learners of the present study, who indicated syntactic transfer effect, might give some enlightenment in the L2 field: First, input factor plays an important role in L2 one-to-one syntactic entrenchment and also syntactic transfer effect, which is in accordance with the studies both in L1 and L2 (Kaschak and Glenberg, 2004; Bybee, 2006; Dabrowska, 2008a,b; Street and Dąbrowska, 2010; Morgan-Short et al., 2012a,b). Even for late L2 learners, input training still matters. Second, the variables in the L2 syntactic field are very complicated and many factors intertwine with each other. Maybe in the future experimental techniques could be used to disentangle the complicated relationships among the variables to acquire relatively pure results.

The Transfer Effect and Its Probable Occurrence Mechanism

In Deng et al. (2015), no ERP components were observed in either EG or GG upon syntactic violations in the pre-test. Then a significant difference was observed between the EG and the CG in an immediately post-test after input training, with a P600 elicited in the EG but not in the CG. This revealed the importance of input training in L2 representation entrenchment, and also that the entrenched representation effect can last a relatively long time (Deng and Chen, 2019). However, questions still remain: Is the entrenched representation only limited to the structure trained, or can it be transferred to the similar structure? What might give an explanation for this transfer effect?

In the present transfer study, a significant P600 was still elicited in the EG but not in the CG, indicating the obvious transfer effect to the subject-verb agreement with plural head nouns. We attempt to explain our results within usage-based theory and similarity theory.

According to usage-based theory, input or usage strengthens the memory representations, making them easier to access (Goldberg, 1995; Croft, 2000; Bybee, 2006; Schmid, 2007). Every exemplar of a language use or input encountered by a speaker has an effect on cognitive representation in memory. As claimed by Bybee (2006), the linguistic memories represented as exemplars can undergo considerable reorganization. Exemplars of phrases or sentences that are similar on different dimensions are grouped together in cognitive representation. Similar but not identical exemplars are stored and represented to constitute a cluster or categories. Eventually, as long as the cognitive representation has been entrenched to a certain degree, it turns out to be highly effective, accessible, and autonomous (Goldberg, 1995; Croft, 2000; Bybee, 2006). Increase in input or exposure leads to representation entrenchment, accessibility of a preexisting representation, learning or acquisition of a new representation, or reorganization or modifications to existing representations (Luka and Choi, 2012). Thus, speakers might improve their performance in comprehending such exemplar categories or cluster based on the entrenched representation, which might be in accord with the previous studies in that appropriate representations increase positive transfer (Luchins, 1942; Chen and Daehler, 1989; Singley and Anderson, 1989).

As to the participants here, although they belong to the participants with relatively low proficiency, they have systematically learned the grammar about the subject-verb agreement structures explicitly in classroom teaching. This kind of input is not sufficient enough for them to form deep representation, which is evident in the lack of P600 in the pretest in the study of Deng et al. (2015). Then, the EG received the specific input training on the subject-verb agreement structure with singular head nouns, which entrenched their relatively shallow representation. The P600 of the EG in the posttest provided the evidence. This kind of non-local subject-verb agreement structure with singular head noun provided in the training sessions, and the materials about the non-local subject-verb agreement structure with plural head noun, belong to this same grammatical category or cluster. According to the usage-based theory (Goldberg, 1995; Croft, 2000; Langacker, 2000; Bybee, 2006; Haskell et al., 2010), exemplars provided in the input training sessions, together with the previous subject-verb agreement structures that the participants learned in the classroom, are grouped together in cognitive representation, where linguistic memories represented as exemplars can undergo considerable reorganization or reanalysis. That is, the syntactic representation about subject-verb agreement expressed in similar or different exemplars together might have been entrenched, through the two sessions of input training provided (Goldberg, 1995; Langacker, 2000; Bybee, 2006). Then, the entrenched representation gave rise to autonomy and efficacy in comprehending the subject-verb agreement structure with plural head nouns, where the P600 to the agreement violations in the EG might provide evidence.

The second explanation for this occurrence mechanism of the transfer effect might be the high similarity between the materials tested in the present study and those trained in the study of Deng et al. (2015). According to Thorndike' s classical view on the transfer effect (Thorndike and Woodworth, 1901), the likelihood of the occurrence of the transfer effect is directly related to the similarity between the situations trained and the situations tested. As claimed by Thorndike (1913), transfer could take place in cases where common elements were shared between the source and the target. Specifically, the materials' structure of the present study contained a plural countable noun modified by a prepositional phrase (PP), the verb "were" or "was" as the predicate that either agreed or disagreed with the plural subject noun in number, and other sentence constituents. For example, "The girls of the family were (was) very beautiful and polite." That structure of Deng et al. (2015) included a single countable noun modified by a prepositional phrase (PP), the verb "was" or "were" as the predicate that either agreed or disagreed with the singular subject noun in number, and other sentence constituents. For example, "The price of the car was (were) very high." They were quite similar in structure expressions. Then, according to the similarity theory, the greater the similarity between the learning task and the test task, the higher the possibility that the transfer could take place (Thorndike, 1913). The only difference between the materials of the present study and the previous study of Deng et al. (2015) was whether the head noun was plural or singular. Therefore, it might be impossible to tease apart the effect of the similarity between the materials on the transfer performance.

In brief, the present study not only provided the evidence for the transfer effect but also tried to give possible explanation about how this transfer effect happened both from the perspective of the usage-based theory (Bybee, 2006) and the similarity theory (Thorndike and Woodworth, 1901). Also, it will be interesting to design some experiments to explore what decides the occurrence of the transfer effect in the future.

The Essence of the L2 Grammar System and Its Broad Significance

In L2 syntactic acquisition, the question of whether the current attainment state of the late L2 learners is the ultimate attainment is still open, as late L2 learners began their L2 learning late in life. This question concerns the possibility for development of late L2 learners. One of the prevalent opinions is that maturational changes lead to discontinuity in the neurocognitive architecture in language development. As a consequence, native-like outcomes in language acquisition are argued to be limited in a biologically circumscribed period of time when language acquisition needs to begin, usually taken to end in late childhood or around puberty (Singleton and Ryan, 2004; DeKeyser, 2012, 2013). That is to say, according to this critical period hypothesis, it might be the ultimate attainment or the end-state of the late L2 learners who began to learn the L2 after puberty, indicating that L2 syntax of the late L2 learners is considered to have a developmental endpoint. Some studies indicated that even advanced late L2 learners occasionally showed slip-ups or even protracted variability in subject-verb agreement, tense, and gender marking (Jiang, 2004, 2007; Silva and Clahsen, 2008), which might be attributed to age constraints. In contrast to this opinion, some researchers believe that it is the proficiency or input factor that affects the acquisition of the L2. As the proficiency or input improves, their difficulty in syntactic comprehension can be relieved (Steinhauer et al., 2009; White et al., 2012). For instance, Deng et al. (2015) conducted an experiment to explore the role of the input in L2 syntactic processing (Deng et al., 2015). The results indicated that the input plays an important role in improving structure-specific proficiency and entrenching syntactic representation. And in Deng and Chen (2019), the results showed that the entrenched representation can even be maintained in a relatively long period. Both of these studies, coupled with the transfer effect of the present study, show the plasticity and dynamic nature of the L2 grammar system. Even for late learners, their grammar system can be in constant grammaticalization of syntactic rules through language exposure.

The transfer effect reported in the present study provides evidence for the malleability of late L2 learners' grammatical system, which is in line with the results of L1 studies (Kaschak and Glenberg, 2004; Kaschak, 2006; Kaschak and Borreggine, 2008; Wells et al., 2009; Haskell et al., 2010; Kaschak et al., 2011; Luka and Choi, 2012). The malleability of the adult language in the aspect of syntax development is robustly evident in a set of phenomena broadly called structural priming, which shows that exposure to a given syntactic construction can affect the subsequent processing of the same or related constructions (Bock, 1986; Branigan et al., 1999, 2000; Pickering and Ferreira, 2008). According to these theories, learning should be reserved for more enduring changes. These observations suggest that incremental adjustments to the language processing system occur continuously and may even extend to acquisition of novel syntactic structures. For instance, Kaschak and Glenberg (2004) carried out a series of experiments to explore how adults learned to comprehend a new syntactic construction in their native language. In experiments 1 and 2, the adults quickly learned to comprehend the new "need" construction and generalized it to new verbs. The transfer effect was evident in which the participants learned to comprehend a novel syntactic pattern from only a few exposures buried within about 10–12 min of conversation. Participants who learned to comprehend the "need" construction were able to generalize this learning to processing the same construction with a new verb. The results indicated that the mechanism that functions in child language acquisition may play an important role in adults' continued ability to learn new constructions in their native language (Seidenberg and MacDonald, 1999). In other words, for adults, syntactic learning is a continuous, dynamic process throughout the whole life span. The results of the present study are consistent with that of Kaschak and Glenberg (2004). According to our results, for late L2 learners, the current state of the grammatical knowledge in memory is not the end-state of the attainment. Their syntactic representation can be entrenched dynamically with the joint forces of input or exposure, cognition, and their interaction.

In short, the transfer effect in the present study indicated that even for late L2 learners, the malleability and learnability of syntax is possible. The nature of grammatical knowledge in memory is dynamic. For late L2 learners, their learning grammatical and morphological knowledge can bring changes and might continue throughout the lifetime.

What's more, the results of the transfer effect might provide some pedagogical implications as to the training paradigm. First, intensive training with specific syntactic structure is like a structure-oriented approach that can contribute to proceduralizing the known syntactic rules. This kind of approach might help participants rediscover already known language in their direct contact with new content input. Second, the training method, the self-paced reading, feedback-facilitated method, not only closely resembled natural reading but also involves a higher level of linguistic and cognitive processes such as inference-making (Just et al., 1982), as the participants have to summon their existing syntactic knowledge and reading strategies to correctly comprehend the sentences in this moving window condition. Third, this kind of task focuses the participants' attention on each word they read, which might indicate the important role of attention in L2 syntactic acquisition. As claimed by Devos (2016), "to encourage FL practice and simultaneously mitigate fossilization, specific attention needs to be paid to the language. The self-paced reading paradigm," the structure-oriented approach with specific subject-verb agreement structures, together with the grammatical judgment task after reading, specifically directed the participants' attention to the head noun and its corresponding predicate. Maybe, in the future, when the teachers design the tasks to help students acquire the syntax, they might take into the consideration such elements as cognition, attention, etc.

Nevertheless, there are important limitations to the present study. First, our studies concerned the transfer effect of the same subject-verb agreement structure with the only difference in head noun. To gain a broader picture of the late L2 learners' ability in syntactic acquisition, similar studies should be done with a range of other transfer conditions. For example, could the participants show sensitivity to other subject-verb agreement expressions after being trained with the materials in the present study? Or was the transfer effect specific to this experimental context? Without a definitive answer, the conclusion that can be drawn from these data is that training with specific structures can lead to easier comprehension of the similar structure. Second, because there are high-level similarities between the trained materials and those of the present study, it might be difficult to differentiate what leads to the present results. Is the transfer effect attributed to similarity or to entrenched representation or to both?

The present experiment takes an important step toward understanding how late L2 learners learn to process similar structure. Of course, our results might only scratch the surface about the interaction between learning, memory, representation, language acquisition, and language processing. More research is needed to shed light on these interactions with the late L2 learners to clarify the internal language acquisition mechanism.

For example, it might be interesting and meaningful to explore, how long does this kind of syntactic transfer effect last and what factors are likely to affect the sustainability? The present study makes an effort to shed some light on the complex L2 syntax acquisition abilities of late L2 learners.

CONCLUSION

In conclusion, the present study set to explore such questions as: Could the relationship between the input and syntactic representation be one-to-one correspondence? Or could it be extended to the broader syntactic category? Could these late L2 learners still show plasticity in L2 syntactic acquisition? The current results of the present study suggest that linguistic input training contributes to the transfer effect, indicating the malleability and dynamic nature of the L2 syntactic acquisition. Input plays an important role not only in one-to-one L2 syntactic representation entrenchment but also in the entrenchment of a broader category of the syntactic representation. Last but not least, even for the late L2 learners, learning and plasticity in L2 syntactic aspect can continue. The results of the present study provide not only theoretical implications on the learnability of late L2 learners but also the pedagogical implications for the teachers on the task designs.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Beijing Normal university. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TD and QF contributed to the design and manuscript drafting. DD provided help in data collecting and processing. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by Social Science Foundation Project of Hebei Province (HB18JY044) to QF.

REFERENCES

Barlow, M., and Kemmer, S. (eds) (2000). *Usage-Based Models of Language.* Stanford, CA: CSLI Publications.

Bertram, R., and Kuperman, V. (2019). The English disease in Finnish compound processing: backward transfer effects in Finnish–English bilinguals. *Bilingualism* 23, 1–12. doi: 10.1017/S1366728919000312

Bock, J. K. (1986). Syntactic persistence in language production. *Cogn. Psychol.* 18, 355–387. doi: 10.1016/0010-0285(86)90004-6

Bowden, H. W., Sanz, C., Steinhauer, K., and Ullman, M. (2007). An ERP study of proficiency in second language. *J. Cogn. Neurosci.* E61:170.

Bowden, H. W., Steinhauer, K., Sanz, C., and Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia* 51, 2492–2511. doi: 10.1016/j.neuropsychologia.2013.09.004

Boxell, O., and Felser, C. (2017). Sensitivity to parasitic gaps inside subject islands in native and non-native sentence processing. *Bilingualism* 20, 494–511. doi: 10.1017/S1366728915000942

Branigan, H. P., Pickering, M. J., and Cleland, A. A. (1999). Syntactic priming in written production: evidence for rapid decay. *Psychon. Bull. Rev.* 6, 635–640. doi: 10.3758/BF03212972

Branigan, H. P., Pickering, M. J., and Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition* 75, 13–25. doi: 10.1016/S0010-0277(99)00081-5

Brooks, P., and Sekerina, I. (2006). Shortcuts to quantifer interpretation in children and adults. *Lang. Acquis.* 13, 177–206. doi: 10.1207/s15327817la1303_2

Brown, A., and Gullberg, M. (2008). Bidirectional crosslinguistic influence in l1-l2 encoding of manner in speech and gesture. *Stud. Sec. Lang. Acquis.* 30, 225–251. doi: 10.1017/S0272263108080327

Bybee, J. (2006). From usage to grammar: the mind's response to repetition. *Language* 82, 711–733.

Chee, M. W., Hon, N., Lee, H. L., and Soon, C. S. (2001). Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. *NeuroImage* 13, 1155–1163. doi: 10.1016/S1053-8119(01)91859-2

Chen, L., Shu, H., Liu, Y., Zhao, J., and Li, P. (2007). ERP signatures of subject–verb agreement in L2 learning. *Bilingualism* 10, 161–174. doi: 10.1017/S136672890700291X

Chen, Z., and Daehler, M. W. (1989). Positive and negative transfer in analogical problem solving by 6-year-old children. *Cogn. Dev.* 4, 327–344. doi: 10.1016/S0885-2014(89)90031-2

Chipere, N. (2001). Native speaker variations in syntactic competence: implications for first language teaching. *Lang. Aware.* 10, 107–124. doi: 10.1080/09658410108667029

Croft, W. (2000). *Explaining Language Change: An Evolutionary Approach*. Harlow: Longman.

Cunnings, I. (2017). Interference in native and non-native sentence processing. *Bilingualism* 20, 712–721. doi: 10.1017/S1366728916001243

Cunnings, I., and Clahsen, H. (2007). The time-course of morphological constraints: evidence from eye-movements during reading. *Cognition* 104, 476–494. doi: 10.1016/j.cognition.2006.07.010

Dabrowska, E. (1997). The LAD goes to school: a cautionary tale for nativists. *Linguistics* 35, 735–766. doi: 10.1515/ling.1997.35.4.735

Dabrowska, E. (2008a). The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflections: an empirical test of usage-based approaches to morphology. *J. Mem. Lang.* 58, 931–951. doi: 10.1016/j.jml.2007.11.005

Dabrowska, E. (2008b). The later development of an early-emerging system: the curious case of the Polish genitive. *Linguistics* 46, 629–650. doi: 10.1515/LING.2008.021

De Haene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., et al. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport* 8, 3809–3815. doi: 10.1097/00001756-199712010-00030

DeKeyser, R. M. (2012). "Age effects in second language learning," in *Handbook of Second Language Acquisition*, eds S. Gass and A. Mackey (London: Routledge), 442–460.

DeKeyser, R. M. (2013). Age effects in second language learning: stepping stones toward better understanding. *Lang. Learn.* 63, 52–67.

Deng, T. P., and Chen, B. G. (2019). Input training matters in L2 syntactic representation entrenchment: evidence from a follow-up ERP study. *J. Psycholinguist. Res.* 48, 729–745. doi: 10.1007/s10936-019-09628-z

Deng, T. P., Dunlap, S., and Chen, B. G. (2017). Effects of input training on second language syntactic representation

entrenchment. *Int. J. Biling.* 21, 3–20. doi: 10.1177/1367006915589423
Deng, T., Zhou, H., Bi, H. Y., and Chen, B. (2015). Input-based structure-specific proficiency predicts the neural mechanism of adult L2 syntactic processing. *Brain Res.* 1610, 42–50. doi: 10.1016/j.brainres.2015.03.039
Deutsch, A., and Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement in Hebrew: evidence from ERPs and eye movements. *J. Mem. Lang.* 45, 200–224. doi: 10.1006/jmla.2000.2768
Devos, N. J. (2016). *Peer Interactions in New Content and Language Integrated Settings*. Cham: Springer International Publishing.
Dussias, P. E., and Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism* 10, 101–116. doi: 10.1017/S1366728906002847
Ellis, N. (2002). Frequency effects in language processing. *Stud. Second Lang. Acquis.* 24, 143–188. doi: 10.1017/S0272263102002024
Ellis, N., and Collins, L. (2009). Input and second language acquisition: the roles of frequency, form, and function introduction to the special issue. *Mod. Lang. J.* 93, 329–335. doi: 10.1111/j.1540-4781.2009.00893.x
Felser, C., and Drummer, J.-D. (2017). Sensitivity to crossover constraints during native and non-native pronoun resolution. *J. Psycholinguist. Res.* 46, 771–789. doi: 10.1007/s10936-016-9465-8
Felser, C., and Roberts, L. (2007). Processing wh-dependencies in a second language: a cross-modal priming study. *Second Lang. Res.* 23, 9–36. doi: 10.1177/0267658307071600
Foote, R. (2011). Integrated knowledge of agreement in early and late English–Spanish bilinguals. *Appl. Psycholinguist.* 32, 187–220. doi: 10.1017/S0142716410000342
Foucart, A., and Frenck-Mestre, C. (2012). Can late l2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *J. Mem. Lang.* 66, 226–248. doi: 10.1016/j.jml.2011.07.007
Friederici, A. D., Steinhauer, K., and Pfeifer, E. (2002). Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 529–534. doi: 10.1073/pnas.012611199
Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago, IL: University of Chicago Press.
Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *J. Psycholinguist. Res.* 30, 251–266. doi: 10.1023/A:1010490917575
Hahne, A., Mueller, J. L., and Clahsen, H. (2006). Morphological processing in a second language: behavioral and event-related brain potential evidence for storage and decomposition. *J. Cogn. Neurosci.* 18, 121–134. doi: 10.1162/089892906775250067
Haskell, T. R., Thornton, R., and MacDonald, M. C. (2010). Experience and grammatical agreement: statistical learning shapes number agreement production. *Cognition* 114, 151–164. doi: 10.1016/j.cognition.2009.08.017
Havas, V., Todriguez-Fornells, V., and Clahsen, H. (2012). Brain potentials for derivational morphology: an ERP study of deadjectival nominalizations in Spanish. *Brain Lang.* 120, 332–344. doi: 10.1016/j.bandl.2011.10.008
Hopp, H. (2013a). Grammatical gender in adult L2 acquisition: relations between lexical and syntactic variability. *Second Lang. Res.* 29, 33–56. doi: 10.1177/0267658312461803
Hopp, H. (2013b). The development of L2 morphology. *Second Lang. Res.* 29, 3–6. doi: 10.1177/0267658312465304
Hopp, H. (2015). Individual differences in the second language processing of object–subject ambiguities. *Appl. Psycholinguist.* 36, 129–173. doi: 10.1017/S0142716413000180
Jiang, N. (2004). Morphological insensitivity in second language processing. *Appl. Psycholinguist.* 25, 603–634. doi: 10.1017/S0142716404001298
Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Lang. Learn.* 57, 1–33. doi: 10.1111/j.1467-9922.2007.00397.x
Juffs, A., and Harrington, M. (1995). Parsing effects in second language sentence processing. *Stud. Second Lang. Acquis.* 17, 483–516. doi: 10.1017/S027226310001442X
Just, M. A., Carpenter, P. A., and Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *J. Exp. Psychol. Gen.* 111, 228–238. doi: 10.1037//0096-3445.111.2.228
Kaschak, M. P. (2006). What this construction needs is generalized. *Mem. Cogn.* 34, 368–379. doi: 10.3758/BF03193414
Kaschak, M. P., and Borreggine, K. L. (2008). Is long-term structural priming affected by patterns of experience with individual verbs? *J. Mem. Lang.* 58, 862–878. doi: 10.1016/j.jml.2006.12.002
Kaschak, M. P., and Glenberg, A. M. (2004). This construction needs learned. *J. Exp. Psychol. Gen.* 133, 450–467. doi: 10.1037/0096-3445.133.3.450
Kaschak, M. P., Kutta, T. J., and Schatschneider, C. (2011). Long-term cumulative structural priming persists for (at least) one week. *Mem. Cogn.* 39, 381–388. doi: 10.3758/s13421-010-0042-3
Kayne, R. S. (2000). *Parameters and Universals*. Oxford: Oxford University Press.
Kellerman, E. (1979). Transfer and non-transfer: where we are now. *Stud. Second Lang. Acquis.* 2, 37–57. doi: 10.1017/S0272263100000942
Langacker, R. W. (2000). "A dynamic usage-based model," in *Usage-Based Models of Language*, eds M. Barlow and S. Kemmer (Stanford, CA: CSLI Publications), 1–63.
Lenneberg, E. (1967). *Biological Foundations of Language*. New York, NY: Wiley.
Li, P., Bates, E., and MacWhinney, B. (1993). Processing a language without inflections: a reaction time study of sentence interpretation in Chinese. *J. Mem. Lang.* 32, 169–192. doi: 10.1006/jmla.1993.1010
Li, P., Jin, Z., and Tan, L. (2004). Neural representations of nouns and verbs in Chinese: an fMRI study. *NeuroImage* 21, 1533–1541. doi: 10.1016/j.neuroimage.2003.10.044
Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychol. Monogr.* 54, i–95. doi: 10.1037/h0093502
Luka, B. J., and Barsalou, L. W. (2005). Structural facilitation: mere exposure effects for grammatical acceptability as evidence for syntactic priming in comprehension. *J. Mem. Lang.* 52, 436–459. doi: 10.1016/j.jml.2005.01.013
Luka, B., and Choi, H. (2012). Dynamic grammar in adults: incidental learning of natural syntactic structures extends over 48 hours. *J. Mem. Lang.* 66, 345–360. doi: 10.1016/j.jml.2011.11.001
Moll, T. M. (1999). A linguistic needs analysis for EFL at the university level. *Rev. Alicant. Estud. Ingl.* 12, 117–126. doi: 10.14198/raei.1999.12.09
Montrul, S., Israel de la, F., Davidson, J., and Foote, R. (2013). The role of experience in the acquisition and production of diminutives and gender in Spanish: evidence from L2 learners and heritage speakers. *Second Lang. Res.* 29, 87–118. doi: 10.1177/0267658312458268
Morgan-Short, K., Finger, I., Grey, S., and Ullman, M. T. (2012a). Second language processing shows increased native-like neural responses after months of no exposure. *PLoS One* 7:e32974. doi: 10.1371/journal.pone.0032974
Morgan-Short, K., Steinhauer, K., Sanz, C., and Ullman, M. T. (2012b). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *J. Cogn. Neurosci.* 24, 933–947. doi: 10.1162/jocn_a_00119
Morgan-Short, K., Sanz, C., Steinhauer, K., and Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: an event-related potential study. *Lang. Learn.* 60, 154–193. doi: 10.1111/j.1467-9922.2009.00554.x
Ojima, S., Nakata, H., and Kakigi, R. (2005). An ERP study on second language learning after childhood: effects of proficiency. *J. Cogn. Neurosci.* 17, 1212–1228. doi: 10.1162/0898929055002436
Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., et al. (2008). Second-language learning and changes in the brain. *J. Neurolinguist.* 21, 509–521. doi: 10.1016/j.jneuroling.2008.01.001
Pavlenko, A., and Jarvis, S. (2002). Bidirectional transfer. *Appl. Linguist.* 23, 190–214. doi: 10.1093/applin/23.2.190
Pickering, M. J., and Ferreira, V. S. (2008). Structural priming: a critical review. *Psychol. Bull.* 134, 427–459. doi: 10.1037/0033-2909.134.3.427
Reiterer, S., Berger, M. L., Hemmelmann, C., and Rappelsberger, P. (2005). Decreased EEG coherence between prefrontal electrodes: a correlate of high language proficiency? *Exp. Brain Res.* 163, 109–113. doi: 10.1007/s00221-005-2215-z
Reiterer, S., Pereda, E., and Bhattacharya, J. (2009). Measuring second language proficiency with EEG synchronization: how functional cortical networks and hemispheric involvement differ as a function of proficiency level in second language speakers. *Second Lang. Res.* 25, 77–106. doi: 10.1177/0267658308098997

Roberts, L., and Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Appl. Psycholinguist.* 32, 299–331. doi: 10.1017/S0142716410000421

Sabourin, L., and Stowe, L. A. (2008). Second language processing: when are first and second languages processed similarly? *Second Lang. Res.* 24, 397–430.

Schmid, H. J. (2007). "Entrenchment, salience and basic levels," in *The Oxford Handbook of Cognitive Linguistics*, eds D. Geeraerts and H. Cuyck-ens (Oxford: Oxford University Press), 117–138.

Seidenberg, M. S., and MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing. *Cogn. Sci.* 23, 569–588. doi: 10.1016/S0364-0213(99)00016-6

Sharwood Smith, M. (1986). "The competence/control model, crosslinguistic influence and the creation of new grammars," in *Crosslinguistic Influence in Second Language Acquisition*, eds E. Kellerman and M. Sharwood Smith (Oxford: Pergamon Press), 10–20.

Silva, R., and Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: evidence from masked priming experiments in English. *Bilingualism* 11, 245–260. doi: 10.1017/S1366728908003404

Singleton, D., and Ryan, L. (2004). *Language Acquisition: The Age Factor*. Clevedon: Multilingual Matters.

Singley, K., and Anderson, J. R. (1989). *The Transfer of Cognitive Skills*. Cambridge, MA: Harvard University Press.

Steiner, G. (2001). "Transfer of learning, cognitive psychology of," in *International Encyclopedia of the Social & Behavioral Sciences*, Vol. 11, eds N. J. Smelser and P. B. Baltes (Oxford: Pergamon Press), 15845–15851. doi: 10.1016/B0-08-043076-7/01481-9

Steinhauer, K., White, E. J., and Drury, J. E. (2009). Temporal dynamics of late second language acquisition: evidence from event-related brain potentials. *Second Lang. Res.* 25, 13–41. doi: 10.1177/0267658308098995

Street, J. A., and Dąbrowska, E. (2010). More individual differences in language attainment: how much do adult native speakers of English know about passives and quantifiers? *Lingua* 120, 2080–2094. doi: 10.1016/j.lingua.2010.01.004

Thorndike, E. L. (1913). *The Psychology of Learning, Educational Psychology*. New York, NY: Teachers College.

Thorndike, E. L., and Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychol. Rev.* 8, 247–261. doi: 10.1037/h0071280

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: the case of Dutch gender. *Bilingualism* 16, 86–110. doi: 10.1017/S1366728912000284

Van Hell, J. G., and Tokowicz, N. (2010). Event-related brain potentials and second language learning: syntactic processing in late L2 learners at different L2 proficiency levels. *Second Lang. Res.* 26, 43–74. doi: 10.1177/0267658309337637

Weber-Fox, C. M., and Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *J. Cogn. Neurosci.* 8, 231–256. doi: 10.1162/jocn.1996.8.3.231

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., and MacDonald, M. C. (2009). Experience and sentence processing: statistical learning and relative clause comprehension. *Cogn. Psychol.* 58, 250–271. doi: 10.1016/j.cogpsych.2008.08.002

White, E. J., Genesee, F., and Steinhauer, K. (2012). Brain responses before and after intensive second language learning: proficiency based changes and first language background effects in adult learners. *PLoS One* 7:e52318. doi: 10.1371/journal.pone.0052318

Zawiszewski, A., and Friederici, A. D. (2009). Processing canonical and non-canonical sentences in basque: the case of object-verb agreement as revealed by event-related brain potentials. *Brain Res.* 1284, 161–179. doi: 10.1016/j.brainres.2009.05.099

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Psychology

ORIGINAL RESEARCH
published: 07 January 2022
doi: 10.3389/fpsyg.2021.796700

Masked Translation Priming With Concreteness of Cross-Script Cognates in Visual Word Recognition by Chinese Learners of English: An ERP Study

Shifa Chen†, Tingting Fu†, Minghui Zhao†, Yuqing Zhang*, Yule Peng*, Lianrui Yang* and Xiaolan Gu

College of Foreign Languages, Ocean University of China, Qingdao, China

Edited by:
Yan Jing Wu,
Ningbo University, China

Reviewed by:
Xiaowei Zhao,
Emmanuel College, United States
Xiaofei Lu,
The Pennsylvania State University (PSU), United States

****Correspondence:***
Yuqing Zhang
zyq8995@stu.ouc.edu.cn
Yule Peng
1443905124@qq.com
Lianrui Yang
lryang@ouc.edu.cn

†These authors share first authorship

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *17 October 2021*
Accepted: *13 December 2021*
Published: *07 January 2022*

Citation:
Chen S, Fu T, Zhao M, Zhang Y, Peng Y, Yang L and Gu X (2022) Masked Translation Priming With Concreteness of Cross-Script Cognates in Visual Word Recognition by Chinese Learners of English: An ERP Study. Front. Psychol. 12:796700. doi: 10.3389/fpsyg.2021.796700

Translation equivalents for cognates in different script systems share the same meaning and phonological similarity but are different orthographically. Event-related potentials were recorded during the visual recognition of cross-script cognates and non-cognates together with concreteness factors while Chinese learners of English performed a lexical decision task with the masked translation priming paradigm in Experiment 1 (forward translation: L1–L2) and Experiment 2 (backward translation: L2–L1). N400 effect was found to be closely related to priming effects of cross-script cognate status and concreteness in Experiment 1; and in Experiment 2, N150 and N400 effects were related to priming effects of cross-script cognate status and concreteness, and greater priming effects of cross-script cognate status in cognates than in non-cognates for abstract words were found in the time window of 100–200 ms. Meanwhile, the asymmetry of translation directions was observed in smaller priming effects in forward translation than in backward translation in the time window of 100–200 ms for abstract cognates, and in larger priming effects in forward translation than in backward translation in the time window of 350–550 ms for each type of words. We discussed the roles of phonological activation and concreteness effects in view of the function of N150 and N400 components as well as the relevant models, mainly the Distributed Feature Model and Bilingual Interactive Activation (BIA+) model.

Keywords: priming effect of cross-script cognate status, priming effect of concreteness, translation asymmetry, N150, N400

INTRODUCTION

In the domain of psycholinguistic research on bilingualism, endeavors have been taken to answer the question of whether lexical representations from both languages are simultaneously activated during the processing of the word input. Two competing theories are proposed to account for this issue. The language selective hypothesis assumes that the lexical candidates from the given language are only limited to compete, which corresponds to the view of independent lexicons of two languages, while the language non-selective hypothesis, supporting the view of an integrated lexicon, claims that lexical representations from both languages are accessed simultaneously with respect to bilingual word recognition.

Although the agreement has not been reached on how lexicons of two languages are represented in bilingual memory, much of the previous research has supported a shared conceptual system (see Francis, 2005). Based on this assumption, several models have been proposed to account for the structure of bilingual memory. For instance, the Revised Hierarchical Model (RHM, Kroll and Stewart, 1994) assumes that the meaning of L2 words is accessed through their L1 translation equivalents, and with the improvement of second language proficiency, L2 words can directly be accessed without the assistance of L1, indicating that L1 and L2 shared a common conceptual system. Another model also addressing the issue of L1 and L2 representations is the Distributed Feature Model (DFM, de Groot, 1992; van Hell and de Groot, 1998) which believes that the conceptual representations are distributed in one common conceptual system as meaning elements/nodes are shared by words within the same language and across languages. In other words, translation equivalents from two languages share a certain amount of meaning components, depending on the degree of meaning overlap. In the DFM, concrete words are assumed to have a larger meaning overlap across languages than abstract words since abstract words are more context-dependent and have rather different interpretations in different contexts. In the same vein, cognate words (i.e., translation equivalents with a similar form) have more common conceptual components than non-cognate words. Although the models mentioned above can to some extent explain the empirical data obtained by using different paradigms and experimental tasks, they failed to give a detailed description of the processing of word identification, from the onset of a word to the time when it is accessed. Language retrieval models such as the Bilingual Interactive Activation (BIA+) model (Dijkstra and van Heuven, 2002) are implemented in the studies of bilingual memory to describe the processing of word identification. The BIA+ model is in support of the hypothesis that the bilingual lexicon is integrated and accessed in a non-selective manner. The original BIA (Dijkstra and van Heuven, 1998) model assumes that when a string is presented, activation spreads through four layers of connected nodes, from sub-lexical features to letters, words, and finally the language node. However, the BIA model is only concerned with the recognition of orthographic representations and has limitations in accounting for some empirical results. The BIA+ model, on the other hand, incorporates different levels of representations (i.e., phonological, orthographic, and semantic representations) and predicts a sequential activation of the three levels. Subsequently, other models emerge to deal with the issues related to the processing and representation of mental lexicon, such as the Shared Asymmetrical Model (SAM, Dong et al., 2005), the Sense Model (Finkbeiner et al., 2004), the Modified Hierarchical Model (MHM, Pavlenko, 2009), and DevLex-II model on simulating cross-language semantic priming effects (Zhao and Li, 2013). The fact that so many theoretical models are proposed to account for the processing and representation of mental lexicon leads to many experiments conducted to investigate variables that can modulate the recognition of L1 and L2 words, including word type (cognates and non-cognates), concreteness and translation direction (Ferré et al., 2017).

For the research investigating bilingual word recognition based on the BIA+ model, challenges are presented by cognates that have a certain degree of overlap between two languages in any of the three levels: phonology, orthography, and semantics. In several studies examining whether word type affected the reaction time, more robust priming effects have been observed in cognates than in non-cognates, suggesting a facilitation effect of the cognate status (Duñabeitia et al., 2010; Ferré et al., 2017). The DFM (de Groot, 1992; van Hell and de Groot, 1998) accounts for the cognate facilitation effect by assuming that there is a greater degree of meaning overlap in cognates than in non-cognates, which leads to larger priming effects between their corresponding translation equivalents for cognates than for non-cognates. Generally, when an L2 cognate is learned, it is easy to associate this L2 word to the conceptual representation that has already existed in the memory (van Hell and de Groot, 1998). However, even if cognate translation equivalents that are orthographically and phonologically similar have been selected as critical materials to explore cross-language words recognition, it is still troublesome to disentangle phonological status from orthographic similarity in the translation pairs of alphabetic languages. According to Kim and Davis (2003), due to the orthographic similarity between the same-script languages, there is often a competition between the prime and the target. Since phonological information can be clearly separated from orthographic similarity in the translation pairs of cross-script languages, it is easier to investigate the cognate effects by distinguishing the phonological similarity from orthographic information within the cross-script languages.

A few researchers have investigated cross-script cognate priming effects using words from languages that do not share orthographic identity in behavioral experiments (Zhou et al., 2010; Nakayama et al., 2012, 2013; Ando et al., 2014; Zhang et al., 2018) and event-related potential (ERP) studies (Hoshino et al., 2010; Ando et al., 2015). For example, in a lexical decision task and a naming task, Zhou et al. (2010) observed priming effects of homophones with Chinese–English bilinguals, supporting the non-selective mechanism in phonological representation, which is in line with the BIA+ model as there was no orthographic similarity between Chinese and English. Zhang et al. (2018) investigated Chinese–English cognates and found that there was no advantage for Chinese–English cognates in forward translation whereas only English–Chinese cognates produced facilitation effects in backward translation, which may be attributed to different mappings from orthography to phonology between English and Chinese. Nakayama et al. (2012) examined cognates and phonological similar non-cognates for Japanese–English bilinguals in masked phonological priming paradigm, and found that the priming effects of cognates, but not of phonological similarity, were influenced by target frequency and L2 proficiency. Similar results from ERP experiments also confirmed that phonological priming occurred prior to and independent of the influence of word frequency. Ando et al. (2015) investigated the cross-script phonological activation in Japanese–English bilinguals by recording both the ERP data and response data in a lexical decision task. They found a facilitation effect of Katakana primes to phonologically similar English target words, which indicated that there was a shared store of sublexical phonological representations by both Japanese and English, and the cross-script phonological priming effects were the consequence of the activation of the shared sublexical phonological representations. Therefore, it is important to disentangle the phonological factor from orthographic representation during visual word recognition within cross-script languages by exploring the processing of Chinese loan words and their English equivalents (i.e., Chinese–English cognates, like "幽默-humor") because they are phonologically and semantically similar, but orthographically different, and can be utilized in the investigation of phonological activation in word recognition. Nevertheless, whether the cognate effects disappear or not on the recognition of the targets based on ERP technology for Chinese loan words and their English equivalents remains unclear.

The factor of concreteness has been well-acknowledged in the previous research, either as an independent variable or as a control variable. It has been suggested that concrete words performed discordantly with abstract words in response latencies and N400 amplitudes in both monolingual and bilingual related studies (Zhang et al., 2006; Tolentino and Tokowicz, 2009; Huang et al., 2010; Barber et al., 2013; Palmer et al., 2013; Ferré et al., 2017). The facilitation effect of concreteness in bilingual word representation has been explained by the DFM (de Groot, 1992; van Hell and de Groot, 1998), which states that concrete words share more semantic components than abstract words. van Hell and de Groot (1998) employed a word association task and found that in both within- and between-language associations, cognates and concrete words were more often associated with their translations relative to non-cognates and abstract words. However, some researchers argued that concrete and abstract words shared equivalent concept overlap across languages in view of the similar priming effects observed in experiments, which discredits the claims of the DFM (Francis and Goldmann, 2011; Chen et al., 2014). One reason for the discrepancy in results may be that concreteness effects could only be observed within a certain range of stimulus-onset asynchronies (SOA). Research by Schoonbaert et al. (2009) examined two SOAs (250 and 100 ms) and their results showed that although the main effect of concreteness did not reach significance in both SOAs, the concrete words but not the abstract words produced a significant priming effect with the 100 ms SOA. Ferré et al. (2017) investigated cognate status and concreteness effects in two SOAs (50 and 100 ms), and found concreteness priming effects only in the longer SOA (100 ms). Thus, the concreteness priming effects are sensitive to SOA duration so that researchers need to carefully consider the factor of SOA when exploring the influence of concreteness, cognate status and their interaction effect on the processing of Chinese–English cross-script cognates.

The robustness of priming effects in the related studies often varies with translation directions. Faster responses were observed when L2 target words were preceded by their L1 translation equivalents (Gollan et al., 1997; Kim and Davis, 2003; Basnight-Brown and Altarriba, 2007), while evidence for L2–L1 priming effect in backward translation was not very consistent, with sometimes null priming effects (Schoonbaert et al., 2009; Dimitropoulou et al., 2011). Chen et al. (2020) found the asymmetrical priming effect between Chinese–English and English–Chinese translation directions with a larger N400 amplitude and a longer N400 latency in Chinese–English translation. The asymmetric effect can be explained by the RHM (Kroll and Stewart, 1994), in which the representations of L1 and L2 are qualitatively different, with L2 words less directly connected to the semantics. In contrast, DFM (de Groot, 1992; van Hell and de Groot, 1998) explains this result in a quantitative way in which L1 words have richer semantic representations than L2 words and thus can activate more features within a shorter time, resulting in stronger priming effects in forward translation (de Groot, 1992). The asymmetry is also predicted by the BIA+ model, which assumes that the speed of activation can be influenced by factors such as subjective frequency, and since L2 words have lower accessibilities than L1 words (L2 words are less frequently or recently used), activation spreads more slowly in L2 access than in L1 access. However, most of the previous studies about the asymmetry of translation directions mainly focused on non-cognate translation equivalents. Cognates with both the semantic and phonological overlap between cross-script languages may shed more light on the studies of translation directions.

In light of the research gaps identified above, previous studies mainly concentrated on languages with the same writing system, and it is difficult to clarify whether the facilitation effect of the cognate status is caused by phonological similarity or orthographic information. In addition, although some related studies have used cross-script languages to distinguish between phonological and orthographic promotion of cognates (Zhang et al., 2018), they do not distinguish between abstract cognates and concrete cognates. With high temporal resolution, ERP technology can provide us a more complete picture by showing the processing of the target words in real time, and has been employed to measure the cross-script phonological activation in Japanese–English bilinguals (Ando et al., 2015). Therefore, the present study is to use Chinese–English cross-script cognates with similar pronunciation and meaning but different orthographic information to examine the roles of phonology as well as concreteness effects with masked translation priming paradigm based on ERP technology in two experiments with different translation directions. It aims to examine whether there are translation priming effects for cross-script cognate status and concreteness in both forward and backward translation directions, whether cross-script phonological similarity and concreteness can elicit greater priming effects for cognates and concrete words than for non-cognates and abstract words in two translation directions, respectively, and whether there exists the translation asymmetry in terms of priming effect magnitudes between the two translation directions.

EXPERIMENT 1: L1–L2 (CHINESE–ENGLISH FORWARD TRANSLATION)

Methods

Experiment 1 examined the role of phonology as well as concreteness effects for Chinese learners of English with a lexical decision task in the masked translation priming paradigm in the L1–L2 translation direction.

Participants

Twenty-five Chinese–English bilinguals (14 females; mean age 20.68, $SD = 0.79$) were recruited from a public university in China to participate in the experiment. They were native Chinese speakers majoring in English and all of them had passed the Test for English Majors-Band 4 (TEM4). No immerse experience to learn English for all the participants and they have been learning English in the classroom environment for 10–12 years. A seven-point Likert scale assessment (1 for "quite poor," 7 for "highly proficient") was conducted to evaluate their L1 and L2 proficiency, and their self-reported rating for listening, speaking, reading and writing in L1 (Chinese) were 6.48 ($SD = 0.65$), 6.12 ($SD = 0.97$), 6.16 ($SD = 0.94$), 5.40 ($SD = 1.12$), and in L2 (English) with 4.84 ($SD = 0.99$), 4.68 ($SD = 1.07$), 5.44 ($SD = 0.92$), 4.36 ($SD = 0.76$), respectively. A paired-sample t-test showed that there were significant differences between L1 and L2 in listening, speaking, reading and writing [$ps < 0.001$, *Cohen's* $d(s) \geq 1.044$]. Therefore, our participants can be treated as unbalanced Chinese–English bilinguals. All of them had normal or corrected-to-normal vision and were right-handed without neurological disease.

Materials

Critical stimuli in Experiment 1 were 40 Chinese–English cognate pairs and 40 Chinese–English non-cognate pairs. All the Chinese–English cognate translation pairs were selected from *A Dictionary of Loan Words and Hybrid Words in Chinese* (Liu et al., 1984). Since there were no already existing common corresponding Chinese–English transliterated pairs, Chinese words were coined based on the pronunciation of the English translation equivalents. To make sure that the Chinese and English cognate pairs were indeed translation equivalents to each other, twenty students in English major who did not participate in the experiment were asked to translate them. Half of the students translated English into Chinese, while the other half translated words in the opposite direction. Only when 60% of the students gave the same translations for a given word were considered as translation equivalents of each other.

Meanwhile, another 20 Chinese learners of English from the same population were recruited to rate the concreteness and familiarity of English words on a five-point scale (1 for "quite abstract" and "very unfamiliar," and 5 for "quite concrete" and "very familiar"). Finally, these 80 Chinese–English pairs, which were categorized into four different sets: 20 cognate abstract word pairs, 20 cognate concrete word pairs, 20 non-cognate abstract word pairs, and 20 non-cognate concrete word pairs, were chosen for the present experiment. A paired-sample t-test was conducted to examine the variables of familiarity and concreteness. An independent-sample t-test was conducted to examine the length of English and the stroke of Chinese. In cognate and non-cognate trials, the 80 English targets were matched in subjective familiarity and concreteness [$ps \geq 0.347$, *Cohen's* $d(s) \leq 0.142$]. The length of the English targets and the number of strokes of the Chinese primes were matched between cognates and non-cognates [$ps \geq 0.681$, *Cohen's* $d(s) \leq 0.019$]. The concrete words and abstract words were matched in subjective familiarity ($p = 0.109$, *Cohen's* $d = 0.377$), length of the English targets and the number of strokes of the Chinese primes [$ps \geq 0.571$, *Cohen's* $d(s) \leq 0.058$].

As the present study attempts to find out whether phonological similarity and concreteness can affect the magnitudes of priming effects, another 80 words were selected as control primes to constitute the unrelated trials. The control primes were matched with the translation (related) primes in terms of the numbers of characters and strokes ($p = 0.680$, *Cohen's* $d = 0.065$) as well as concreteness ($p = 0.937$, *Cohen's* $d = 0.018$). In the experiment, the targets were presented under two conditions, the related condition in which the primes and the targets were translations of each other, and the unrelated condition in which the primes and the targets were not related in meaning. Additionally, to complete the yes or no response in the lexical decision, task additional 80 Chinese primes were paired with English pseudowords as targets. The English pseudowords were selected from Macquarie Online Test Interface[1] or generated by Wuggy (Keuleers and Brysbaert, 2010), and they were pronounceable sequences that followed the rules of English orthography. Their average length was matched with that of real English targets. Examples of the stimuli in the experiment were presented in **Table 1**.

Procedure

All the participants were tested in front of a computer in a sound attenuated room while the experimenter could monitor the process in another room. The experimental program was designed by E-Prime 3.0. In the experiment, participants needed to respond to 240 trials in total (160 with a word target, 80 with

[1] www.motif.org.au

TABLE 1 | Stimuli examples in Experiments 1 and 2.

Priming direction	Condition	Prime	Control	Target
L1–L2	Abstract cognates	逻辑	情感	Logic
	Concrete cognates	沙发	肌肉	Sofa
	Abstract non-cognates	心情	本能	Mood
	Concrete non-cognates	裤子	泥沙	Pants
L2–L1				
	Abstract cognates	Logic	Genre	逻辑
	Concrete cognates	Sofa	Coin	沙发
	Abstract non-cognates	Mood	Fate	心情
	Concrete non-cognates	Pants	Scarf	裤子

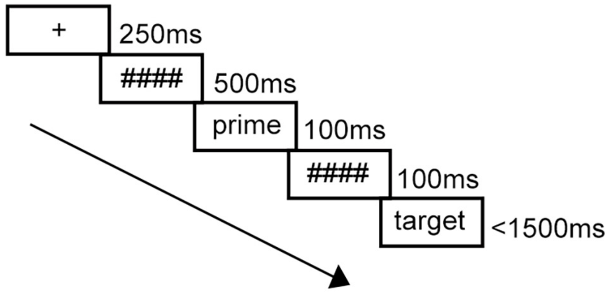


FIGURE 1 | Trial structure of the experiment.

a pseudoword target). Each item was presented at the center of the monitor. First, the fixation point "+" was displayed for 250 ms followed by a row of hash masks (#) for 500 ms. Next, a prime word appeared for 100 ms before it was replaced by the backward mask which lasted 100 ms. The length of pre- and post-masks for Chinese primes was matched with two hash masks for one Chinese character. Then the English target word was presented until the participant made a response, but for no more than 1,500 ms (see **Figure 1**). Participants were asked to decide whether the target words were real words or not. They could indicate their answers by pressing two keys on the keyboard, "J" or "F." The assignment of which key represented real words was counterbalanced across participants. There was a random interval of 300–500 ms after each trial. Before the experiment, 12 practice trials were constructed to help participants get familiar with the experimental process. Every target appeared twice, once in the related condition and once in the unrelated condition. The order of the two conditions was counterbalanced for a given target word and the presentation of trials was randomized.

Electroencephalogram Recording Procedure

The electroencephalogram (EEG) was recorded from 32 scalp sites by an electrodes cap following a revised standard International 10–20 system. Two electrodes were located next to the canthus to monitor horizontal activity, and vertical eye movement was monitored by another two electrodes next to the left eye up and down. All the data were re-referenced to the mean electric activity of the mastoids. The digitizing computer continuously sampled the EEG at a rate of 1,000 Hz. Scalp electrodes impedances were maintained below 5 kΩ and Bandpass was filtered between 0.01 and 100 Hz. The EEG was collected online and analyzed offline by Neuroscan Curry 8.

Data Analyses and Results

The entire data of three participants were excluded from analyses in forward translation due to higher error rates (over 30%) in behavioral and ERP analyses. The mean response times and error rates (E%) in Experiment 1 across each experimental condition were presented in **Table 2**.

Behavioral Analyses

The mean reaction times (RTs) and error rates (E%) were submitted to a 4 (word type: abstract cognates, concrete cognates, abstract non-cognates and concrete non-cognates) × 2 (relatedness: related and unrelated) design. Repeated-measures analysis of variance (ANOVA) by subjects and univariate ANOVA by items examined translation priming effects of each word type. The Greenhouse–Geisser correction was applied to all repeated-measures with more than one degree of freedom in the numerator in the present study.

The behavioral data on RT analysis showed that there was a significant main effect of word type [$F_1(3,63) = 18.304$, $p < 0.001$, $\eta_p^2 = 0.466$, $F_2(3,76) = 4.750$, $p = 0.004$, $\eta_p^2 = 0.158$], and of relatedness [$F_1(1,21) = 85.218$, $p < 0.001$, $\eta_p^2 = 0.802$, $F_2(1,76) = 296.164$, $p < 0.001$, $\eta_p^2 = 0.796$]. The interaction between word type and relatedness was (marginally) significant [$F_1(3,63) = 2.555$, $p = 0.067$, $\eta_p^2 = 0.108$, $F_2(3,76) = 3.192$, $p = 0.028$, $\eta_p^2 = 0.112$]. Simple effect comparisons showed that the relatedness effects were significant for all word types [$F_1s(1,21) \geq 21.141$, $ps \leq 0.001$, $\eta_p^2s \geq 0.502$, $F_2s(1,76) \geq 40.905$, $ps \leq 0.001$, $\eta_p^2s \geq 0.350$]. The response times of the unrelated condition were much longer than that of the related condition.

For the analysis of error rate, there was a significant main effect of word type [$F_1(3,63) = 18.563$, $p < 0.001$, $\eta_p^2 = 0.469$, $F_2(3,76) = 8.086$, $p < 0.001$, $\eta_p^2 = 0.242$], and of relatedness [$F_1(1,21) = 20.171$, $p < 0.001$, $\eta_p^2 = 0.490$, $F_2(1,76) = 29.949$, $p < 0.001$, $\eta_p^2 = 0.283$]. The interaction between word type and relatedness was also significant [$F_1(3,63) = 12.380$, $p < 0.001$, $\eta_p^2 = 0.371$, $F_2(3,76) = 11.102$, $p < 0.001$, $\eta_p^2 = 0.305$]. Simple effect comparisons revealed (marginally) significant differences for relatedness factor in all word types in analyses by subjects [$F_1s(1,21) \geq 4.074$, $ps \leq 0.057$, $\eta_p^2s \geq 0.162$] and by items [$F_2s(1,76) \geq 4.364$, $ps \leq 0.040$, $\eta_p^2s \geq 0.054$], except for the cognate concrete words which failed to reach significance by items [$F_2(1,76) = 1.017$, $p = 0.316$, $\eta_p^2 = 0.013$)].

Following the previous study (Ferré et al., 2017), the magnitudes of priming effects in the present study were calculated by subtracting the response times and error rates of the related conditions from the unrelated conditions for detecting

TABLE 2 | Mean RTs/error rates (E%) as a function of translation direction, cognate status and concreteness.

	Translation	Control	Priming effect
Experiment 1 forward translation (Chinese–English)			
Abstract cognates	614.1/4.1	701.1/18.0	87.1*/13.9
Concrete cognates	570.3/2.0	650.0/7.0	79.8*/5.0
Abstract non-cognates	604.3/2.0	657.5/6.1	53.3*/4.1
Concrete non-cognates	590.3/1.8	657.0/7.3	66.7*/5.5
Experiment 2 backward translation (English–Chinese)			
Abstract cognates	580.0/1.4	602.3/2.3	22.2*/0.9
Concrete cognates	557.6/1.1	589.4/2.3	31.8*/1.1
Abstract non-cognates	564.5/2.0	582.5/2.5	18.1*/0.5
Concrete non-cognates	552.9/0.2	577.9/1.6	25.1*/1.4

**p < 0.05. Significant difference between the translation (related) condition and the control (unrelated) condition.*

the greater phonological and conceptual overlaps in translation pairs than unrelated pairs in DFM (de Groot, 1992; van Hell and de Groot, 1998). Separate ANOVAs on the magnitude of priming effects were conducted for RT and E% data with the independent factors of cognate status (cognate and non-cognate) and concreteness (abstract and concrete) by subjects (F_1) and by items (F_2).

The behavioral data on RT analysis showed that there was a significant effect for cognate status [$F_1(1,21) = 6.480$, $p = 0.019$, $\eta_p^2 = 0.236$, $F_2(1,76) = 7.900$, $p = 0.006$, $\eta_p^2 = 0.094$], reflecting that the priming magnitude of cognates was greater than that of non-cognates. The main effect of concreteness was not significant [$F_1(1,21) = 0.172$, $p = 0.682$, $\eta_p^2 = 0.008$, $F_2(1,76) = 0.134$, $p = 0.716$, $\eta_p^2 = 0.002$]. The interaction between cognate status and concreteness failed to reach the statistical significance [$F_1(1,21) = 0.898$, $p = 0.354$, $\eta_p^2 = 0.041$, $F_2(1,76) = 1.541$, $p = 0.218$, $\eta_p^2 = 0.020$].

E% data were submitted to the same analysis. There was a significant main effect for cognate status [$F_1(1,21) = 13.174$, $p = 0.002$, $\eta_p^2 = 0.385$, $F_2(1,76) = 4.785$, $p = 0.032$, $\eta_p^2 = 0.059$], indicating that the error rate of cognates was higher than non-cognates. The main effect of concreteness was (marginally) significant [$F_1(1,21) = 16.047$, $p = 0.001$, $\eta_p^2 = 0.433$, $F_2(1,76) = 3.100$, $p = 0.082$, $\eta_p^2 = 0.039$], and the error rate of abstract words was significantly higher than that of concrete words. The interaction between cognate status and concreteness was significant [$F_1(1,21) = 9.682$, $p = 0.005$, $\eta_p^2 = 0.316$, $F_2(1,76) = 5.764$, $p = 0.019$, $\eta_p^2 = 0.070$]. Simple effect comparisons revealed that the error rate in abstract cognates was significantly higher than that in abstract non-cognates [$F_1(1,21) = 16.733$, $p = 0.001$, $\eta_p^2 = 0.443$, $F_2(1,76) = 10.526$, $p = 0.002$, $\eta_p^2 = 0.122$]. In addition, the error rate in abstract cognates was significantly higher than that in concrete cognates [$F_1(1,21) = 22.234$, $p < 0.001$, $\eta_p^2 = 0.514$, $F_2(1,76) = 8.659$, $p = 0.004$, $\eta_p^2 = 0.102$].

Electroencephalogram Data Analyses

As shown in **Figure 2**, visual inspection of the grand mean ERP components elicited by target presentation revealed a negative peak in the 100–200 ms post-stimulus time window (N150), a positive peak in the 200–350 ms time window (P250), and a negative peak in the 350–550 ms time window (N400). In the previous studies, a component in the time window of 100–200 ms was usually considered as the mapping of visual features onto prelexical features in the word-base process (Holcomb and Grainger, 2006; Hoshino et al., 2010). In the present study, N150 can be regarded to reflect the processing of phonological information, and the mean amplitudes of the electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, Oz, and O2 in the 100–200 ms time window for each participant in all conditions were extracted based on Kong et al. (2010). According to Kutas and Federmeier (2011), the N400 component was an indicator of semantic processing, largest over centro-parietal sites. Therefore, we extracted the mean amplitudes of the electrodes C3, Cz, C4, CP3, CPz, CP4, P3, Pz, and P4 in the 350–550 ms time window. In a Chinese–English non-cognates translation priming experiment, Chen et al. (2020) identified N300 component between P200 and P400 as the index of the morphological-semantic interface. Therefore, as an ERP component between N150 and N400, the present P250 component was thought to be an index of the phonological-semantic interface, and the mean amplitudes of the electrodes F3, Fz, F4, C3, Cz, and C4 in the 200–350 ms time window were extracted.

We first examined whether there were translation priming effects of each word type in forward translation.

N150: The mean amplitudes were subjected to a 4 (word type: abstract cognates, concrete cognates, abstract non-cognates and concrete non-cognates) × 2 (relatedness: related and unrelated) × 12 (electrode: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, Oz, and O2) repeated-measure ANOVA. Statistical analysis showed that there was no main effect of either word type [$F(3,63) = 0.798$, $p = 0.471$, $\eta_p^2 = 0.037$], or relatedness [$F(1,21) = 0.001$, $p = 0.972$, $\eta_p^2 < 0.001$]. Additionally, the interaction between word type and relatedness was not significant [$F(3,63) = 0.494$, $p = 0.688$, $\eta_p^2 = 0.023$]. Planned comparisons indicated that none of the word types showed significant relatedness effects (ps ≥ 0.484, η_p^2s ≤ 0.024).

P250: The analysis in the 200–350 ms time window with six electrodes (F3, Fz, F4, C3, Cz, and C4) showed that there was no main effect of either word type [$F(3,63) = 0.800$, $p = 0.470$, $\eta_p^2 = 0.037$], or relatedness [$F(1,21) = 0.008$, $p = 0.931$, $\eta_p^2 < 0.001$]. In addition, the interaction between word type and relatedness was not significant [$F(3,63) = 0.639$, $p = 0.593$, $\eta_p^2 = 0.030$]. Planned comparisons indicated that none of the word types showed significant relatedness effects. (ps ≥ 0.304, η_p^2s ≤ 0.050).

N400: The analysis in the 350–550 ms time window with 9 electrodes (C3, Cz, C4, CP3, CPz, CP4, P3, Pz, and P4) showed that there was a significant main effect of word type [$F(3,63) = 3.310$, $p = 0.035$, $\eta_p^2 = 0.136$], and of relatedness [$F(1,21) = 20.770$, $p < 0.001$, $\eta_p^2 = 0.497$]. The interaction between word type and relatedness was not significant [$F(3,63) = 0.396$, $p = 0.681$, $\eta_p^2 = 0.018$]. Planned comparisons showed that the relatedness effects were significant for all word types [Fs$(1,21) \geq 5.262$, ps ≤ 0.032, η_p^2s ≥ 0.200], and related condition elicited significantly larger N400 than unrelated condition.

Then, we examined whether cross-script phonological similarity and concreteness could elicit greater priming effects in forward translation. The difference waves (the mean amplitudes of the unrelated condition minus the mean amplitudes of the related condition) of the same electrodes as the three ERP components mentioned above in three time windows (100–200, 200–350, and 350–550 ms) were submitted to statistical analyses, respectively.

100–200 ms: The difference waves were subjected to a 2 (cognate status: cognate and non-cognate) × 2 (concreteness: abstract and concrete) × 12 (electrode: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, Oz, and O2) repeated-measure ANOVA. Statistical analysis showed that there was no main effect of either cognate status [$F(1,21) = 0.016$, $p = 0.901$, $\eta_p^2 = 0.001$],

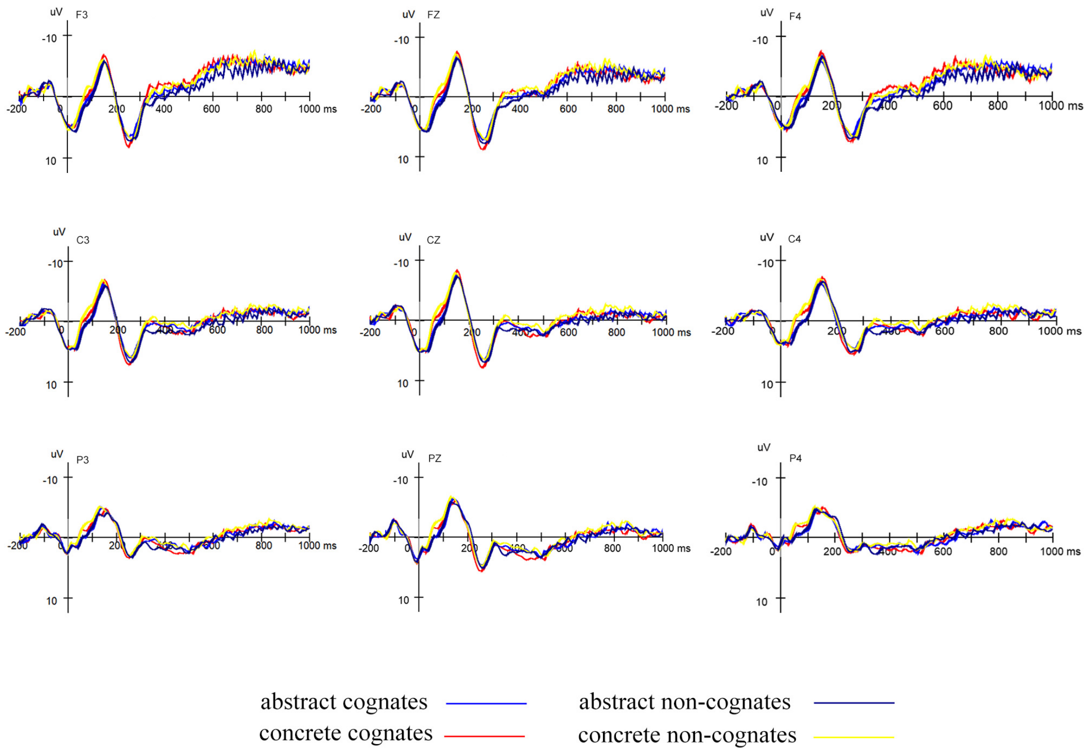


FIGURE 2 | Grand average ERPs elicited by targets primed by abstract cognates, abstract non-cognates, concrete cognates, and concrete non-cognates in L1–L2 translation direction.

or concreteness [$F(1,21) = 0.890$, $p = 0.356$, $\eta_p^2 = 0.041$]. The interaction between cognate status and concreteness was not significant either [$F(1,21) = 1.866$, $p = 0.186$, $\eta_p^2 = 0.082$].

200–350 ms: The analysis in the 200–350 ms time window with six electrodes (F3, Fz, F4, C3, Cz, and C4) showed that there was no main effect of either cognate status [$F(1,21) = 1.012$, $p = 0.326$, $\eta_p^2 = 0.046$], or concreteness [$F(1,21) = 2.158$, $p = 0.157$, $\eta_p^2 = 0.093$]. Additionally, the interaction between cognate status and concreteness was not significant [$F(1,21) = 0.024$, $p = 0.878$, $\eta_p^2 = 0.001$].

350–550 ms: The analysis in the 350–550 ms time window with nine electrodes (C3, Cz, C4, CP3, CPz, CP4, P3, Pz, and P4) showed that there was no main effect of either cognate status [$F(1,21) = 1.268$, $p = 0.273$, $\eta_p^2 = 0.057$], or concreteness [$F(1,21) = 0.062$, $p = 0.806$, $\eta_p^2 = 0.003$]. In addition, the interaction between cognate status and concreteness was not significant [$F(1,21) = 0.004$, $p = 0.952$, $\eta_p^2 < 0.001$].

In summary, the results of Experiment 1 indicated that the translation priming effects from Chinese to English were reflected in RT data, E% data and N400 component. The priming effects of cognate status were shown in RT data and E% data, whereas the priming effects of concreteness and interaction between cognate status and concreteness were only sensitive to E% data. No ERP evidence was found for the greater priming effects of cognate status and concreteness in forward translation since no main effects nor interaction effects in the three time windows (100–200, 200–350, and 350–550 ms) were observed on the difference waves, respectively.

EXPERIMENT 2: L2–L1 (ENGLISH–CHINESE BACKWARD TRANSLATION)

Methods

Experiment 2 explored the role of phonology as well as concreteness effects for Chinese learners of English with a lexical decision task in the masked translation priming paradigm in the L2–L1 translation direction.

Participants

This experiment had the same participants as Experiment 1.

Materials

The experimental materials were the same as in Experiment 1 except for the priming direction in which the primes were English and the targets were presented in Chinese. In the backward direction, translation primes and control primes (unrelated primes) were matched in length, concreteness, and familiarity [$ps \geq 0.119$, *Cohen's* d(s) ≤ 0.161]. In addition, there were also 80 Chinese pseudowords, which were meaningless words comprised of two or three characters. The real words and pseudowords were also matched in the number of strokes.

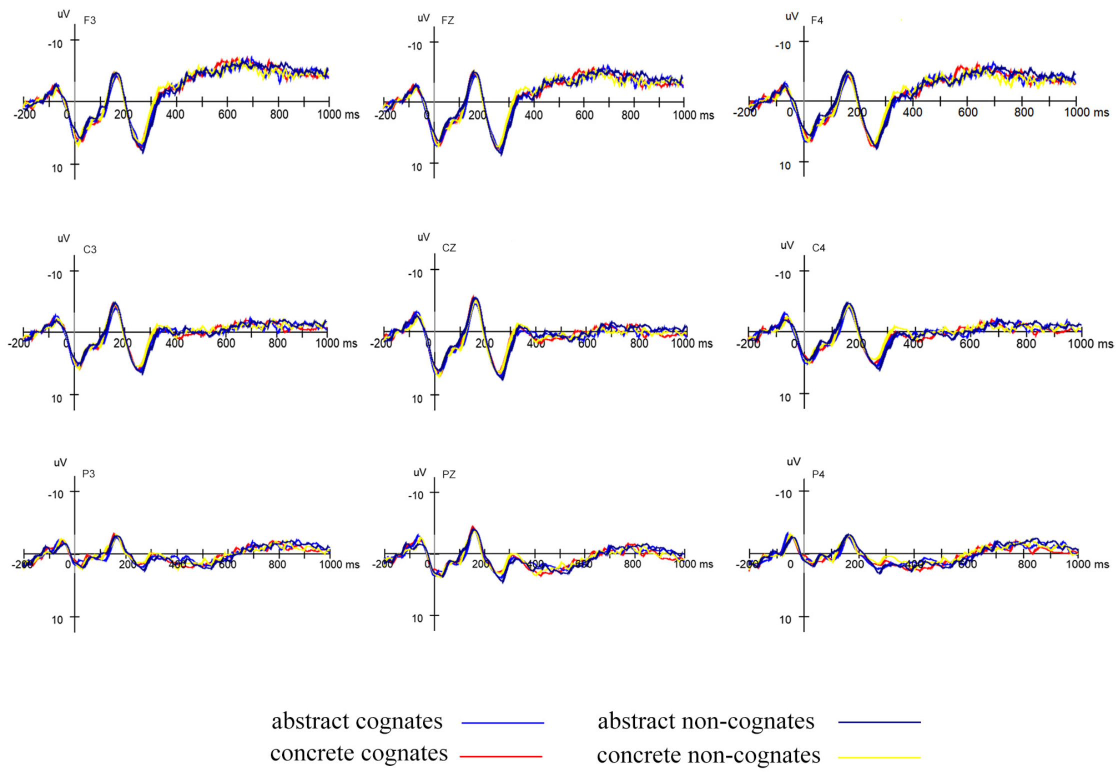


FIGURE 3 | Grand average ERPs elicited by targets primed by abstract cognates, abstract non-cognates, concrete cognates, and concrete non-cognates in L2–L1 translation direction.

Examples of the stimuli in the experiment were presented in **Table 1**.

Procedure

The procedure of Experiment 2 replicates the experimental procedure of Experiment 1, except that the length of pre- and post-masks for English primes was matched with one hash mask for one English letter. Then the Chinese target word was presented until the participant made a response, but for no more than 1,500 ms (see **Figure 1**). There was 1-h interval between Experiments 2 and 1, during which an experiment unrelated to the present two experiments was conducted in order to avoid the mutual influence of the present two experiments.

Electroencephalogram Recording Procedure

The EEG recording procedure was the same as in Experiment 1.

Data Analyses and Results

The three participants whose data were deleted in Experiment 1 due to the high error rates (over 30%), and their data in Experiment 2 were also discarded due to high error rates (over 30%) in data analyses. The mean RTs and error rates in Experiment 2 across each experimental condition are presented in **Table 2**.

Behavioral Analyses

Similar to Experiment 1, the mean reaction times and error rates were submitted to 4 (word type: abstract cognates, concrete cognates, abstract non-cognates, and concrete non-cognates) × 2 (relatedness: related and unrelated) separate ANOVAs by subjects and by items to examine backward translation priming effects of each word type.

The data analysis of RT showed that there was a significant main effect of word type [$F_1(3,63) = 14.080$, $p < 0.001$, $\eta_p^2 = 0.401$, $F_2(3,76) = 5.709$, $p = 0.001$, $\eta_p^2 = 0.184$], and of relatedness [F_1 $(1,21) = 44.362$, $p < 0.001$, $\eta_p^2 = 0.679$, $F_2(1,76) = 73.623$, $p < 0.001$, $\eta_p^2 = 0.492$]. The interaction between word type and relatedness failed to reach significance [$F_1(3,63) = 0.903$, $p = 0.434$, $\eta_p^2 = 0.041$, $F_2(3,76) = 1.046$, $p = 0.377$, $\eta_p^2 = 0.040$]. Planned comparisons revealed that there were significant relatedness effects for all word types [F_1s$(1,21) \geq 7.089$, $ps \leq 0.015$, η_p^2s ≥ 0.252, F_2s$(1,76) \geq 10.160$, $ps \leq 0.020$, η_p^2s ≥ 0.118], with longer response times in unrelated condition than that in related condition.

For the error rate data, the main effect of relatedness was significant by items [$F_2(1,76) = 5.609$, $p = 0.020$, $\eta_p^2 = 0.069$], but not significant by subjects [$F_1(1,21) = 0.606$, $p = 0.445$, $\eta_p^2 = 0.028$]. There was no significant main effect of word type [$F_1(3,63) = 1.035$, $p = 0.367$, $\eta_p^2 = 0.047$, $F_2(3,76) = 1.903$, $p = 0.136$, $\eta_p^2 = 0.070$]. The interaction between word type and relatedness failed to reach significance [$F_1(3,63) = 0.287$, $p = 0.712$, $\eta_p^2 = 0.013$, $F_2(3,76) = 0.226$, $p = 0.878$, $\eta_p^2 = 0.009$].

The magnitudes of priming effects were submitted to 2 (cognate status: cognates and non-cognates) × 2 (concreteness: abstract words and concrete words) separate ANOVAs by subjects and by items to examine the cognate effects and concreteness effects.

The data analyses of RT showed that no significant main effects nor interaction effects were found by subjects ($ps \geq 0.232$, $\eta_p^2 s \leq 0.067$) and by items ($ps \geq 0.219$, $\eta_p^2 s \leq 0.020$).

Meanwhile, the same analysis was conducted on the data of error rate, and no significant main effects nor interaction effects were found in analyses by subjects ($ps \geq 0.303$, $\eta_p^2 s \leq 0.050$) and by items ($ps \geq 0.488$, $\eta_p^2 s \leq 0.006$).

Electroencephalogram Data Analyses

As shown in **Figure 3**, visual inspection of the grand mean ERP components elicited by target presentation revealed a negative peak in the 100–200 ms time window (N150), a positive peak in the 200–350 ms time window (P250), and a negative peak in the 350–550 ms time window (N400). We selected the same electrodes and conducted the same statistical analyses as in Experiment 1.

We first examined whether there were translation priming effects of each word type in backward translation.

N150: The mean amplitudes were subjected to a 4 (word type: abstract cognates, concrete cognates, abstract non-cognates and concrete non-cognates) × 2 (relatedness: related and unrelated) × 12 (electrode: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, Oz, and O2) repeated-measure ANOVA. Statistical analysis showed that there was no main effect of either word type [$F(3,63) = 0.787$, $p = 0.496$, $\eta_p^2 = 0.036$], or relatedness [$F(1,21) = 2.122$, $p = 0.160$, $\eta_p^2 = 0.092$]. There was a significant interaction between word type and relatedness [$F(3,63) = 4.041$, $p = 0.015$, $\eta_p^2 = 0.161$]. Simple effect comparisons revealed that related abstract cognates elicited significantly larger N150 than unrelated abstract cognates [$F(1,21) = 9.122$, $p = 0.007$, $\eta_p^2 = 0.303$], and the mean amplitudes were −1.862 and −1.016 (μV, respectively; that related concrete non-cognates elicited marginally larger N150 than unrelated concrete non-cognates [$F(1,21) = 4.015$, $p = 0.058$, $\eta_p^2 = 0.161$, and the mean amplitudes were −1.682 and −1.033 μV, respectively. There were no significant effects of relatedness for the concrete cognate and abstract non-cognates ($ps \geq 0.107$, $\eta_p^2 s \leq 0.119$).

P250: The analysis in the 200–350 ms time window with six electrodes (F3, Fz, F4, C3, Cz, and C4) showed that there was a main effect of word type [$F(3,63) = 4.551$, $p = 0.009$, $\eta_p^2 = 0.178$]. There was no significant main effect of relatedness [$F(1,21) = 1.760$, $p = 0.199$, $\eta_p^2 = 0.077$]. The interaction between word type and relatedness was not significant [$F(3,63) = 0.550$, $p = 0.611$, $\eta_p^2 = 0.026$]. Planned comparisons indicated that none of the word types showed significant relatedness effects ($ps \geq 0.125$, $\eta_p^2 s \leq 0.136$).

N400: The analysis in the 350–550 ms time window with nine electrodes (C3, Cz, C4, CP3, CPz, CP4, P3, Pz, and P4) revealed no significant main effect of either word type [$F(3,63) = 0.928$, $p = 0.418$, $\eta_p^2 = 0.042$], or relatedness [$F(1,21) = 1.685$, $p = 0.208$, $\eta_p^2 = 0.074$]. Additionally, the interaction between word type and relatedness was not significant [$F(3,63) = 1.271$, $p = 0.293$, $\eta_p^2 = 0.057$]. The planned comparisons revealed that related concrete cognates elicited marginally significantly larger N400 than unrelated concrete cognates [$F(1,21) = 3.344$, $p = 0.082$, $\eta_p^2 = 0.137$], and the mean amplitudes were 1.343 and 0.635 (μV, respectively. There were no significant effects of relatedness for the abstract cognates, concrete non-cognates and abstract non-cognates [$ps \geq 0.172$, $\eta_p^2 s \leq 0.087$].

Then, we examined whether cross-script phonological similarity and concreteness could elicit greater priming effects in backward translation. The difference waves (the mean amplitude of the unrelated condition minus the mean amplitude of the related condition) of the same electrodes as the three ERP components mentioned above in three time windows (100–200, 200–350, and 350–550 ms) were submitted to statistical analyses, respectively.

100–200 ms: The difference waveforms were subjected to a 2 (cognate status: cognate, non-cognates) × 2 (concreteness: abstract and concrete) × 12 (electrode: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, Oz, and O2) repeated-measure ANOVA. Statistical analysis showed that there was no significant main effect of either cognate status [$F(1,21) = 2.134$, $p = 0.159$, $\eta_p^2 = 0.092$], or concreteness [$F(1,21) = 0.355$, $p = 0.558$, $\eta_p^2 = 0.017$]. There was a significant interaction between cognate status and concreteness [$F(1,21) = 14.061$, $p = 0.001$, $\eta_p^2 = 0.401$]. Simple effect comparisons revealed that abstract non-cognates produced significantly smaller difference waves than concrete non-cognates [$F(1,21) = 7.270$, $p = 0.014$, $\eta_p^2 = 0.257$], and the mean amplitudes were −0.573 and 0.649 μV, respectively; abstract cognates produced significantly larger difference waves than abstract non-cognates [$F(1,21) = 16.881$, $p = 0.001$, $\eta_p^2 = 0.446$], and the mean amplitudes were 0.846 and −0.573 μV, respectively.

200–350 ms: The analysis in the 200–350 ms time window with six electrodes (F3, Fz, F4, C3, Cz, and C4) showed that there was no main effect of either cognate status [$F(1,21) = 0.168$, $p = 0.686$, $\eta_p^2 = 0.008$], or concreteness [$F(1,21) = 0.653$, $p = 0.428$, $\eta_p^2 = 0.030$]. In addition, the interaction between cognate status and concreteness was not significant [$F(1,21) = 0.895$, $p = 0.355$, $\eta_p^2 = 0.041$].

350–550 ms: The analysis in the 350–550 ms time window with nine electrodes (C3, Cz, C4, CP3, CPz, CP4, P3, Pz, and P4) showed that there was no main effect of either cognate status [$F(1,21) = 2.812$, $p = 0.108$, $\eta_p^2 = 0.118$], or concreteness [$F(1,21) = 0.458$, $p = 0.506$, $\eta_p^2 = 0.021$]. Additionally, the interaction between cognate status and concreteness was not significant [$F(1,21) = 0.151$, $p = 0.702$, $\eta_p^2 = 0.007$].

In summary, the results of Experiment 2 demonstrated that backward translation priming effects were obtained in RT data for each type of words. The ERP evidence for translation priming effects was obtained in terms of the N150 for abstract cognates and concrete non-cognates, as well as the N400 for concrete cognates. The interaction effect between cognate status and concreteness in the time window of 100–200 ms indicated that concrete non-cognates had greater priming effects than abstract non-cognates and abstract cognates had greater priming effects than abstract non-cognates.

JOINT ANALYSES

In order to investigate whether there existed the asymmetry of translation direction in terms of the priming effect magnitudes between forward translation and backward translation, joint analyses were conducted by comparing behavioral and ERP data in Experiment 1 and Experiment 2.

Behavioral Analyses

Separate ANOVAs were performed for the magnitudes of priming with the factors of translation direction (2: forward direction and backward direction) and word type (4: abstract cognates, concrete cognates, abstract non-cognates, and concrete non-cognates) for the RT to examine the existence of translation priming asymmetry. The main effect of direction reached significance [$F_1(1,21) = 24.387$, $p < 0.001$, $\eta_p^2 = 0.537$, $F_2(1,152) = 88.496$, $p < 0.001$, $\eta_p^2 = 0.368$], and there were larger priming effects in forward translation than in backward translation. The main effect of word type was also significant [$F_1(3,63) = 3.155$, $p = 0.032$, $\eta_p^2 = 0.131$, $F_2(3,152) = 3.435$, $p = 0.019$, $\eta_p^2 = 0.063$]. The interaction between translation direction and word type failed to reach significance [$F_1(3,63) = 0.850$, $p = 0.458$, $\eta_p^2 = 0.039$, $F_2(3,152) = 1.592$, $p = 0.194$, $\eta_p^2 = 0.030$]. Planned comparisons revealed that the direction effects were significant for all word types [F_1s$(1,21) \geq 5.600$, ps ≤ 0.028, η_p^2s ≥ 0.211, F_2s$(1,76) \geq 12.228$, ps ≤ 0.001, η_p^2s ≥ 0.074], and that larger priming effects were found for each word type in forward translation than in backward translation.

For the analysis of error data, the main effect of direction was significant [$F_1(1,21) = 16.179$, $p = 0.001$, $\eta_p^2 = 0.435$, $F_2(1,152) = 28.954$, $p < 0.001$, $\eta_p^2 = 0.160$], and the error rate was higher in forward translation than that in backward translation. The main effect of word type was also significant [$F_1(3,63) = 11.709$, $p < 0.001$, $\eta_p^2 = 0.358$, $F_2(3,152) = 4.015$, $p = 0.009$, $\eta_p^2 = 0.073$]. There was a significant interaction between translation direction and word type [$F_1(3,63) = 6.251$, $p = 0.004$, $\eta_p^2 = 0.229$, $F_2(3,152) = 3.978$, $p = 0.009$, $\eta_p^2 = 0.073$]. Simple effect comparisons revealed that priming effects of all word types were (marginally) significantly larger in forward translation than those in backward translation in analyses by participants [F_1s$(1,21) \geq 4.841$, ps ≤ 0.039, η_p^2s ≥ 0.187] except for concrete non-cognates [$F_1(1,21) = 2.872$, $p = 0.105$, $\eta_p^2 = 0.120$], and by items [F_2s$(1,152) \geq 2.870$, ps ≤ 0.092, η_p^2s ≥ 0.019] except for abstract non-cognates [$F_2(1,152) = 2.542$, $p = 0.113$, $\eta_p^2 = 0.016$].

Electroencephalogram Data Analyses

To examine the existence of the priming asymmetry between forward translation and backward translation, we compared the ERP difference waves in the time windows of 100–200 ms, 200–350 ms, and 350–550 ms between the two directions.

100–200 ms: The difference waves were subjected to a 2 (translation direction: forward and backward) × 4 (word type: abstract cognates, concrete cognates, abstract non-cognates, and concrete non-cognates) × 12 (electrode: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1, Oz, and O2) repeated-measure ANOVA. Statistical analysis showed that there was a significant main effect of word type [$F(3,63) = 4.136$, $p = 0.011$, $\eta_p^2 = 0.165$]. The main effect of direction was not significant [$F(1,21) = 0.987$, $p = 0.332$, $\eta_p^2 = 0.045$]. The interaction between translation direction and word type was not significant [$F(3,63) = 0.733$, $p = 0.510$, $\eta_p^2 = 0.034$]. Planned comparisons revealed that abstract cognates in forward translation elicited marginally significantly smaller difference waves than abstract cognates in backward translation [$F(1,21) = 3.338$, $p = 0.082$, $\eta_p^2 = 0.137$], and the mean amplitudes were 0.072 and 0.846 μV, respectively. There were no translation direction effects in terms of the concrete cognates, abstract non-cognates and concrete non-cognates (ps ≥ 0.662, η_p^2s ≤ 0.009).

200–350 ms: The analysis in the 200–350 ms time window with six electrodes (F3, Fz, F4, C3, Cz, and C4) presented no significant main effect of either direction [$F(1,21) = 0.854$, $p = 0.366$, $\eta_p^2 = 0.039$], or word type [$F(3,63) = 0.993$, $p = 0.401$, $\eta_p^2 = 0.045$]. Additionally, the interaction between translation direction and word type was not significant [$F(3,63) = 0.580$, $p = 0.621$, $\eta_p^2 = 0.027$]. Planned comparisons indicated that none of the word types showed significant direction effects (ps ≥ 0.182, η_p^2s ≤ 0.083).

350–550 ms: The analysis in the 350–550 ms time window with nine electrodes (C3, Cz, C4, CP3, CPz, CP4, P3, Pz, and P4) revealed a significant main effect of direction [$F(1,21) = 20.726$, $p < 0.001$, $\eta_p^2 = 0.497$]. The main effect of word type was not significant [$F(3,63) = 1.549$, $p = 0.217$, $\eta_p^2 = 0.069$]. The interaction between translation direction and word type was not significant [$F(3,63) = 0.150$, $p = 0.903$, $\eta_p^2 = 0.007$]. Planned comparisons revealed that the asymmetry of the priming effects between forward translation and backward translation existed in all word types [Fs$(1,21) \geq 4.564$, ps ≤ 0.045, η_p^2s ≥ 0.179] with larger priming effects in forward translation than in backward translation.

In summary, the differences between L1–L2 direction and L2–L1 direction in behavioral data analyses reflected greater priming effects in forward translation than in backward translation. Meanwhile, the translation priming asymmetry was observed in terms of smaller priming effect for forward translation than for backward translation in the time window of 100–200 ms for abstract cognates, and in terms of larger priming effects for forward translation than for backward translation in the time window of 350–550 ms for each type of words.

GENERAL DISCUSSION

The present study investigated the impact of cross-script cognate phonological activation and concreteness with Chinese–English cognates that shared similar pronunciation and concept simultaneously in masked translation priming paradigm based on ERP technology. The roles of cross-script cognate status and concreteness were investigated in forward translation and backward translation throughout the analyses of the behavioral data and ERP data. The results of behavioral data analyses showed the translation priming effects for four types of word in both translation directions, and greater priming effects were observed for cross-script cognate status with larger priming effects for cognates than for non-cognates in forward translation, but not in backward translation, and the translation priming asymmetry was found. However, the ERP evidence from the results of data analyses in Experiment 1, Experiment 2 and joint analyses showed different influences of cognate status and concreteness on cross-script language processing, and confirmed the existence of the asymmetry of translation directions indicated by different ERP indices. As a whole, N400 effect was found to be closely related to cross-script cognate status advantage and the role of concreteness effect in forward translation, and in the reverse direction N150 and N400 effects were related to the roles of cross-script cognate effect and concreteness effect. In the time window of 100–200 ms for backward translation, we found greater priming effects in concrete words than in abstract words for non-cognates and greater priming effects in cognates than in non-cognates for abstract words. Meanwhile, the asymmetry of translation directions was observed with smaller priming effects in forward translation than in backward translation in the time window of 100–200 ms for abstract cognates, and with larger priming effects in forward translation than in backward translation in the time window of 350–550 ms for each type of words.

Priming Effects of Cross-Script Cognates

In the previous studies of cognate status, phonological information failed to disentangle from orthographic similarity within same-script languages. Chinese–English cognates that shared similar semantic and phonological representation without orthographic links showed strong evidence for the phonological advantage with respect to cognate status for cross-script languages in the present study.

In the present study, priming effects of cognate status were observed in the N150 in backward translation, and greater priming effects of cross-script cognate status in cognates than in non-cognates for abstract words were found in the time window of 100–200 ms also in backward translation. On the contrary, neither N150 for priming effects of Chinese–English translation nor greater priming effects of phonological information between cognates and non-cognates in the time window of 100–200 ms were found in forward translation. And the translation asymmetry caused by the priming effects of cross-script cognates was indicated by larger amplitudes in the time window 100–200 ms for backward translation than for forward translation. The discrepancy in cognate status due to phonological similarity reflected by N150 component might be interpreted as an indicator of phonological processing during sub-lexical phase since this component was regarded as the mapping of visual features onto prelexical features during word-base process in other studies (Holcomb and Grainger, 2006; Hoshino et al., 2010). The results in backward translation provided the evidence for more phonological overlaps in cognates than in non-cognates.

On the other hand, no ERP evidence was observed for priming effects of Chinese–English translation or for greater priming effects of phonological information between cognates and non-cognates in forward translation. The results might show that the phonological priming effects between cross-script cognates and non-cognates with respect to their corresponding translation equivalents keep similar in forward translation (from L1 Chinese to L2 English). It can be found that the present findings extend the cognate hypothesis stated in DFM (de Groot, 1992; van Hell and de Groot, 1998). While DFM emphasizes the importance of the semantic features in bilingual mental lexicon, it pays less attention to other linguistic features such as phonological features, orthographic features. More phonological features are activated for cross-script cognates than for cross-script non-cognates in masked translation priming paradigm. More activated phonological features in Chinese–English mental lexicon lead to greater phonological priming effects in English–Chinese (L2–L1) priming pairs, not in Chinese–English (L1–L2) priming pairs. The English learning environment for Chinese learners of English may account for the lack of role of phonological similarity in L1–L2 translation. In English classroom, English learners are usually taught to learn L2 English words by remembering their Chinese equivalents, not vice versa. Thus a more frequent repetition from English to Chinese, not from Chinese to English may form a strong phonological activation for English. Therefore, compared with Chinese primes in the L1–L2 direction, English primes as phonograms in the L2–L1 direction gave more direct prompt to activate phonological representation of the target. Meanwhile, the results of the present study are in line with other empirical studies. For example, Zhang et al. (2018) examined the translation priming for cross-script cognates within behavioral data and found that in the L1–L2 priming direction, there was no priming advantage for cognates over non-cognates, and both L1–L2 cognate and non-cognate primes similarly facilitated L2 word recognition and that in the L2–L1 priming direction, only cognate primes facilitated L1 word processing while non-cognates primes failed to generate priming effects. Therefore, the present findings partly support RHM (Kroll and Stewart, 1994) in terms of the weak link from the L1–L2 direction and a strong link from the L2–L1 direction.

The BIA+ model assumes that the appearance of primes leads to activation of phonology, which could render the phonological representations of the targets more easily activated (if the prime and target have phonological similarity). Thus, the pre-activated phonology could accelerate the process of word recognition. There are two routes in the BIA+ model when the lexical phonology is activated, the lexical route and the prelexical route. In the former, activation spreads from sub-lexical orthography to lexical orthography and then to lexical phonology, whereas in the latter, sub-lexical orthography activates sub-lexical phonology which subsequently activates lexical phonology (Dijkstra and van Heuven, 2002). It is possible for phonological activation to occur in the recognition of alphabetical languages since they have regular grapheme-to-phoneme conversion rules. However, in the present cross-script study, the phonological priming effects occurred in the time window of 100–200 ms for abstract words. This time course is ahead of the modulation in 200–250 ms proposed by Ando et al. (2015) in spite of different translation directions. After all, the activation of lexical stage could not accomplish as early as 250 ms after the onset of the stimuli during masked onset priming (Jouravlev et al., 2014). In addition, it has been suggested that N150 component might be interpreted as the sub-lexical phase of lexical processing in mental lexicon in which phonemes or graphemes are activated. Thus the N150 might result from the priming effects that occurred at prelexical stage due to the similar phonological activation in the lexical decision task. Pinyin, a system of Romanized spelling which describes how each Chinese character is pronounced, is in daily use for students in China mainland (for example, typing). Zhou et al. (2010) argued that the pinyin of a given Chinese word could have orthographic overlap with its phonologically similar English word. For instance, "dao" is the Chinese character "道" in pinyin, and there are two overlapping letters in "dao" and its phonologically similar English word "door." This explanation is also applicable to the present study in that most pinyin of the loan words and their English equivalents are similar to some degree. The processing for the pinyin of "nacui" ("纳粹" in Chinese) was accelerated by its English translation "Nazi" with greater phonological overlap as a prime at the sub-lexical processing phase.

Priming Effects of Concreteness

The priming effects of concreteness were observed in N400 component in terms of translation priming effects for the four types of words in forward translation, and for cognate concrete words in backward translation, and also in larger priming effects in forward translation than in backward translation in the time window of 350–550 ms for each type of words in the present study. However, no greater priming effects of concreteness between concrete words and abstract words with respect to their corresponding translation equivalents in time window of 350–550 ms were found in forward translation and backward translation.

It has been found that N400 component was sensitive to semantic cognition load. As concreteness can be regarded as one part of semantic information, the N400 component is closely related to the priming effect of concreteness. The translation priming pairs elicited greater N400 than control pairs (non-translation priming pairs) in forward translation. One possible explanation is that the priming effects of concreteness in terms of N400 amplitudes come from the greater semantic overlap within translation pairs than within control pairs, which leads to the activation of more semantic features for Chinese primes than for English primes. Larger priming effects in forward translation than in backward translation in the time window of 350–550 ms for each type of words in the present study provided ERP evidence for the existence of translation asymmetry caused by the priming effects of concreteness.

However, no greater priming effects of concreteness between concrete words and abstract words with respect to their corresponding translation equivalents in time window of 350–550 ms were observed in two translation directions. This finding demonstrated that the priming effects of both concrete words and abstract words keep balanced in the two directions, and further suggested similar conceptual overlap between concrete words and abstract words with respect to their corresponding translation equivalents regardless of their concreteness. Indeed, the masked priming translation paradigm conducted in the present study is distinctive from the single lexical decision or semantic categorization task in which no context information was provided for the semantic knowledge of the target. More specifically, participants could only see the target without the primes in the single lexical decision or semantic categorization task. Concrete words would elicit greater semantic processing than abstract words (Barber et al., 2013). In the present study, it is assumed that compared with abstract primes, concrete primes may provide more specific semantic information for the targets to facilitate the semantic processing. However, both Chinese concrete primes and abstract primes offered quantitatively equal semantic clues to the English targets in L1–L2 translation direction, and both English concrete primes and abstract primes offered quantitatively equal semantic clues to the Chinese targets in the L2–L1 direction.

The balanced priming effects between concrete words and abstract words in ERP data analyses may be caused by SOA between primes and targets in the masked translation priming paradigm. Till now, it is still under debate whether or not concreteness of words can modulate the priming effects, since concreteness effects were SOA-sensitive, and only the priming paradigms within a certain range of SOAs could produce the facilitation effect of concreteness (Ferré et al., 2017). Chen et al. (2014) designed a study with 50 ms for the primes and 150 ms for the backward masks to investigate the concreteness effects in lexical decision task and semantic categorization task, and found no significant difference between concrete words and abstract words. Ferré et al. (2017) failed to find concreteness effects with the 50 ms SOA in the masked priming paradigm, but the concrete words showed greater advantages relative to abstract words in a 100 ms SOA. With the observed priming effects of concreteness in terms of N400 and the balanced concreteness effect in the 350–550 ms time window, we may have found the appropriate SOA for the studies of concreteness effect. In Chen et al. (2014) and Ferré et al. (2017), the primes lasting for 50 ms might not be so long enough to activate the targets, so the discrepancy between concrete words and abstract words disappeared in response latencies. The other possibility for the discrepancy of the related studies might ascribe to the technology. After all, ERP based studies are more sensitive to measuring the time course of processing, while behavioral studies mainly focus on the results of processing.

The Role of Interplay Between Cross-Script Cognate Status and Concreteness

As discussed previously, it seemed that N150 is closely related to the processing of phonological information, while N400 is associated with concreteness. In the previous studies, the N250 component was thought to reflect the mapping of prelexical representations onto whole-word form representations (Holcomb and Grainger, 2006; Grainger and Holcomb, 2009). Chen et al. (2020) identified N300 component between P200 and P400 as the index of the morphological-semantic interface in Chinese–English non-cognates translation priming experiment. Therefore, it is possible that the P250 component elicited between N150 and N400 in the present study is closely related to the processing of phonological-semantic interface, and can be thought of as an index of phonological-semantic interface, reflecting the mapping of phonological information onto semantic representation.

For the P250 effect, the present study only found the main effect of word type in Experiment 2 (L2–L1 translation experiment), and no interaction effect between phonological similarity and concreteness effects was found in terms of P250 or in the time window of 200–350 ms. It seemed that the phonological and semantic features of English–Chinese cognates may not be closely related, and the phonological activation and concreteness representation were independent of each other regardless of translation directions in time window 200–350 ms, which may be explained by differences between English and Chinese. Unlike the close relationship between phonemes and meanings of phonography in English, Chinese characters are hieroglyphs and thus have relatively loose relation with the phonological features. Therefore, in Chinese–English cognate translation direction, no priming effects of the phonological-semantic interface were observed, but in English–Chinese cognate translation direction, P250 for main effect of word type was detected perhaps because more phonological-semantic overlapping information was activated in English cognates than in their Chinese equivalents. Therefore, it is crucial to further explore the interplay between cross-script cognate status and concreteness factors considering that lexical phonology is one of the routes accessing to semantic representation as illustrated in the BIA+ model.

CONCLUSION

The present study investigated the concreteness effects of cross-script phonological activation with masked translation priming paradigm based on ERP technology in two experiments. N400 effect was found to be closely related to concreteness effects in Experiment 1. N150 and N400 effects were related to cross-script cognate effects and concreteness effects in Experiment 2. Greater priming effects of cross-script cognate status in cognates than in non-cognates for abstract words were found in the time window of 100–200 ms. Meanwhile, the translation asymmetry was observed in the time window of 100–200 ms with smaller priming effects for abstract cognates in forward translation than in backward translation, and in the time window of 350–550 ms with larger priming effects for each type of words in forward translation than in backward translation. We discussed the phonological activation and concreteness effects as well as translation asymmetry in view of the function of N150 and N400 components and the relevant models, mainly the Distributed Feature Model and Bilingual Interactive Activation (BIA+) model.

The present study only focused on the influence of cognate status and concreteness on bilingual memory, which cannot give a whole picture of bilingual memory. Additionally, we cannot deny that the development issues such as the age of acquisition may have great influence on the vocabulary learning of L1 and L2, and impact bilingual structure. Further cross-script studies might employ other techniques such as computational models to deal with as many variables as possible to examine phonological similarity and concreteness as computational models offer particular advantages in dealing with complex interactions between variables that are often confounded in natural language situations (Li and Zhao, 2018), which may shed more light on the principle of phonological activation and concreteness feature in bilingual visual word recognition.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The study involving human participants were reviewed and approved by the Ethics Review Committee of College of Foreign Languages, Ocean University of China. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

SC, TF, and MZ put forward the idea of the study and contributed to the research design, experiments, and manuscript revision. YZ, YP, and LY contributed to the data collection, data analysis, and manuscript drafting. XG contributed to the research design, data analysis, and manuscript revision. All authors contributed to this article and approved the submitted version.

FUNDING

This research was supported by the National Social Science Fund of China (No: 21BYY114) and by the Fundamental Research Funds for the Central Universities (No: 202061042).

ACKNOWLEDGMENTS

The authors would like to thank the editors and reviewers for their valuable comments.

REFERENCES

Ando, E., Jared, D., Nakayama, M., and Hino, Y. (2014). Cross-script phonological priming with Japanese Kanji primes and English targets. *J. Cogn. Psychol.* 26, 853–870. doi: 10.1080/20445911.2014.971026

Ando, E., Matsuki, K., Sheridan, H., and Jared, D. (2015). The locus of Katakana-English masked phonological priming effects. *Biling. Lang. Cogn.* 18, 101–117. doi: 10.1017/S1366728914000121

Barber, H. A., Otten, L. J., Kousta, S. T., and Vigliocco, G. (2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain Lang.* 125, 47–53. doi: 10.1016/j.bandl.2013.01.005

Basnight-Brown, D. M., and Altarriba, J. (2007). Differences in semantic and translation priming across languages: the role of language direction and language dominance. *Mem. Cogn.* 35, 953–965. doi: 10.3758/BF03193468

Chen, B., Liang, L., Cui, P., and Dunlap, S. (2014). The priming effect of translation equivalents across languages for concrete and abstract words. *Acta Psychol.* 153, 143–152. doi: 10.1016/j.actpsy.2014.10.004

Chen, S., Qiu, J., Peng, Y., Zhang, Y., and Yang, L. (2020). A translation priming ERP study on bilingual mental lexical representation for Chinese learners of English. *Foreign Lang. Teach. Res.* 3, 423–434. doi: 10.19923/j.cnki.fltr.2020.03.009

de Groot, A. M. B. (1992). "Bilingual lexical representation: A closer look at conceptual representations," in *Orthography, phonology, morphology, and meaning*, eds R. Frost and L. Kaatz (Amsterdam: Elsevier), 389–412.

Dijkstra, T., and van Heuven, W. J. B. (1998). "The BIA model and bilingual word recognition," in *Localist connectionist approaches to human cognition*, eds J. Grainger and A. M. Jacobs (Manwah, NJ: Lawrence Erlbaum Associates), 189–225.

Dijkstra, T., and van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Biling. Lang. Cogn.* 5, 175–197. doi: 10.1017/S1366728902003012

Dimitropoulou, M., Duñabeitia, J. A., and Carreiras, M. (2011). Two words, one meaning: Evidence of automatic co-activation of translation equivalents. *Front. Psychol.* 2:1–20. doi: 10.3389/fpsyg.2011.00188

Dong, Y., Gui, S., and MacWhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Biling. Lang. Cogn.* 8, 221–238. doi: 10.1017/S1366728905002270

Duñabeitia, J. A., Perea, M., and Carreiras, M. (2010). Masked translation priming effects with highly proficient simultaneous bilinguals. *Exp. Psychol.* 57, 98–107. doi: 10.1027/1618-3169/a000013

Ferré, P., Sanchez-Casas, R., Comesana, M., and Demestre, J. (2017). Masked translation priming with cognates and noncognates: is there an effect of words' concreteness? *Biling. Lang. Cogn.* 20, 770–782. doi: 10.1017/S1366728916000262

Finkbeiner, M., Forster, K., Nicol, J., and Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *J. Mem. Lang.* 51, 1–22. doi: 10.1016/j.jml.2004.01.004

Francis, W. S. (2005). "Bilingual semantic and conceptual representation," in *Handbook of bilingualism: Psycholinguistic approaches*, eds J. F. Kroll and A. B. M. de Groot (New York, NY: Oxford University Press), 251–267.

Francis, W. S., and Goldmann, L. L. (2011). Repetition priming within and between languages in semantic classification of concrete and abstract words. *Memory* 19, 653–663. doi: 10.1080/09658211.2011.595724

Gollan, T. H., Forster, K. I., and Frost, R. (1997). Translation priming with different scripts: masked priming with cognates and non-cognates in Hebrew-English bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 23, 1122–1139. doi: 10.1037/0278-7393.23.5.1122

Grainger, J., and Holcomb, P. J. (2009). Watching the word go by: on the time-course of component processes in visual word recognition. *Lang. Linguist. Compas.* 3, 128–156. doi: 10.1111/j.1749-818X.2008.00121.x

Holcomb, P. J., and Grainger, J. (2006). On the time course of visual word recognition: an event-related potential investigation using masked repetition priming. *J. Cogn. Neurosci.* 18, 1631–1643. doi: 10.1162/jocn.2006.18.10.1631

Hoshino, N., Midgley, K. J., Holcomb, P. J., and Grainger, J. (2010). An ERP investigation of masked cross-script translation priming. *Brain Res.* 1344, 159–172. doi: 10.1016/j.brainres.2010.05.005

Huang, H., Lee, C., and Federmeier, K. D. (2010). Imagine that! ERPs provide evidence for distinct hemispheric contributions to the processing of concrete and abstract concepts. *NeuroImage* 49, 1116–1123. doi: 10.1016/j.neuroimage.2009.07.031

Jouravlev, O., Lupker, S. J., and Jared, D. (2014). Cross-language phonological activation: Evidence from masked onset priming and ERPs. *Brain Lang.* 134, 11–22. doi: 10.1016/j.bandl.2014.04.003

Keuleers, E., and Brysbaert, M. (2010). Wuggy: a multilingual pseudoword generator. *Behav. Res. Methods* 42, 627–633. doi: 10.3758/BRM.42.3.627

Kim, J., and Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *J. Mem. Lang.* 49, 484–499. doi: 10.1016/S0749-596X(03)00093-7

Kong, L., Zhang, J. X., Kang, C., Du, Y., Zhang, B., and Wang, S. (2010). P200 and phonological processing in Chinese word recognition. *Neurosci. Lett.* 473, 37–41. doi: 10.1016/j.neulet.2010.02.014

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174. doi: 10.1006/jmla.1994.1008

Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Ann. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123

Li, P., and Zhao, X. W. (2018). "Computational modeling," in *Research Methods in Psycholinguistics and the Neurobiology of Language: A Practical Guide*, eds A. M. B. de Groot and P. Hagoort (Hoboken, NJ: John Wiley & Sons, Inc), 208–229.

Liu, Z. T., Gao, M. K., Mai, Y. Q., and Shi, Y. W. (1984). *A Dictionary of Loan Words and Hybrid Words in Chinese*. Shang Hai: Shanghai Lexicographical Publishing House.

Nakayama, M., Sears, C. R., Hino, Y., and Lupker, S. J. (2012). Cross-script phonological priming for Japanese-English bilinguals: evidence for integrated phonological representations. *Lang. Cogn. Process.* 27, 1563–1583. doi: 10.1080/01690965.2011.606669

Nakayama, M., Sears, C. R., Hino, Y., and Lupker, S. J. (2013). Masked translation priming with Japanese–English bilinguals: interactions between cognate status, target frequency and L2 proficiency. *J. Cogn. Psychol.* 25, 949–981. doi: 10.1080/20445911.2013.839560

Palmer, S. D., MacGregor, L. J., and Havelka, J. (2013). Concreteness effects in single-meaning, multi-meaning and newly acquired words. *Brain Res.* 1538, 135–150. doi: 10.1016/j.brainres.2013.09.015

Pavlenko, A. (2009). "Conceptual Representation in the Bilingual Lexicon and Second Language Vocabulary Learning," in *The Bilingual Mental Lexicon: Interdisciplinary Approaches*, ed. A. Pavlenko (Clevedon: Multilingual Matters), 125–160.

Schoonbaert, S., Duyck, W., Brysbaert, M., and Hartsuiker, R. J. (2009). Semantic and translation priming from a first language to a second and back: aaking sense of the findings. *Mem. Cogn.* 37, 569–586. doi: 10.3758/MC.37.5.569

Tolentino, L. C., and Tokowicz, N. (2009). Are pumpkins better than heaven? An ERP investigation of the concrete-word advantage using order effects. *Brain Lang.* 110, 12–22. doi: 10.1016/j.bandl.2009.01.001

van Hell, J. G., and de Groot, A. M. B. (1998). Conceptual representation in bilingual memory: effects of concreteness and cognate status in word association. *Biling. Lang. Cogn.* 1, 193–211. doi: 10.1017/s1366728998000352

Zhang, J., Wu, C., Zhou, T., and Meng, Y. (2018). Cognate facilitation priming effect is modulated by writing system: Evidence from Chinese-English bilinguals. *Int. J. Biling.* 23, 553–566. doi: 10.1177/1367006917749062

Zhang, Q., Guo, C. Y., Ding, J. H., and Wang, Z. Y. (2006). Concreteness effects in the processing of Chinese words. *Brain Lang.* 96, 59–68. doi: 10.1016/j.bandl.2005.04.004

Zhao, X., and Li, P. (2013). Simulating cross-language priming with a dynamic computational model of the lexicon. *Biling. Lang. Cogn.* 16, 288–303. doi: 10.1017/S1366728912000624

Zhou, H., Chen, B., Yang, M., and Dunlap, S. (2010). Language nonselective access to phonological representations: evidence from Chinese-English bilinguals. *Q. J. Exp. Psychol.* 63, 2051–2066. doi: 10.1080/17470211003718705

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Psychology

ORIGINAL RESEARCH
published: 24 March 2022
doi: 10.3389/fpsyg.2022.803518

Influences of First and Second Language Phonology on Spanish Children Learning to Read in English

Carmen Hevia-Tuero[1], Sara Incera[2] and Paz Suárez-Coalla[1]*

[1]Laboratorio de Psicología del Lenguaje, Department of Psychology, University of Oviedo, Oviedo, Spain, [2]Multilingual Laboratory, Department of Psychology, Eastern Kentucky University, Richmond, KY, United States

Children learning to read in two different orthographic systems are exposed to cross-linguistic interferences. We explored the effects of school (Monolingual, Bilingual) and grade (2nd, 4th, and 6th) on phonological activation during a visual word recognition task. Elementary school children from Spain completed a lexical decision task in English. The task included real words and pseudohomophones following Spanish or English phonological rules. Using the mouse-tracking paradigm, we analyzed errors, reaction times, and computer mouse movements. Children in the bilingual school performed better than children in the monolingual school. Children in higher grades performed better than children in lower grades. The interference effect of Spanish phonology was weak and became weaker in higher grades. Spanish children differentiate between first and second language grapheme-to-phoneme correspondences since early on in the educational process. In 6th grade, children from the bilingual school responded better to words and Spanish pseudohomophones, while children from the monolingual school were less distracted by the English pseudohomophones. Children in the bilingual school had stronger inhibition of Spanish (L1) phonology and stronger activation of English (L2) phonology. Instructional method plays an important role on the processing strategies Spanish children rely on when reading in English. School and grade influence the link between orthographic and phonological representations.

Keywords: orthography, phonology, bilingual reading, pseudohomophones, mouse-tracking

OPEN ACCESS

Edited by:
Koji Miwa,
Nagoya University, Japan

Reviewed by:
Adrian Pasquarella,
University of Delaware, United States
Marina Ferroni,
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina
Lama K. Farran,
University of West Georgia, United States

****Correspondence:***
Carmen Hevia-Tuero
heviacarmen@uniovi.es

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *28 October 2021*
Accepted: *07 March 2022*
Published: *24 March 2022*

Citation:
Hevia-Tuero C, Incera S and Suárez-Coalla P (2022) Influences of First and Second Language Phonology on Spanish Children Learning to Read in English. Front. Psychol. 13:803518. doi: 10.3389/fpsyg.2022.803518

INTRODUCTION

Learning to read is a key foundation for education, and much effort is invested in ensuring all children are able to read properly. Learning a second language is also important, as it allows worldwide communication and it improves professional development. Thus, how children learn to read in a second language is an important topic to investigate.

Speaking more than one language is an important skill highly valued within the European educational systems (Council of Europe, 2001). In Spain, studying a foreign language at school is compulsory for all children. English is by far the most popular, and the number of schools implementing Spanish–English bilingual programs is increasing. Many bilingual education programs are being developed, but English as a second language instructional methods vary across schools (Hélot and Cavalli, 2017). The consequences of this variety of educational approaches have not been fully investigated, but these different techniques could be influencing how children learn to read in a second language. For instance, it has been demonstrated that being more exposed to second language impacts positively on language learning (Farukh and Vulchanova, 2015). Thus, it is essential to determine the role of school and to ensure teachers know how to help students read in their second language. Despite the undeniable benefits of being exposed to a second language since early stages (Winsler et al., 1999; Larson-Hall, 2008; Olulade et al., 2016), children face the challenge of simultaneously learning to read and write in two different orthographies. The purpose of this study is to determine how Spanish children learn to read in English. In particular, we examine the effect of school and grade on word processing during second language reading.

When children learn to read in their native language, they learn a specific set of grapheme-to-phoneme correspondences. The goal is to connect a grapheme (the letter "a") to its correspondent phoneme (the sound/ʌ/). For instance, in order to form the word cat, the letters c - a - t are processed and connected to the sound each is related to. In languages like English, it may also be necessary to learn correspondences between larger segments of writing, like syllables or rhymes, and their phonological representations. Regardless of the size of the processing units (as long as it is not the whole word) this serial rule-based procedure is known as sublexical decoding (Rau et al., 2014). A stage of sublexical decoding is included in some reading developmental models (Frith, 1985; Ehri, 2005). In later stages, readers transit from this sublexical to a lexical strategy, which improves fluency and efficiency. However, as Share (1995) states in the self-teaching hypothesis, this developmental transition can be different for each word and strategies may overlap. Every time a word is successfully decoded, children acquire specific orthographic information. The orthographic representation of the word will be formed through a self-teaching mechanism after repeated exposures. The coexistence of phonological and lexical processing continues along the reader's life. This highlights the relevance of print-to-sound correspondence knowledge, which is specific to the orthography of each language (Goswami et al., 2001). Children learning to read are influenced by orthographic depth of their native language—the extent to which the orthography is a phonetic representation of speech (Katz and Feldman, 2017). This reliability of print-to-sound correspondences is based on the complexity and unpredictability of the orthography (Schmalz et al., 2015; De Simone et al., 2021). In more shallow orthographies (e.g., Spanish), each grapheme is associated with a single phoneme; there is a one-to-one correspondence with relatively few exceptions. However, in deeper orthographies (e.g., English) each grapheme can be associated with multiple phonemes. In these cases, the formation of strong orthographic representations and the transition from a sublexical to a lexical strategy will be more likely than in shallower orthographies.

Orthographic depth determines the main route (phonological or lexical) children rely on most during literacy acquisition (Ziegler and Goswami, 2005). For instance, children learning to read in a shallower orthography language like Spanish rely heavily on the phonological route and use more frequently grapheme–phoneme decoding strategies (Bhide, 2015). This facilitates code learning, allowing Spanish children to reach accuracy in reading sooner than their counterparts who learn to read in deeper orthography languages like English (Seymour et al., 2003). On the contrary, children learning to read in a deeper orthography like English rely more frequently on the lexical route (Defior and Serrano, 2005). Because not all graphemes correspond to a unique phoneme in English, children's sublexical decoding is based on units bigger than graphemes (e.g., syllables). The orthographic context, as well as other sublexical elements like syllables or rhymes, must be taken into consideration in more deep orthographies. This makes decoding a more complex task for English than for Spanish readers, which results in children who are learning to read in English reaching reading accuracy about a year later than their Spanish counterparts.

In bilingual programs children are exposed to another language and must learn an additional set of grapheme-to-phoneme mappings. While English and Spanish share the same alphabet, the grapheme–phoneme equivalences are not the same. For instance, the sound /i/ is represented with i in Spanish and ee or ea in English. This sound is perceived in English as a long vowel, but vowel length is not a relevant aspect in Spanish (Fox et al., 1995). Furthermore, other phonemes may be perceived as two separate sounds in English but a single sound for Spanish speakers. For instance, the /ʤ/ in jeans (which is not contrastive with the /j/ in yellow) or the /i/ and /ɪ/, which are both perceived and represented as the same grapheme i. This substitution of the spelling of an English specific phoneme (like /iː/ or /ʌ/) for the spelling of the closest phoneme in Spanish (like/i/ or/a/) has been frequently reported (Cronnell, 1985; Zutell and Allen, 1988; Fashola et al., 1996; Sun-Alperin and Wang, 2008; Howard et al., 2012). In the case of cheese, for example, its transcription following Spanish rules would be chis. This lack of discrimination affects not only the vowel sound, but the final/z/ phoneme as well. This voice alveolar fricative does not exist in Spanish, and its closest phoneme is a voiceless alveolar fricative (/s/). Moreover, in Spanish the letter "z" represents the sound/θ/, which is normally spelled as "th" in English. These inconsistencies help illustrate the incongruences that Spanish children encounter when learning to read in English.

While understanding the orthography of each language is essential to learn how to read, the corresponding phonology also plays an important role in literacy acquisition. For instance, the triangle model (Seidenberg and McClelland, 1989; Harm and Seidenberg, 2004) suggests a cooperation between orthography and phonology to read words. Nevertheless, exposure to the phonology of both languages can lead to cross-linguistic interferences between first language (L1) and second language (L2; Akamatsu, 2003; Lemhöfer et al., 2008; Sun-Alperin and Wang, 2008; Deacon et al., 2009; Ota et al., 2010; Howard et al., 2012; Bhide, 2015). As posited by the language non-selective lexical access hypothesis (Dijkstra and van Heuven, 2002), lexical and sublexical information from both languages is coactivated during word reading. The strength of these influences depend on variables like exposure (Brysbaert et al., 2017), amount of use (Flege et al., 1997; Luk and Bialystok, 2013), proficiency in L1, L2, or both languages (Haigh and Jared, 2007; Van Hell and Tanner, 2012), age (Howard et al., 2012), and the specific orthography (Beauvillain, 1992; Bialystok et al., 2005a; Hamada and Koda, 2008; Lemhöfer et al., 2008; Sun-Alperin and Wang, 2008; Ota et al., 2010; Lallier and Carreiras, 2018) and phonology (Sun-Alperin and Wang, 2008; Ota et al., 2009, 2010) of the L1 and L2 languages. Confusion between decoding rules (e.g., reading an English word by applying Spanish phonological rules) is likely to influence bilingual readers when the languages differ in terms of orthographic depth (Goswami et al., 1998). Many authors suggest that early phonological activation of both L1 and L2 phonological codes overlap during reading (Jared and Szucs, 2002; Duyck, 2005; Jared et al., 2012). This overlap of the two languages happens even in skilled readers that rely on lexical strategies (Perfetti and Bell, 1991; Grainger et al., 2005; Braun et al., 2009).

The pseudohomophone effect provides consistent evidence of phonological activation during reading. Pseudohomophones are non-words that sound like real words (e.g., pseudohomophones of the real English word cheese would be /chease/ or /chis/). Pseudohomophones are orthographically different from words, but phonologically equivalent. In native speakers, pseudohomophones yield faster responses in naming, which reflects a facilitating effect of familiar pronunciations (McCann and Besner, 1987; Seidenberg et al., 1996; Goswami et al., 2001). In addition, pseudohomophones delay responses in lexical decision tasks; since they sound like real words it is more difficult to discard them efficiently (McCann et al., 1988; Seidenberg et al., 1996; Goswami et al., 2001; Pexman et al., 2001; Ziegler et al., 2001; Briesemeister et al., 2009).

The pseudohomophone effect can be explained by computational models of visual word recognition like the multiple read-out model (MROM-p; Jacobs et al., 1998) or the dual-route cascaded model (DRC; Coltheart et al., 2001). In the MROM-p, a stimulus is rejected as a non-word when a threshold is not reached within a certain amount of time. During the processing of a pseudohomophone, there is a mismatch in the activation of the phonological and orthographical nodes, which requires a readjustment that results in delays in the response. The DRC, implemented with the MROM-p, is based on the double-route model (Coltheart, 1978). According to this model, activation in early modules flows to later modules, which receives excitation or inhibition from feedback pathways. In this model, a pseudohomophone activates a lexical entry in the phonological lexicon that does not match with any input in the orthographical lexicon, producing an incongruity. Both models describe a conflict between the "real word" phonological information and the "non-word" orthographical information. Readers are able to resolve this conflict, but the time needed to do so results in delayed responses.

As it happens in monolinguals, the pseudohomophone effect also results in a processing advantage (naming) or disadvantage (lexical decision) in second language readers. In lexical decision tasks, cross-lingual pseudohomophones rely on phonological transference across languages (Duyck, 2005). The phonological activation of a real word in either language competes with the orthographical activation of a non-word. In the case of bilinguals, the coactivation of L1 and L2 phonologies must be handled by activating the target language and inhibiting the non-target language (Grainger and Dijkstra, 1992; van Heuven et al., 1998). Thus, pseudohomophones can have the phonology-to-orthography correspondences of the target (/dreem/ for dream) or the non-target (/drim/for dream) language of the bilingual.

To date, research about pseudohomophone interference effects in second language learners of English has focused mainly on native speakers of orthographies like Dutch or French (Nas, 1983; Duyck, 2005; Haigh and Jared, 2007; Jared et al., 2012; Commissaire et al., 2019). These authors describe pseudohomophone effects as a result of the coactivation of both languages. However, Dutch and French orthographies are not as shallow as Spanish (Seymour et al., 2003), so there is no information about how readers of more shallow orthographies behave when learning to read using a deeper orthography. The present investigation is designed to provide new insights on this topic.

Furthermore, most of the research of pseudohomophones in second language learners has been conducted in adult populations (Nas, 1983; Haigh and Jared, 2007), with a smaller number investigating teenagers (Commissaire et al., 2019) or children (Jared et al., 2012). These studies did not systematically evaluate the developmental evolution of bilingual reading acquisition. Pseudohomophone effects might not emerge in beginner readers because their orthographic representations are not formed yet. In those without orthographic representations, the conflict between phonological and orthographical information would not exist, and therefore, the incongruity that leads to a delayed response would not emerge. Changes across grades in literacy patterns have been documented in Spanish children learning English as a second language (Howard et al., 2012; Hevia-Tuero et al., 2021), and just before middle-childhood, there is a key period in which children are proficient enough to rely on lexical retrieval and they depend less on sublexical decoding (Rau et al., 2014). Nevertheless, there is no information about how this pattern may affect performance in a pseudohomophone task.

Differences between languages may lead to different reading strategies during literacy acquisition, especially when native language orthography is shallower (Spanish), and second language orthography is deeper (English). A better understanding of the factors that affect word recognition across languages with different orthographies will lead to better approaches to reading instruction in second language learners. The present investigation contributes to the literature by measuring the effect of phonological cross-linguistic interferences in Spanish children learning English (a deeper orthography language). Studies that have investigated how Spanish influences English in second language learners have focused on vocabulary, morphological awareness, reading-aloud, or spelling (Zutell and Allen, 1988; Fashola et al., 1996; Sun-Alperin and Wang, 2008; Howard et al., 2012; Goodwin et al., 2015). To our knowledge, this is the first study to investigate the effects of L1 and L2 phonology in Spanish children learning to read in English.

This research has numerous educational implications. Instructional methods influence bilingual children reading abilities (Bialystok et al., 2005b). Depending on the school's characteristics, instructional methods expose children to different amounts of oral and written input in their different languages. For instance, reading skills in first language are important (Cummins, 1979; Maurer et al., 2021), but the amount of input received in second language also has a strong impact on reading proficiency (Matusevych et al., 2017; Mahmoud Al-Zoubi, 2018). Increased exposure to a language would mean more opportunities to process words, which may facilitate the formation of orthographic representations, as well as consolidate grapheme-to-phoneme correspondences. In Spain, there are different approaches to help children become proficient in English; however, not all of them seem to be successful (Martínez Agudo, 2019). Developing empirically validated instructional methods that are effective at teaching children to understand and read English are essential (Freeman and Freeman, 2006).

A novelty of the present study is that we measured participants responses using the computer software MouseTracker (J. Freeman and Ambady, 2010). The mouse-tracking paradigm has been extensively used in psycholinguistics research (Spivey et al., 2005; Barca and Pezzulo, 2015; Incera and McLennan, 2016; Incera, 2018). In line with previous research, the mouse-tracking paradigm measures errors and reaction times, so direct comparisons with other studies can be performed. In addition, it measures mouse trajectories (i.e., participant's computer mouse movements as they respond to the task), which provide detailed information about the online decision-making processes taking place. Through the analysis of x-coordinates over time (how close the mouse is from the correct response) it is possible to visualize the slope of the mouse trajectory. Steeper mouse trajectories mean that responses are more efficient (the computer mouse moves faster/straighter toward the correct response). Less steep mouse trajectories mean that responses are less efficient (the computer mouse moves slower/deviates more when moving toward the correct responses).

In the present investigation, children responded to a visual lexical decision task that included English words (dream), pseudohomophones following Spanish (L1) phonological rules (drim), and pseudohomophones following English (L2) phonological rules (dreem). Children were asked to click on the green tick when reading a real word and to click on the red cross when reading a string of letters that was not an English word. Clicking on the red cross (non-word) when reading a pseudohomophone is likely to take additional time, as children would be activating the real word phonology and the incorrect orthography. Thus, using the mouse-tracking paradigm we expect responses to pseudohomophones to result in more errors, slower reaction times, and less efficient mouse trajectories for children with less English proficiency (younger children, children attending the monolingual school). We want to determine the extent to which school (monolingual, bilingual) and grade (2nd, 4th, and 6th) influence second language reading.

The present study is the first to investigate the combine effects that grade and school have on the phonological development of Spanish children learning to read in English. Grade is an important factor to consider, as reading processes quickly evolve during the elementary school years. Furthermore, instructional method is likely to have a big impact on the ability of Spanish children to read in English. While all children in Spain are required to learn English, those in schools with bilingual instructional methods are likely to be exposed to English more often than those in other schools. For each of the three types of stimuli (English Words, Pseudohomophones following Spanish phonological rules, and Pseudohomophones following English phonological rules), our predictions are:

1. Children in higher grades will perform better than children in lower grades (Main effect of Grade).
2. Children in the bilingual school will perform better than children in the monolingual school (Main effect of School).
3. The effect of school (better performance in the bilingual school) will be larger for children in higher grades (Grade by School Interaction).

MATERIALS AND METHODS

Participants

Spanish native children from second, fourth, and sixth grade who attended two different types of schools participated in the study. All the schools that agreed to participate in the experiment were located in Spain, and they declared having a Spanish–English bilingual learning program. They were similar in terms of educational approaches during lessons taking place in Spanish. However, distinct instructional methods with respect to English were applied, and the hours per week that children were exposed to English differed. Henceforward, we will refer to them as monolingual (with less exposure to English) and bilingual (with more exposure to English) schools.

Monolingual School

In the monolingual school type, all the staff are Spanish native speakers. Children attend 4 h of English lessons per week and follow a Content and Language Integrated Learning methodology (CLIL; Martínez Agudo, 2019). Lessons of two other subjects, which vary depending on the grade (e.g., arts or science), also take place in English. Children are exposed to oral English during kindergarten stages through songs and letter names learning, but English instruction begins to place value on grammar and written vocabulary at Elementary levels. No specific reading instructional method is followed for English.

Bilingual School

The bilingual school type has some native English speakers as staff members. Lessons are taught 50% of the time in Spanish and 50% of the time in English. The instructional method emphasizes oral communication during English lessons. During kindergarten stages, children learn phonics, with explicit

TABLE 1 | Age and sex per school and grade.

School	Grade	Age Mean (SD)	Sex
Monolingual	Second	7.74 (0.29)	15 F/14 M
	Fourth	9.61 (0.28)	13 F/16 M
	Sixth	11.63 (0.28)	14 F/14 M
Bilingual	Second	7.67 (0.30)	12 F/15 M
	Fourth	9.61 (0.27)	15 F/14 M
	Sixth	11.63 (0.28)	14 F/14 M

instruction of phonological correspondences and decoding skills. Teaching of foundation skills of reading continues in later stages, where reading and writing is combined with oral communication.

The sample included 168 participants between 7 and 12 years old ($M_{age} = 9.60$; $SD_{age} = 1.60$). Children were randomly recruited from both types of school, and samples were equivalent −84 from the monolingual and 84 from the bilingual school. Across both types of school, the sample included 54 children from second grade (27 males and 27 females), 58 children from fourth grade (30 males, 28 females), and 56 children from sixth grade (28 males and 28 females; see **Table 1**). All of them had Spanish as their native language, and they had been studying English for at least 4 years by the time of data collection. None of the participants had cognitive or behavioral impairments. Children from both types of school were socio-economically equivalent.

Materials

A total of 24 words were selected, avoiding cognates and words that could be similar in Spanish and English. The mean length was 4.54 ($SD = 0.72$) characters and the mean frequency was 55,722 according to the Subtlex-UK database (van Heuven et al., 2014). Each word (e.g., *cheese*) was manipulated in order to create a pseudohomophone with a transcription that followed Spanish phonological rules (e.g., *chis*), and a pseudohomophone with a transcription that followed English phonological rules (e.g., *chease*). Four different versions of the experiment were created in order to counterbalance the stimuli across conditions. Every participant answered to all words, but within each version of the experiment, each word appeared only in one format (word, Spanish pseudohomophone, English pseudohomophone, illegal non-word). Furthermore, stimuli were randomly presented and the position of the response options was counterbalanced. For half the participants, the "it is a word" response (green tick image) was placed on the top left corner of the screen, while for the other half the correct response was placed on the top right corner of the screen (see **Figure 1**). Each participant responded to 42 trials (six baseline trials, six words, six English pseudohomophones, six Spanish pseudohomophones, six illegal non-words, and 12 filler words) for a total of 7,056 observations. The illegal version of each word and other English words were included as fillers. This was necessary to balance the amount of trial types answered by each participant (same amount of real word/non-word trials).

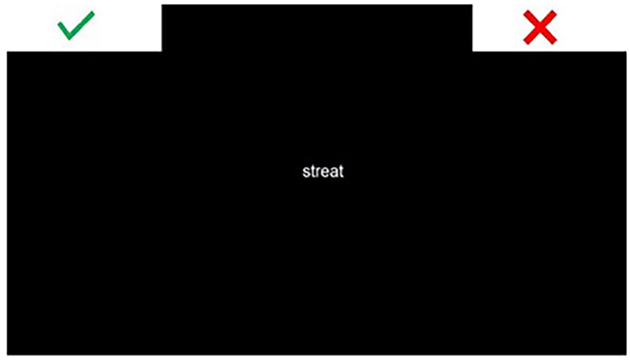


FIGURE 1 | Screenshot of the participants' view of the lexical decision task.

Procedure

The task was created with the computer software MouseTracker (Freeman and Ambady, 2010). An HP x360 Stream laptop was used to present the stimuli to the participants. Participants were asked to answer using a computer mouse and a large mouse pad (17.8 by 15.5 inches). Participants were tested individually, and performance feedback was not provided. Testing took place in a room free of noise and distracting elements to ensure the accuracy of the results. Each participant was randomly assigned to one of the eight versions of the experiment (to counterbalance the stimuli type and the response position).

Before the experiment, children were asked to complete a baseline task (Incera and McLennan, 2018; Hevia-Tuero et al., 2021). Non-linguistic trials (click on the smiley face on the top right or left corners of the screen) were included as a baseline motor task to measure the basic mouse movement abilities of the children. Furthermore, training trials were presented with the purpose of familiarizing the children with the computer program and the task before presenting the target trials.

At the beginning of each trial, START appeared at the bottom-center of the screen and the response options appeared on the top left and right corners. The written word or non-word was displayed in the center of the screen as soon as participants clicked START. The stimuli remained on the screen until participants clicked on one of the two response alternatives (green tick for real words, red cross for non-words). Children were told to click on one of the two response options as quickly and accurately as possible. Once they answered, the START button appeared and they had to click on it to initiate the next trial. If participants took more than 750 milliseconds to initiate a mouse movement, a warning appeared instructing them to start moving the mouse earlier on in future trials.

Analysis Plan

R-software (version 4.0.2) was used to run the mixed model analyses using the lme4 package (version 1.1–21; Bates et al., 2015). To analyze number of errors, we combined the advantages of ordinary logit models with the ability to account for random subject and item effects (Jaeger, 2008). The independent variables included in the analyses were grade (2nd, 4th, and 6th) and school (monolingual, bilingual). We performed separate analyses for each of the three types of stimuli in the lexical decision task: Words, Pseudohomophones following Spanish phonological rules, and Pseudohomophones following English phonological rules. The dependent variables included in the analyses were number of errors, reaction times, and mouse trajectory (x-coordinates over time).

The MouseTracker program measures participants' mouse positions over time, which includes three variables: *y*-coordinates, *x*-coordinates, and time (in milliseconds). Since three-dimensional graphs are hard to visualize, the standard in the field is to report *x*-coordinates over time [see (Incera, 2018), for a detailed discussion of methodological concerns and practical recommendations when using the mouse-tracking paradigm with bilingual populations]. While all participants move the mouse upwards (START is at the bottom and the response options are at the top of the screen) the way in which the task is set up results in the manipulation influencing whether participants move right or left (toward the response options on the right or left corner). Thus, we report mouse trajectories as *x*-coordinates over time.

Outliers were filtered, deleting correct responses with reaction times over and under 2 SD for each school, grade, and type of stimuli. First, we performed the Grade by School analysis on the baseline, in order to determine whether children in both schools are equivalent at the motor level. The baseline analysis does not include the random effect of items because all trials are the same (at baseline there is no item variability to account for). Second, we performed the Grade by School analysis on words, pseudohomophones following Spanish rules, and pseudohomophones following English rules. The goal was to test the effect of Grade (children in higher grades perform better), the effect of School (children in the bilingual school perform better), and the Grade by School interaction (the effect of school—bilingual better—is larger in higher grades). Random effects of participants and items were included crossed in all models testing Words, Spanish Pseudohomophones, and English Pseudohomophones. Models were compared using the Chi-square test; only factors that significantly contributed to model fit, as determined by a significant value of *p* in the chi-square test, were included in the final model. The estimate (effect size) and standard error of each effect was reported for all factors included in the final model for each dependent variable.

RESULTS

The data and the R Notebook with the analyses can be found at the Open Science Framework.

Errors

Errors are calculated by counting the number of times children clicked on the incorrect response (red cross for words, green tick for pseudohomophones). Error analyses cannot be conducted for the baseline task since there are no errors; all children were able to click on the smiley face at the top right/left corner of the screen without making any mistakes.

When analyzing number of errors for *words*, model comparisons indicated that there was a main effect of Grade [$\chi^2_{(2)} = 80.44$, $p < 0.001$] and a main effect of School [$\chi^2_{(1)} = 13.32$, $p < 0.001$], in line with our first and second hypotheses. Furthermore, the Grade by School interaction [$\chi^2_{(2)} = 11.56$, $p = 0.003$] also improved model fit, in line with our third hypothesis. The final model for errors for words as modeled in R is as follows: Error ~ Grade*School + (1|Participant) + (1|Stimuli). The effect of Grade emerged because second graders had more errors than fourth (*Estimate* = −0.84, *SE* = 0.33) and sixth (*Estimate* = −2.24, *SE* = 0.39) graders. The interaction emerged because, while the number of errors in words was equivalent for monolingual and bilingual children in second grade (second grade monolinguals 42.59%; second grade bilinguals 41.97%), the monolingual children had more errors than the bilingual children in fourth (fourth grade monolinguals 28.16%; fourth grade bilinguals 9.19%; *Estimate* = −1.70, *SE* = 0.52) and sixth (sixth grade monolinguals 11.90%; sixth grade bilinguals 4.16%; *Estimate* = −1.28, *SE* = 0.62) grades (**Figure 2**).

In the *Spanish pseudohomophones* analysis, model comparisons indicated that there was a main effect of Grade [$\chi^2_{(2)} = 40.41$, $p < 0.001$]. However, there was no effect of School [$\chi^2_{(1)} = 0.04$, $p = 0.841$] and there was no Grade by School interaction [$\chi^2_{(2)} = 2.38$, $p = 0.303$]. The final model for errors for Spanish pseudohomophones as modeled in R is as follows: Error ~ Grade + (1|Participant) + (1|Stimuli). The effect of Grade emerged because children in second grade had more errors than children in fourth (*Estimate* = −0.90, *SE* = 0.26) and sixth (*Estimate* = −1.92, *SE* = 0.31) grades. In the *English pseudohomophones* analysis, model comparisons indicated that there was a main effect of Grade [$\chi^2_{(2)} = 11.13$, $p = 0.003$]. However, the effect of School [$\chi^2_{(1)} = 1.58$, $p = 0.208$] and the Grade by School interaction [$\chi^2_{(2)} = 0.917$, $p = 0.631$] did not emerge. The final model for errors for English pseudohomophones as modeled in R is as follows: Error ~ Grade + (1|Participant) + (1|Stimuli). While there were no differences between children in second and fourth grade (*Estimate* = −0.24, *SE* = 0.22), the effect of Grade emerged because there were differences between children in second and sixth grade (*Estimate* = −0.74, *SE* = 0.22), the older children had less errors (see **Figure 3**).

In sum, error analyses for Words supported Hypothesis 1 (Effect of Grade), Hypothesis 2 (Effect of School), and Hypothesis 3 (Grade by School Interaction). Furthermore, error analyses for Pseudohomophones supported Hypothesis 1 (Effect of Grade). However, the effect of School (Hypothesis 2) and the Grade by School Interaction (Hypothesis 3) did not emerge in error analyses for pseudohomophones. Children from both schools (monolingual, bilingual) were equally likely to consider the pseudohomophones incorrect.

Reaction Times

Reaction times were measured from the moment the stimulus appeared on the screen to the moment participants clicked

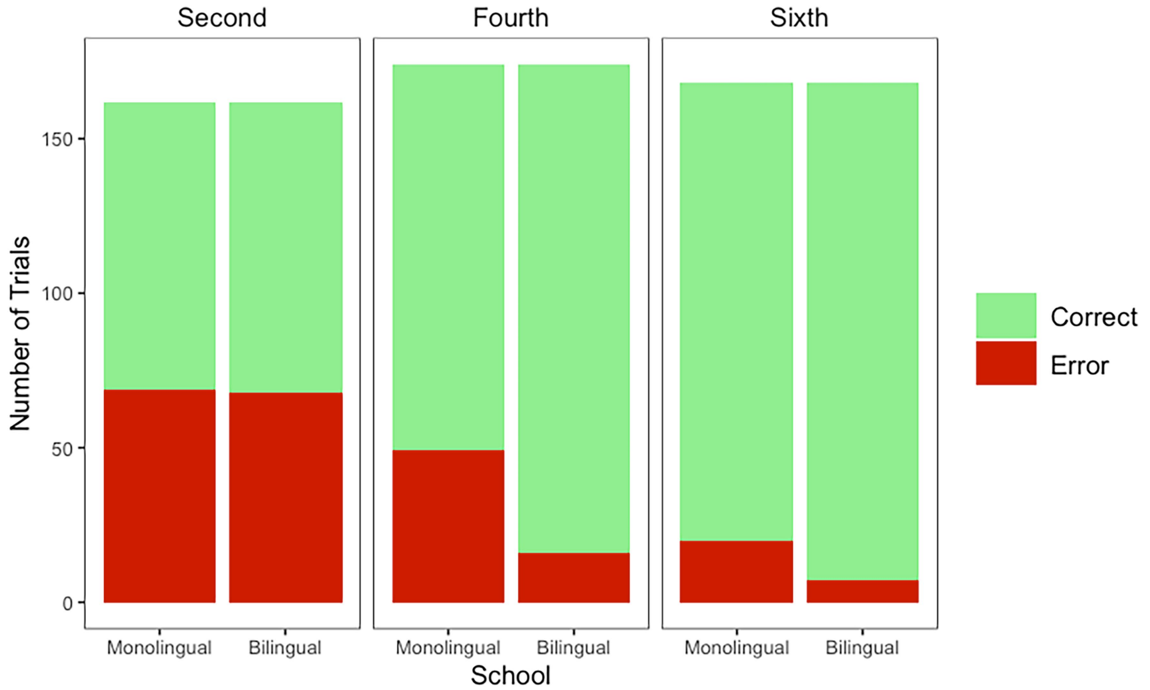


FIGURE 2 | Correct answers (green) and errors (red) per grade and school for words.

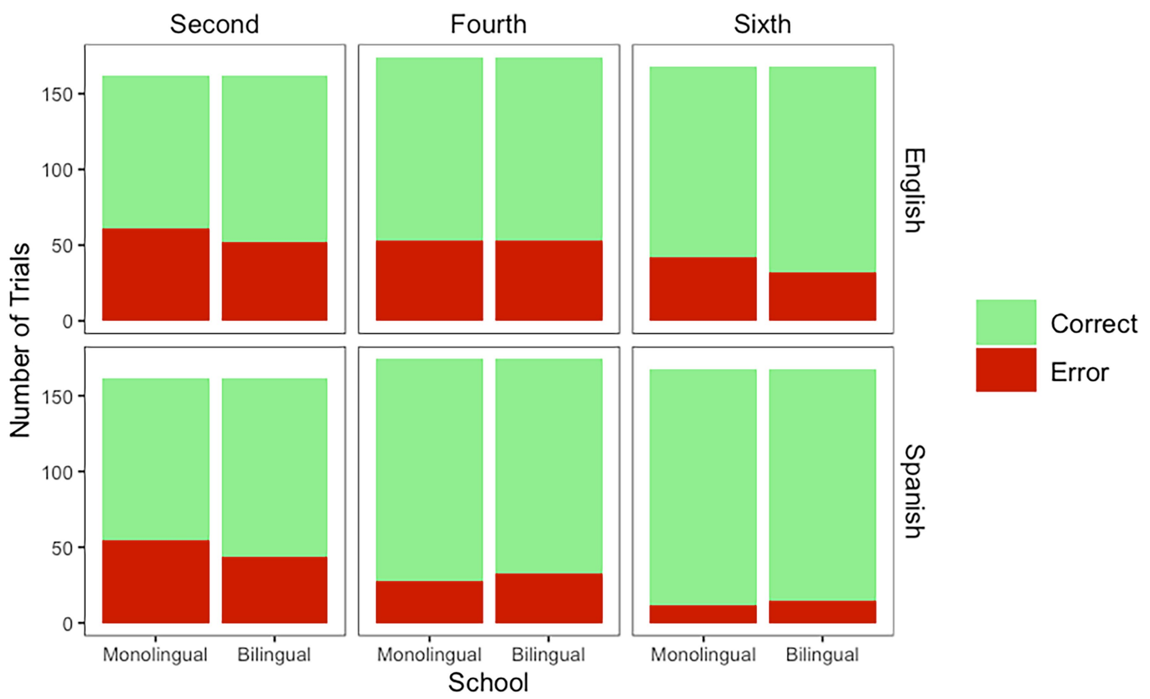


FIGURE 3 | Correct answers (green) and errors (red) per grade and school for Spanish and English pseudohomophones.

on the response. When analyzing the *baseline*, model comparisons indicated that there was a main effect of Grade [$\chi^2_{(2)} = 73.01$, $p < 0.001$] and a main effect of School [$\chi^2_{(1)} = 4.61$, $p = 0.031$]. The Grade by School interaction did not emerge [$\chi^2_{(2)} = 1.49$, $p = 0.472$]. The final model for reaction times for baseline as modeled in R is as follows: RT ~ Grade + School + (1|Participant).

TABLE 2 | Descriptive statistics (means and standard deviations) for reaction times responding to each condition (English Words, Spanish Pseudohomophones, English Pseudohomophones) per grade and school.

School	Grade	Baseline	Word	Spanish	English
Monolingual	Second	2,177 (886)	3,161 (1147)	3,441 (1203)	3,760 (1350)
	Fourth	1,579 (652)	2,484 (962)	2,619 (884)	2,698 (919)
	Sixth	1,293 (406)	1,942 (528)	2,133 (609)	2,293 (705)
Bilingual	Second	1,912 (679)	2,750 (1111)	3,256 (1303)	3,559 (1376)
	Fourth	1,510 (1016)	1,977 (546)	2,497 (881)	2,438 (718)
	Sixth	1,179 (450)	1,710 (433)	1,870 (452)	2,019 (519)

The effect of Grade emerged because second graders responded 499 ms (*SE* = 83) slower than fourth graders, and 808 ms (*SE* = 84) slower than sixth graders. The effect of School emerged because children from the bilingual school responded 147 ms (*SE* = 68) faster than children from monolingual school (see **Table 2**).

When analyzing *words*, model comparisons indicated that there was a main effect of Grade [$\chi^2_{(2)} = 86.60$, $p < 0.001$] and a main effect of School [$\chi^2_{(1)} = 17.61$, $p < 0.001$]. However, the Grade by School interaction did not emerge [$\chi^2_{(2)} = 1.91$, $p = 0.385$]. The final model for reaction times for words as modeled in R is as follows: RT ~ Grade + School + (1|Participant) + (1|Stimuli). The effect of Grade emerged because second graders responded 782 ms (*SE* = 109) slower than fourth graders, and 1,208 ms (*SE* = 108) slower than sixth graders. The effect of School emerged because children from bilingual school responded 374 ms (*SE* = 86) faster than children from monolingual school (see **Table 2**).

When analyzing *Spanish pseudohomophones*, there was a main effect of Grade [$\chi^2_{(2)} = 88.03$, $p < 0.001$]. The main effect of School [$\chi^2_{(1)} = 3.78$, $p = 0.051$] and the Grade by School interaction did not emerge [$\chi^2_{(2)} = 0.28$, $p = 0.868$]. The final model for reaction times for Spanish pseudohomophones as modeled in R is as follows: RT ~ Grade + (1|Participant) + (1|Stimuli). Overall, children took more than 3,000 ms to respond (*Estimate* = 3,400, *SE* = 96). Second graders were 780 ms slower than fourth graders (*SE* = 129) and 1,389 ms slower than sixth graders (*SE* = 129). When analyzing *English pseudohomophones*, there was a main effect of Grade [$\chi^2_{(2)} = 105.85$, $p < 0.001$] and a main effect of School [$\chi^2_{(1)} = 7.46$, $p = 0.006$]. The Grade by School interaction [$\chi^2_{(2)} = 0.03$, $p = 0.984$] did not emerge. The final model for reaction times for English pseudohomophones as modeled in R is as follows: RT ~ Grade + School + (1|Participant) + (1|Stimuli). Overall, children took more than 3,500 ms to respond (*Estimate* = 3,869, *SE* = 111). Second graders were 1,111 ms (*SE* = 129) slower than fourth graders and 1,555 ms (*SE* = 128) slower than sixth graders. Children attending a bilingual school were 286 ms (*SE* = 103) faster than children attending a monolingual school (see **Table 2**).

In sum, reaction time analyses for Words supported Hypothesis 1 (Effect of Grade) and Hypothesis 2 (Effect of School), but not Hypothesis 3 (Grade by School Interaction). Furthermore, reaction time analyses for Pseudohomophones supported Hypothesis 1 (Effect of Grade). Interestingly, the reaction time effect of School (Hypothesis 2) emerged in English but not in Spanish Pseudohomophones. Finally, the Grade by School Interaction (Hypothesis 3) did not emerge for Pseudohomophones.

Mouse Trajectories

Mouse trajectories are measured with *x*-coordinates over time. When analyzing the *baseline*, model comparisons indicated that on the slope of the mouse trajectory there was a main effect of Grade [$\chi^2_{(2)} = 48.10$, $p < 0.001$]. However, the main effect of School [$\chi^2_{(1)} = 1.53$, $p = 0.215$] and the Grade by School [$\chi^2_{(5)} = 2.26$, $p = 0.811$] interaction did not emerge. The final model for mouse trajectories for baseline as modeled in R is as follows: X100 ~ Time*Grade + (Time|Participant). The effect of Grade emerged on the slope of the mouse trajectories (Time*Grade) because, when compared to children in second grade, the mouse trajectories were steeper (better performance) for children in fourth (*Estimate* = −2.14, *SE* = 0.68) and sixth grade (*Estimate* = −5.09, *SE* = 0.68).

In *words*, model comparisons indicated that on the slope of the mouse trajectory there was a main effect of Grade [$\chi^2_{(2)} = 38.97$, $p < 0.001$] and a Grade by School [$\chi^2_{(5)} = 15.03$, $p = 0.010$] interaction. The main effect of School [$\chi^2_{(1)} = 1.43$, $p = 0.231$] did not emerge. The final model for mouse trajectories for words as modeled in R is as follows: X100 ~ Time*Grade*School + (Time|Participant) + (1|Stimuli). The effect of Grade emerged on the slope of the mouse trajectories (Time*Grade) because—compared to children in second grade—the mouse trajectories were steeper (better performance) for children in fourth (*Estimate* = 4.24, *SE* = 2.28) and sixth grade (*Estimate* = 12.30, *SE* = 2.27). The Grade by School interaction emerged on the slope of the mouse trajectories (Time*Grade*School) because the difference between the children attending the monolingual and the bilingual school was larger in fourth than second grade (*Estimate* = 4.56, *SE* = 3.19). However, the difference was smaller in sixth than second grade (*Estimate* = −2.74, *SE* = 3.20). While in sixth grade the children attending the bilingual school still outperformed the children attending the monolingual school (see **Figure 4**), this difference—the effect of school—was not as large in sixth as in fourth grade.

In *Spanish pseudohomophones*, model comparisons indicated that on the slope of the mouse trajectory there was a main effect of Grade [$\chi^2_{(2)} = 47.85$, $p < 0.001$]. However, the main effect of School [$\chi^2_{(1)} = 0$, $p = 0.994$] and the Grade by School [$\chi^2_{(5)} = 6.70$, $p = 0.243$] interaction did not emerge. The final model for mouse trajectories for Spanish pseudohomophones as modeled in R is as follows: X100 ~ Time*Grade + (Time|Participant) + (1|Stimuli). When responding to Spanish pseudohomophones, mouse trajectories were steeper for children in fourth (*Estimate* = 5.39, *SE* = 1.78) and sixth (*Estimate* = 13.23, *SE* = 1.79) grades. In *English pseudohomophones*, there was a main effect of Grade [$\chi^2_{(2)} = 46.27$, $p < 0.001$]. However, the main effect of School [$\chi^2_{(1)} = 0.02$, $p = 0.864$] or the Grade by School [$\chi^2_{(5)} = 3.18$, $p = 0.67$] interaction did not emerge. The final model for mouse trajectories for English pseudohomophones as modeled in R is as follows: X100 ~ Time*Grade + (Time|Participant) + (1|Stimuli). When responding to English pseudohomophones, mouse trajectories were steeper for children in fourth (*Estimate* = 9.27, *SE* = 1.94) and sixth (*Estimate* = 14.12, *SE* = 1.96) grades (see **Figure 5**).

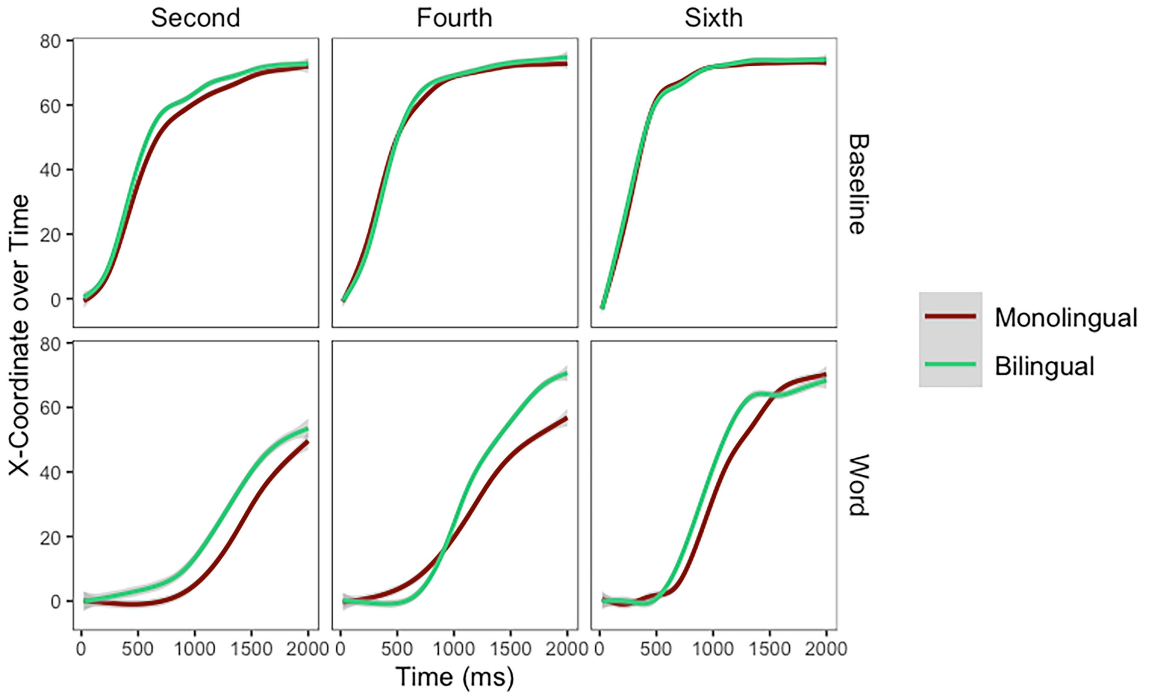


FIGURE 4 | Mouse trajectories for baseline and words per grade and school.

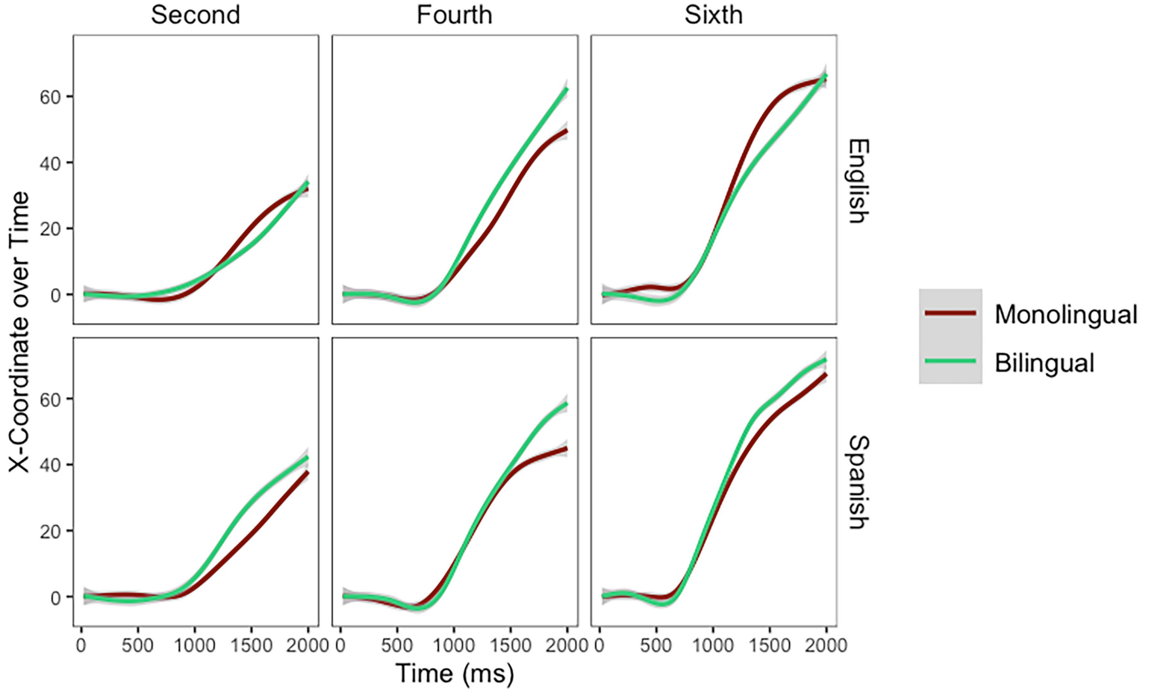


FIGURE 5 | Mouse trajectories for Spanish and English pseudohomophones per grade and school.

In sum, all analyses performed on the Mouse Trajectories supported Hypothesis 1 (Effect of Grade). However, Hypothesis 2 (Effect of School) did not emerge for words or pseudohomophones. Finally, results from Words support Hypothesis 3 (Grade by School interaction), but fourth grade (as opposed to sixth) is where the effect of School is largest.

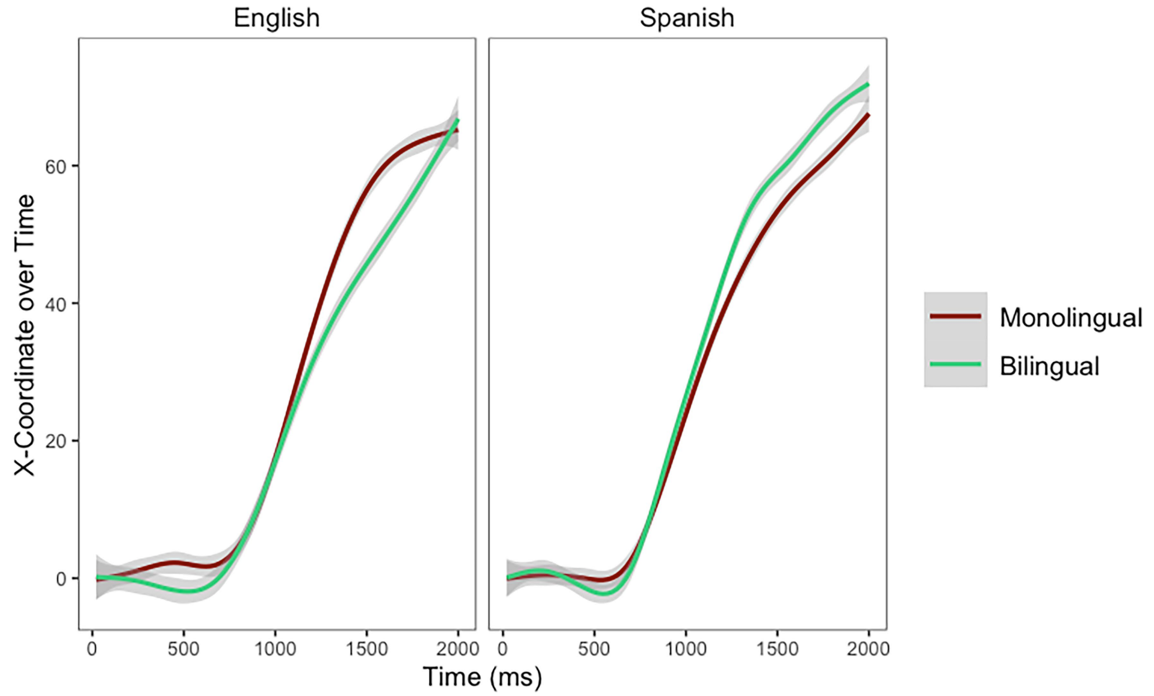


FIGURE 6 | Mouse trajectories for Spanish and English pseudohomophones per school in sixth grade.

Exploratory Analysis

The intriguing pattern of results for sixth graders in **Figure 5** (the children from the monolingual school seem to outperform the children from the bilingual school) led to an *exploratory analysis* performed on the slope of the mouse trajectories. This analysis focuses exclusively on children in sixth grade (the group where this interaction seems to emerge). The goal is to explore the potential Pseudohomophone (Spanish, English) by School (Monolingual, Bilingual) interaction in these skilled children. When sixth graders responded to the pseudohomophones, there was a main effect of Pseudohomophone [$\chi^2_{(1)} = 67.36$, $p < 0.001$] and a Pseudohomophone by School interaction [$\chi^2_{(3)} = 215.87$, $p < 0.001$]. The final model for this exploratory analysis as modeled in R is as follows: X100 ~ Time*Condition*School + (Time|Participant) + (1|Stimuli). The main effect of Pseudohomophone emerged because for all students English pseudohomophones were more distracting than Spanish pseudohomophones (*Estimate* = 0.28, *SE* = 0.39). However, the main effect of School [$\chi^2_{(1)} = 0.06$, $p = 0.805$] did not emerge. The cross-over interaction (see **Figure 6**) emerged because children from the bilingual school outperformed children in the monolingual school when responding to Spanish Pseudohomophones, while children in the monolingual school outperformed children in the bilingual school when responding to English pseudohomophones (*Estimate* = 4.99, *SE* = 0.55).

Close inspection of **Figure 6** indicates that early in the trajectory (around 500 ms after stimulus onset) children in the bilingual school are more distracted by both types of pseudohomophones than children in the monolingual school. Once participants start moving toward the correct response, the interaction emerges. Children in the bilingual school are able to outperform their counterparts in Spanish pseudohomophones, but children in the monolingual school outperform their counterparts in English pseudohomophones.

DISCUSSION

The aim of this study is to determine how native language (i.e., Spanish) interferes with second language reading (i.e., English); especially when L1 is shallower (phonemes and graphemes are more consistently linked) than L2. Importantly, we explored the extent to which grade (2nd, 4th, and 6th) and type of school (Bilingual, Monolingual) play a role in the acquisition of L2 grapheme-to-phoneme correspondence rules.

Words

In line with Hypothesis 1, older children performed better than younger children. Sixth graders had less errors, faster reaction times, and straighter mouse trajectories than fourth and second graders. In line with Hypothesis 2, children attending a bilingual school performed better (less errors, faster reaction times) than children attending a monolingual school. In line with Hypothesis 3, the effect of school (Spanish children attending a bilingual school having less errors) was larger in higher grades. In second grade, children attending the bilingual and the monolingual schools had similar number of errors (close to 40%) when recognizing English words. While students in the second grade performed above chance (60% accuracy), performance substantially improved in higher grades, especially for the children attending the bilingual school.

Our results from number of errors in words point to the conclusion that early on children in both schools perform equally, but as time passes children in the bilingual school outperform those in the monolingual school. Considering the MROM-p (Jacobs et al., 1998), word nodes are strongly activated in children attending a bilingual school, facilitating their responses in the lexical decision task and increasing processing speed in visual word recognition. The fact that the differences between bilinguals and monolinguals are larger in older children indicates that the effect of type of school is cumulative. Those in the bilingual school continue gaining advantages until at least sixth grade, as they are likely to have a higher level of English exposure (in particular, oral exposure).

Regarding the development of the orthography, higher exposure to English among bilingual children could have benefited their formation of strong orthographic representations along their schooling experience. These strong representations could have made their visual word recognition more accurate. An increase of English instructional time, especially the increase in oral instructional time, could have aided the bilingual children in consolidating the grapheme-to-phoneme correspondences. In addition, more opportunities to form orthographic representations are gained with more exposure to written words, which allows for a more efficient transition from serial phonological decoding to lexical processing (Share, 1995). This would be especially relevant for English reading acquisition due to the opacity of the English orthography. In English, phonological decoding is not enough to process words (Cunningham et al., 2002; Ziegler and Goswami, 2005). Finally, having more vocabulary is likely to facilitate recognition of a higher number of words, which makes readers more confident at rejecting non-words. This last possibility is supported by our results as children attending the monolingual school were less confident at rejecting non-words. The number of errors suggests that real English words were not recognized as such by the children in the monolingual school, likely because these children do not know these words yet.

Reaction times indicated that children from the bilingual school were faster than children from the monolingual school. A caveat to claim an advantage is that children in the bilingual school were also faster at baseline. In order to conclude that there are cognitive effects at play, the effect needs to be above and beyond that of the baseline. Indeed, the time difference between children attending the monolingual and bilingual school when responding to English words (374 ms) was more than double that the difference between these two groups at baseline (147 ms). Even though this effect needs to be considered cautiously—the two groups were not equivalent at baseline so the difference could be due (at least in part) to motor influences—the results indicate that the bilingual school has a positive effect on reading performance.

While baseline differences emerged in reaction times they did not emerge in mouse trajectories. Mouse trajectories showed that children were equivalent in terms of baseline mouse movements. When looking at the effect of school on the mouse trajectory, those attending the bilingual school were better at processing English words than those attending the monolingual school. Interestingly, this effect was largest in fourth grade. This is an important finding as it points to a time in development when the effect of School might be maximal (at least when measuring performance using a Lexical Decision task). Alternatively, it is possible that the task was too easy for the older children, thus the difference does not emerge because the sixth graders are performing at ceiling.

Pseudohomophones

In line with Hypothesis 1, when responding to pseudohomophones older children performed better than younger children. Second graders were more affected by the pseudohomophones than older students, probably because they do not have strong English orthographic representations. Second graders had recently started English literacy learning, and correspondence rules might not have been well established at this stage. Additionally, less expertise in L1 inhibition, joined to a lack of reading proficiency, are likely to result in less efficient reading performance. As Hamada (2017) observed, the influence of the native language phonology decreases when learners become more proficient in second language. A difference in knowledge of English phonological rules between children in the bilingual and the monolingual school could also explain these effects. There were no differences between schools with respect to errors in pseudohomophones. The nature of the cognitive processes at play (rejecting a non-word vs. accepting a word) is likely to have influenced these results. All children, even those in the bilingual school, took time and had doubts when rejecting the pseudowords and accidentally accepted some pseudowords as real words. Further research is necessary to determine what additional variables (e.g., oral versus written exposure) are influencing pseudohomophone effects in bilinguals.

While the main effect of school—children in the bilingual school outperforming children in the monolingual school—did not emerge in mouse trajectories, we observed a cross-over interaction. In line with Hypothesis 3, when responding to Spanish pseudohomophones children in the bilingual school outperformed children in the monolingual school. However, against Hypothesis 3, when responding to English pseudohomophones, children in the monolingual school outperformed their peers attending the bilingual school. In sixth grade, children attending a bilingual school are very efficient at rejecting Spanish pseudohomophones, but they get more distracted by the English pseudohomophones than children attending a monolingual school (**Figure 6**). The fact that children in the bilingual school are more confident discarding non-words that clearly follow Spanish rules than children in the monolingual school, support the idea that children in the bilingual school have more experience/practice inhibiting Spanish. The fact that children in the monolingual school get less distracted by the English pseudohomophones than children in the bilingual school indicate that their English phonology might not be as strongly developed.

Implications

Children learning a second language automatically activate L2-specific rules during word reading. All participants had more errors when responding to English than Spanish pseudohomophones (see **Figure 3**). When responding to an English task (in English mode), native speakers of Spanish were more distracted by the English than the Spanish phonology. This pseudohomophone effect is equivalent in other languages (Nas, 1983; Commissaire et al., 2019). In line with other studies, Spanish children develop knowledge of English phonology relatively early during development (Hevia-Tuero et al., 2021). In our study, the pseudohomophone effect emerged even in second grade, and not only in advanced L2 learners like Commissaire and colleagues had previously reported (2019). An emerging knowledge of English orthography is acquired at early stages, with relatively few years of instruction. These results support the idea that phonological information is activated in visual word recognition (Goswami et al., 2001; Ziegler et al., 2001).

The type of school children attend to (bilingual, monolingual) influences word processing. This effect could be due to higher levels of exposure to the second language or to a different approach to reading instruction. Different instructional methods might lead to different ways of processing, altering the orthography–phonology relationship. Indeed, instructional methods and native language characteristics influence reading strategies in both native and second language (Bhide, 2015). Furthermore, phonics instruction facilitates successful learning of relationships between letters and sounds, a requirement for learning to read (Castles et al., 2018). Not having been explicitly taught about English phonics, children in monolingual schools could be building orthographic representations without developing English phonological representations. These children could be relying on the lexical route or on Spanish phonological representations. In this way, they could be processing a whole word unit and rejecting a non-word based on orthographical characteristics. Results from the monolingual school coincide with what Pitts and Hanley (2010) found in their study: Spanish-speaking adults were less reliant on phonology than native speakers, despite knowing well the English grapheme-to-phoneme rules. These findings support the triangle model of cooperation between phonology and orthography to read words (Seidenberg and McClelland, 1989; Harm and Seidenberg, 2004). For those with less knowledge of phonology, a development of a direct orthography-to-semantics pathway would be reasonable (and advantageous in this task). Children attending a bilingual school have a foundation of phonic knowledge, and they are more familiar with English phonology. Therefore, they are likely to have a balanced division of labor. This approach is efficient in some situations (when phonology is helpful). However, the activation of the English phonology makes these bilingual children more sensitive to pseudohomophone effects. The type of school children attend influence their processing strategies during word recognition. There is a shift in the division of labor between the orthographic and the phonological component, which is likely to be influenced by how much written and oral exposure they have in their second language.

The more plausible explanation for our results is that English phonology plays a major role in the way that children in bilingual schools learn. When processing English pseudohomophones, the conflict between the existence of phonological information and the lack of orthographical information of a real word makes them move toward the correct response (rejecting the pseudohomophone) less efficiently. Although they have developed a better "rejection of Spanish" mechanism than children in the monolingual schools, they are still more distracted by the English phonology. These results connect to an increase on the activation of the English language node as described in BIA model (van Heuven et al., 1998) in children attending a bilingual school. Being aware of the stimuli language membership activates the language node. Moreover, for the children attending a bilingual school the higher level of exposure to English is likely to intensify language node activation.

The performance differences found between these instructional methods are remarkable. These results open the possibility for new research in L2 literacy instruction. The goal would be to better understand how instructional methods influence reading proficiency in each language, as authors like Rolla San Francisco et al. (2006) have suggested. Attending a bilingual school may strengthen English phonology activation. While this might constitute a disadvantage in a lexical decision task involving pseudohomophones, this is likely to be helpful when reading. This way of processing written text is closer to the "native" way of processing English words, which speaks to the good job bilingual schools are doing.

The current study is one of the few studies that have investigated L1 and L2 phonology interferences in Spanish children learning English. Moreover, this is the first study to use a pseudohomophone lexical decision task for this purpose. These findings support and complement previous research about phonological activation in second language learners during reading tasks. Furthermore, the present experiment adds to the literature on pseudohomophone effects in orthographic systems with different orthographic depths (shallower, like Spanish and Dutch; or deeper, like French and English). Our results are in line with those reported by Commissaire et al. (2019) and Nas (1983). Nas (1983) focused on adult Dutch L2 learners who had reach a proficient level of reading in their native language. Commissaire et al. (2019) studied adolescent French L2 learners of sixth and eighth grades. The novelty of our study is that participants started English instruction at an early age, and they learnt to read in both languages (L1 and L2) at the same time. In fact, the ages of the children participating in this study match the age for literacy foundation, which is another important contribution of the present investigation. Evaluating children across different grades allowed us to investigate the evolution of L1 and L2 during simultaneous reading learning, shedding light on the processes of literacy acquisition of English learners. However, we do not know to what extent our findings can be extrapolated to other populations of English learners, like Chinese or Hebrew speakers. Spanish and English share the same alphabet, which may have facilitated orthographic rule learning (Pasquarella et al., 2015). Future studies should address this issue, as cross-linguistic transfer is likely to be influenced by the proximity of L1 and L2 orthographies (Geva and Siegel, 2000; Chung et al., 2019).

The mouse-tracking paradigm allowed us to explore children's responses as they unfold over time. This methodology could be used in future studies to investigate automatic phonological activation during reading in tasks like visual masked priming using pseudohomophones (see Duyck, 2005; Ziegler et al., 2013; Sauval et al., 2017). Additionally, it would be interesting to focus on the effect of linguistic variables in order to broaden our knowledge of visual word recognition in L2 learners. Data focused on Spanish speakers learning English are scarce, despite the fact that English and Spanish are the first and fourth most commonly spoken languages in the world (Eberhard et al., 2020). Further investigations are needed to explore how reading mechanisms from the native language interfere with how children learn to read in their second language.

There are additional variables that could be taken into account when investigating these effects. Teachers were asked to select children with average reading skills, and children with difficulties were not included. This study did not assess Spanish and English reading skills, nor did it take into account domain-general abilities like inhibitory control (Bartolotti et al., 2011), which likely influence children's performance. The practical concerns of creating a study short enough for young children, while assessing a wide range of linguistic and cognitive skills, is a real challenge. Furthermore, data were collected during school hours, so students could not be absent from class too long. Additional variables related to the school are likely to influence children's performance. Some examples are the amount of time (only at school, also outside of school) and the type of exposure (oral versus written) to the language, the presence or absence of native speaker teachers, and the instructional methods used during pre-literacy stages. Together, these are factors that may be relevant for Spanish children learning English. It would be interesting to assess the specific weight of these variables in future studies, building on previous research (De Wilde et al., 2020).

CONCLUSION

In conclusion, the aim of this experiment was to understand how Spanish children learn to read in English. We found that Spanish children are able to recognize English orthography independently of their grade and the type of school they attend (monolingual, bilingual). Interestingly, differences in teaching methodologies—like an oral emphasis in bilingual schools versus a written emphasis in monolingual schools, as well as explicit phonics instruction—influence how L2 learners read. Spanish children in the bilingual school are more efficient at recognizing English words and discarding Spanish pseudohomophones, but get more distracted by English pseudohomophones. These results are in line with the idea that children in the bilingual school have better oral English (better English phonological representations) which makes them perform similar to the way in which native English speakers perform. The way in which learners are exposed to a second language determines how they process the orthography and phonology of their languages. Instructional methods influence the strength of the L1 and L2 inhibition processes.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary materials, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comité de Ética de la Investigación del Principado de Asturias. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

CH-T made contributions to the collection, analysis, interpretation of data for the work, and drafted the manuscript. SI and PS-C supervised the study and revised it critically for important intellectual content. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Ministry of Science and Innovation of Spanish Government (predoctoral grant number FPU18/03368 and grant number PID2019-106868GB-I00).

ACKNOWLEDGMENTS

We want to thank the students and staff from the schools that collaborated with the project, as well as all the families that provided consent.

REFERENCES

Akamatsu, N. (2003). The effects of first language orthographic features on second language Reading in text. *Lang. Learn.* 53, 207–231. doi: 10.1111/1467-9922.00216

Barca, L., and Pezzulo, G. (2015). Tracking second thoughts: continuous and discrete revision processes during visual lexical decision. *PLoS One* 10, 1–14. doi: 10.1371/journal.pone.0116193

Bartolotti, J., Marian, V., Schroeder, S. R., and Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Front. Psychol.* 2:324. doi: 10.3389/fpsyg.2011.00324

Bates, D., Mächler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67:1. doi: 10.18637/jss.v067.i01

Beauvillain, C. (1992). Orthographic and lexical constraints in bilingual word recognition. *Adv. Psychol.* 83, 221–235. doi: 10.1016/S0166-4115(08)61497-1

Bhide, A. (2015). Early literacy experiences constrain L1 and L2 reading procedures. *Front. Psychol.* 6, 1–16. doi: 10.3389/fpsyg.2015.01446

Bialystok, E., Luk, G., and Kwan, E. (2005a). Bilingualism, biliteracy, and learning to read: interactions among languages and writing systems. *Sci. Stud. Read.* 9, 43–61. doi: 10.1207/s1532799xssr0901_4

Bialystok, E., McBride-Chang, C., and Luk, G. (2005b). Bilingualism, language proficiency, and learning to read in two writing systems. *J. Educ. Psychol.* 97, 580–590. doi: 10.1037/0022-0663.97.4.580

Braun, M., Hutzler, F., Ziegler, J. C., Dambacher, M., and Jacobs, A. M. (2009). Pseudohomophone effects provide evidence of early lexico-phonological processing in visual word recognition. *Hum. Brain Mapp.* 30, 1977–1989. doi: 10.1002/hbm.20643

Briesemeister, B. B., Hofmann, M. J., Tamm, S., Kuchinke, L., Braun, M., and Jacobs, A. M. (2009). The pseudohomophone effect: evidence for an orthography-phonology-conflict. *Neurosci. Lett.* 455, 124–128. doi: 10.1016/j.neulet.2009.03.010

Brysbaert, M., Lagrou, E., and Stevens, M. (2017). Visual word recognition in a second language: A test of the lexical entrenchment hypothesis with lexical decision times. *Biling. Lang. Congn.* 20, 530–548. doi: 10.1017/S1366728916000353

Castles, A., Rastle, K., and Nation, K. (2018). Ending the Reading wars: Reading acquisition From novice to expert. *Psychol. Sci. Public Interest* 19, 5–51. doi: 10.1177/1529100618786959

Chung, S. C., Chen, X., and Geva, E. (2019). Deconstructing and reconstructing cross-language transfer in bilingual reading development: An interactive framework. *J. Neurolinguistics* 50, 149–161. doi: 10.1016/j.jneuroling.2018.01.003

Coltheart, M. (1978). "Lexical access in simple reading tasks" in *Strategies of Information Processing*. ed. G. Underwood (United States: Academic Press), 151–216.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., and Ziegler, J. C. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychol. Rev.* 108, 204–256. doi: 10.1037/0033-295X.108.1.204

Commissaire, E., Duncan, L. G., and Casalis, S. (2019). Investigating pseudohomophone interference effects in young second-language learners. *J. Exp. Child Psychol.* 180, 1–18. doi: 10.1016/j.jecp.2018.11.010

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. United States: Cambridge University Press.

Cronnell, B. (1985). Language influences in the English writing of third- and sixth-grade Mexican-American students. *J. Educ. Res.* 78, 168–173. doi: 10.1080/00220671.1985.10885594

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Rev. Educ. Res.* 49, 222–251. doi: 10.3102/00346543049002222

Cunningham, A. E., Perry, K. E., Stanovich, K. E., and Share, D. L. (2002). Orthographic learning during reading: examining the role of self-teaching. *J. Exp. Child Psychol.* 82, 185–199. doi: 10.1016/S0022-0965(02)00008-5

De Simone, E., Beyersmann, E., Mulatti, C., Mirault, J., and Schmalz, X. (2021). Order among chaos: cross-linguistic differences and developmental trajectories in pseudoword reading aloud using pronunciation entropy. *PLoS One* 16:e0251629. doi: 10.1371/journal.pone.0251629

De Wilde, V., Brysbaert, M., and Eyckmans, J. (2020). Learning English through out-of-school exposure. Which levels of language proficiency are attained and which types of input are important? *Biling. Lang. Congn.* 23, 171–185. doi: 10.1017/S1366728918001062

Deacon, S. H., Wade-Woolley, L., and Kirby, J. R. (2009). Flexibility in young second-language learners: examining the language specificity of orthographic processing. *J. Res. Read.* 32, 215–229. doi: 10.1111/j.1467-9817.2009.01392.x

Defior, S., and Serrano, F. (2005). The initial development of spelling in Spanish: From global to analytical. *Read. Writ.* 18, 81–98. doi: 10.1007/s11145-004-5893-1

Dijkstra, T., and van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Biling. Lang. Congn.* 5, 175–197. doi: 10.1017/s1366728902003012

Duyck, W. (2005). Translation and associative priming with cross-lingual pseudohomophones: evidence for nonselective phonological activation in bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 1340–1359. doi: 10.1037/0278-7393.31.6.1340

Eberhard, D. M., Simons, G. F., and Fennig, C. D. (2020). *Ethnologue: Languages of the World. 23rd Edn*. United States: SIL International.

Ehri, L. C. (2005). Learning to read words: theory, findings, and issues. *Sci. Stud. Read.* 9, 167–188. doi: 10.1207/s1532799xssr0902_4

Farukh, A., and Vulchanova, M. (2015). *L1, Quantity of Exposure to L2, and Reading Disability as Factors in L2 Literacy Skills*. Germany: Walter de Gruyter, GmbH.

Fashola, O. S., Drum, P. A., Mayer, R. E., and Kang, S.-J. J. (1996). A cognitive theory of orthographic transitioning: predictable errors in how Spanish-speaking children spell English words. *Am. Educ. Res. J.* 33, 825–843. doi: 10.3102/00028312033004825

Flege, J. E., Frieda, E. M., and Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *J. Phon.* 25, 169–186. doi: 10.1006/jpho.1996.0040

Fox, R. A., Flege, J. E., and Munro, M. J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *J. Acoust. Soc. Am.* 97, 2540–2551. doi: 10.1121/1.411974

Freeman, J., and Ambady, N. (2010). MouseTracker: software for studying real-time mental processing using a computer mouse-tracking method. *Behav. Res. Methods* 42, 226–241. doi: 10.3758/BRM.42.1.226

Freeman, Y. S., and Freeman, D. E. (2006). *Teaching Reading and Writing in Spanish and English in Bilingual and Dual Language Classrooms. 2nd Edn*. United States: Heinemann Portsmouth.

Frith, U. (1985). Beneath the surface of developmental dyslexia. *Surface Dyslexia* 32, 301–330.

Geva, E., and Siegel, L. S. (2000). The role of orthography and cognitive factors in the concurrent development of basic reading skills in bilingual children. *Read. Writ.* 12, 1–30. doi: 10.1023/A:1008017710115

Goodwin, A. P., August, D., and Calderon, M. (2015). Reading in multiple orthographies: differences and similarities in Reading in Spanish and English for English learners. *Lang. Learn.* 65, 596–630. doi: 10.1111/lang.12127

Goswami, U., Gombert, J. E., and de Barrera, L. F. (1998). Children's orthographic representations and linguistic transparency: nonsense word reading in English, French, and Spanish. *Appl. Psycholinguist.* 19, 19–52. doi: 10.1017/s0142716400010560

Goswami, U., Ziegler, J. C., Dalton, L., and Schneider, W. (2001). Pseudohomophone effects and phonological recoding procedures in reading development in English and German. *J. Mem. Lang.* 45, 648–664. doi: 10.1006/jmla.2001.2790

Grainger, J., and Dijkstra, T. (1992). On the representation and use of language information in bilinguals. *Adv. Psychol.* 83, 207–220. doi: 10.1016/S0166-4115(08)61496-X

Grainger, J., Muneaux, M., Farioli, F., and Ziegler, J. C. (2005). Effects of phonological and orthographic neighbourhood density interact in visual word recognition. *Q. J. Exp. Psychol.* 58, 981–998. doi: 10.1080/02724980443000386

Haigh, C. A., and Jared, D. (2007). The activation of phonological representations by bilinguals While Reading silently: evidence From Interlingual homophones. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 623–644. doi: 10.1037/0278-7393.33.4.623

Hamada, M. (2017). L2 word recognition: influence of L1 orthography on multi-syllabic word recognition. *J. Psycholinguist. Res.* 46, 1101–1118. doi: 10.1007/s10936-017-9483-1

Hamada, M., and Koda, K. (2008). Influence of first language orthographic experience on second language decoding and word learning. *Lang. Learn.* 58, 1–31. doi: 10.1111/j.1467-9922.2007.00433.x

Harm, M. W., and Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychol. Rev.* 111, 662–720. doi: 10.1037/0033-295X.111.3.662

Hélot, C., and Cavalli, M. (2017). "Bilingual education in Europe: dominant languages," in *Bilingual and Multilingual Education*. ed. O. García (New York: Springer International Publishing), 3–34.

Hevia-Tuero, C., Incera, S., and Suárez-Coalla, P. (2021). Does English orthography influence bilingual Spanish readers? The effect of grapheme crosslinguistic congruency and complexity on letter detection. *Cogn. Dev.* 59:101074. doi: 10.1016/j.cogdev.2021.101074

Howard, E. R., Green, J. D., and Arteagoitia, I. (2012). Can yu rid guat ay rot? A developmental investigation of cross- linguistic spelling errors among Spanish-English bilingual students. *Biling. Res. J.* 35, 164–178. doi: 10.1080/15235882.2012.703637.Can

Incera, S. (2018). Measuring the timing of the bilingual advantage. *Front. Psychol.* 9:1983. doi: 10.3389/fpsyg.2018.01983

Incera, S., and McLennan, C. T. (2016). Mouse tracking reveals that bilinguals behave like experts. *Biling. Lang. Congn.* 19, 610–620. doi: 10.1017/S1366728915000218

Incera, S., and McLennan, C. T. (2018). The time course of within and between-language interference in bilinguals. *Int. J. Biling.* 22, 88–99. doi: 10.1177/1367006916644688

Jacobs, A. M., Rey, A., Ziegler, J. C., and Grainger, J. (1998). "MROM-P: An interactive activation, multiple read-out model of orthographicand phonological processes in visual word recognition" in *Scientific Psychology Series. Localist Connectionist Approaches to Human Cognition*. eds. J. Grainger and A. M. Jacobs (United States: Lawrence Erlbaum Associates Publishers), 147–187.

Jaeger, T. F. (2008). Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. *J. Mem. Lang.* 59, 434–446. doi: 10.1016/j.jml.2007.11.007

Jared, D., Cormier, P., Levy, B. A., and Wade-Woolley, L. (2012). Cross-language activation of phonology in young bilingual readers. *Read. Writ.* 25, 1327–1343. doi: 10.1007/s11145-011-9320-0

Jared, D., and Szucs, C. (2002). Phonological activation in bilinguals: evidence from interlingual homograph naming. *Biling. Lang. Congn.* 5, 225–239. doi: 10.1017/s1366728902003024

Katz, L., and Feldman, L. B. (2017). Linguistic coding in word recognition: comparisons Between a deep and a shallow. *Int. Proc. Reading* 6:85. doi: 10.4324/9781315108506-4

Lallier, M., and Carreiras, M. (2018). Cross-linguistic transfer in bilinguals reading in two alphabetic orthographies: The grain size accommodation hypothesis. *Psychon. Bull. Rev.* 25, 386–401. doi: 10.3758/s13423-017-1273-0

Larson-Hall, J. (2008). Weighing the benefits of studying a foreign language at a younger starting age in a minimal input situation. *Second. Lang. Res.* 24, 35–63. doi: 10.1177/0267658307082981

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., and Zwitserlood, P. (2008). Native language influences on word recognition in a second language: A Megastudy. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 12–31. doi: 10.1037/0278-7393.34.1.12

Luk, G., and Bialystok, E. (2013). Bilingualism is not a categorical variable: interaction between language proficiency and usage. *J. Cogn. Psychol.* 25, 605–621. doi: 10.1080/20445911.2013.795574

Mahmoud Al-Zoubi, S. (2018). The impact of exposure to English language on language acquisition. *J. App. Ling. Lang. Res.* 5, 151–162.

Martínez Agudo, J. D. (2019). Which instructional programme (EFL or CLIL) results in better oral communicative competence? Updated empirical evidence from a monolingual context. *Linguist. Educ.* 51, 69–78. doi: 10.1016/j.linged.2019.04.008

Matusevych, Y., Alishahi, A., and Backus, A. (2017). The impact of first and second language exposure on learning second language constructions. *Biling. Lang. Congn.* 20, 128–149. doi: 10.1017/S1366728915000607

Maurer, U., Jost, L. B., Pfenninger, S. E., and Eberhard-Moscicka, A. K. (2021). Effects of German reading skills and bilingualism on early learning of English as a foreign language in primary school children. *Read. Writ.* 34, 2673–2689. doi: 10.1007/s11145-021-10166-1

McCann, R. S., and Besner, D. (1987). Reading Pseudohomophones: implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 14–24. doi: 10.1037/0096-1523.13.1.14

McCann, R. S., Besner, D., and Davelaar, E. (1988). Word recognition and identification: do word-frequency effects reflect lexical access? *J. Exp. Psychol. Hum. Percept. Perform.* 14, 693–706. doi: 10.1037/0096-1523.14.4.693

Nas, G. (1983). Visual word recognition in bilinguals: evidence for a cooperation between visual and sound based codes during access to a common lexical store. *J. Verbal Learn. Verbal Behav.* 22, 526–534. doi: 10.1016/S0022-5371(83)90319-5

Olulade, O. A., Jamal, N. I., Koo, D. S., Perfetti, C. A., LaSasso, C., and Eden, G. F. (2016). Neuroanatomical evidence in support of the bilingual advantage theory. *Cereb. Cortex* 26, 3196–3204. doi: 10.1093/cercor/bhv152

Ota, M., Hartsuiker, R. J., and Haywood, S. L. (2009). The KEY to the ROCK: near-homophony in nonnative visual word recognition. *Cognition* 111, 263–269. doi: 10.1016/j.cognition.2008.12.007

Ota, M., Hartsuiker, R. J., and Haywood, S. L. (2010). Is a FAN always FUN? Phonological and orthographic effects in bilingual visual word recognition. *Lang. Speech* 53, 383–403. doi: 10.1177/0023830910371462

Pasquarella, A., Chen, X., Gottardo, A., and Geva, E. (2015). Cross-language transfer of word reading accuracy and word reading fluency in Spanish-English and Chinese-English bilinguals: script-universal and script-specific processes. *J. Educ. Psychol.* 107, 96–110. doi: 10.1037/a0036966

Perfetti, C. A., and Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: evidence from backward masking and priming. *J. Mem. Lang.* 30, 473–485. doi: 10.1016/0749-596X(91)90017-E

Pexman, P. M., Lupker, S. J., and Jared, D. (2001). Homophone effects in lexical decision. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 139–156. doi: 10.1037/0278-7393.27.1.139

Pitts, B., and Hanley, J. R. (2010). Reading strategies in English by adults whose first language is Spanish. *Eur. J. Cogn. Psychol.* 22, 596–611. doi: 10.1080/09541440903060296

Rau, A. K., Moeller, K., and Landerl, K. (2014). The transition From sublexical to lexical processing in a consistent orthography: An eye-tracking study. *Sci. Stud. Read.* 18, 224–233. doi: 10.1080/10888438.2013.857673

Rolla San Francisco, A., Mo, E., Carlo, M., August, D., and Snow, C. (2006). The influences of language of literacy instruction and vocabulary on the spelling of Spanish-English bilinguals. *Read. Writ.* 19, 627–642. doi: 10.1007/s11145-006-9012-3

Sauval, K., Perre, L., Duncan, L. G., Marinus, E., and Casalis, S. (2017). Automatic phonological activation during visual word recognition in bilingual children: A cross-language masked priming study in grades 3 and 5. *J. Exp. Child Psychol.* 154, 64–77. doi: 10.1016/j.jecp.2016.10.003

Schmalz, X., Marinus, E., Coltheart, M., and Castles, A. (2015). Getting to the bottom of orthographic depth. *Psychon. Bull. Rev.* 22, 1614–1629. doi: 10.3758/s13423-015-0835-2

Seidenberg, M. S., and McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychol. Rev.* 96, 523–568. doi: 10.1037/0033-295X.96.4.523

Seidenberg, M. S., Petersen, A., MacDonald, M. C., and Plaut, D. C. (1996). Pseudohomophone effects and models of word recognition. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 48–62. doi: 10.1037/0278-7393.22.1.48

Seymour, P. H. K., Aro, M., and Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *Br. J. Psychol.* 94, 143–174. doi: 10.1348/000712603321661859

Share, D. L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition* 55, 151–218. doi: 10.1016/0010-0277(94)00645-2

Spivey, M. J., Grosjean, M., and Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proc. Natl. Acad. Sci. U. S. A.* 102, 10393–10398. doi: 10.1073/pnas.0503903102

Sun-Alperin, M. K., and Wang, M. (2008). Spanish-speaking children's spelling errors with English vowel sounds that are represented by different graphemes in English and Spanish words. *Contemp. Educ. Psychol.* 33, 932–948. doi: 10.1016/j.cedpsych.2007.12.005

Van Hell, J. G., and Tanner, D. (2012). Second language proficiency and cross-language lexical activation. In. *Lang. Learn.* 62, 148–171. doi: 10.1111/j.1467-9922.2012.00710.x

van Heuven, W. J. B., Dijkstra, T., and Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *J. Mem. Lang.* 39, 458–483. doi: 10.1006/jmla.1998.2584

van Heuven, W. J. B., Mandera, P., Keuleers, E., and Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Q. J. Exp. Psychol.* 67, 1176–1190. doi: 10.1080/17470218.2013.850521

Winsler, A., Díaz, M., Espinosa, L., and Rodríguez, J. L. (1999). When learning a second language does not mean losing the first: bilingual language development in low-income, Spanish-speaking children attending bilingual preschool. *Child Dev.* 70, 349–362. doi: 10.1111/1467-8624.t01-1-00026

Ziegler, J. C., Bertrand, D., Lété, B., and Grainger, J. (2013). Orthographic and phonological contributions to Reading development: tracking developmental trajectories using masked priming. *Dev. Psychol.* 50, 1026–1036. doi: 10.1037/a0035187

Ziegler, J. C., and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychol. Bull.* 131, 3–29. doi: 10.1037/0033-2909.131.1.3

Ziegler, J. C., Jacobs, A. M., and Klüppel, D. (2001). Pseudohomophone effects in lexical decision: still a challenge for current word recognition models. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 547–559. doi: 10.1037/0096-1523.27.3.547

Zutell, J., and Allen, V. (1988). The English spelling strategies of Spanish-speaking bilingual children. *TESOL Q.* 22, 333–340. doi: 10.5054/tq.201

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 08 April 2022
doi: 10.3389/fpsyg.2022.871349

The Interactive Model of L2 Listening Processing in Chinese Bilinguals: A Multiple Mediation Analysis

Yilong Yang[1,2]*, Guoying Yang[2] and Yadan Li[3,4]*

[1] Research Center for Linguistics and Applied Linguistics, Xi'an International Studies University, Xi'an, China, [2] School of English Studies, Xi'an International Studies University, Xi'an, China, [3] Ministry of Education (MOE) Key Laboratory of Modern Teaching Technology, Shaanxi Normal University, Xi'an, China, [4] Shaanxi Normal University Branch, Collaborative Innovation Center of Assessment Toward Basic Education Quality at Beijing Normal University, Xi'an, China

OPEN ACCESS

Edited by:
Yan Jing Wu,
Ningbo University, China

Reviewed by:
Wei Ren,
Beihang University, China
Choo Mui Cheong,
The University of Hong Kong,
Hong Kong SAR, China
Kelly A. Vaughn,
University of Texas Health Science Center, United States

***Correspondence:**
Yilong Yang
smartison@outlook.com;
yangyilong@xisu.edu.cn
Yadan Li
liyadan@snnu.edu.cn;
yadanli@outlook.com

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: 08 February 2022
Accepted: 15 March 2022
Published: 08 April 2022

Citation:
Yang Y, Yang G and Li Y (2022) The Interactive Model of L2 Listening Processing in Chinese Bilinguals: A Multiple Mediation Analysis. Front. Psychol. 13:871349. doi: 10.3389/fpsyg.2022.871349

Second language (L2) listening is a common challenge for language learners. It remains largely unknown how bilinguals process L2 listening. The literature has suggested an interactive model of L2 listening processing. However, few studies have examined the model from an experimental approach. The current study tried to provide empirical evidence for the interactive model of L2 listening processing in bilinguals by exploring the relationships among English spoken word segmentation (SWS), cognitive inhibition, cognitive flexibility, and L2 listening proficiency. The results showed positive associations among SWS, cognitive inhibition, cognitive flexibility, and L2 listening proficiency. Mediation analysis suggested that SWS might have a positive influence on L2 listening proficiency both directly and indirectly through cognitive inhibition and cognitive flexibility, respectively. These results imply that both bottom-up (reflected at SWS) and top-down (reflected at cognitive inhibition and flexibility) processes are engaged in bilinguals' L2 listening processing.

Keywords: L2 listening, spoken word segmentation, cognitive inhibition, cognitive flexibility, the interactive model

INTRODUCTION

Listening is an indispensable prerequisite for us to sustain effective communication. It constitutes forty-five percent of our total communication (Feyten, 1991). Unlike reading, it remains an obstacle for language learners to identify and segment L2 utterances into understandable segments during oral communication because there are no obvious signs, pauses, or punctuations signaled within a complete and fluent speech flow (Cole and Jakimik, 1980). Additionally, listening to a language that is rhythmically different from one's first language (L1) can be particularly challenging (Vandergrift, 2008), for it requires second language (L2) listeners' to carry out additional processes to overcome comprehension barriers (Flowerdew and Miller, 2005). Moreover, listening provides the basis for the development of the other main language skills, i.e., speaking, reading, and writing (Murphy, 1991; Vandergrift, 1997; Fotos, 2001; Snow, 2005; Hinkel, 2006). L2 listening, therefore, has been considered lying at "the heart of second language learning" (Vandergrift, 2007).

The Interactive Model of Second Language Listening

Given the importance of L2 listening, scholars have made continuous efforts to unravel the underlying mechanisms of L2 listening processing. Several models have been developed from both the linguistic and cognitive perspectives, namely, the bottom-up and top-down models of L2 listening processing.

The bottom-up model of L2 listening processing was first developed in the 1940s and 1950s under the influence of behaviorism. It follows the traditional idea that communication is a means of information transmission; listeners accrete each basic linguistic unit (e.g., individual sounds or phonemes) within the speech into increasingly larger meaningful units, e.g., clauses, sentences, or discourses (Vandergrift, 2011). The bottom-up listening processing involves decoding. It indicates that the listeners have to segment the speech into meaningful units during communication (Vandergrift, 2011). Studies have revealed that L2 learners are more prone to segment speech by invoking their L1 segmentation procedures (Cutler, 2000), and this phenomenon is more prominent in low L2 proficiency listeners (Goh, 2000; Graham, 2006). Evidence has also shown that listeners with lower L2 proficiency need to put more effort into the bottom-up processing than those with higher proficiency (Li et al., 2020; Yang et al., 2021). Field (2008) even ascribed the failure of L2 listening comprehension to the incorrect segmentation of speech by L2 listeners. These studies together imply that spoken word segmentation (hereafter SWS) is an important way of bottom-up processing in successful L2 listening. Although it is plausible that the pure bottom-up model explains the mechanism of language learners listening and combining discrete segments to form meaning, it cannot account for the situation that listeners may still achieve successful listening comprehension without having to identify every single word of the interlocutor's utterance.

Another well-established model is the top-down L2 listening model. It appeared in the 1980s under the influence of constructivism, following the concept that considers listening as a purpose-driven process (Flowerdew and Miller, 2005). The top-down listening processing includes three main categories of listening strategies, i.e., metacognitive strategies, cognitive strategies, and socio-affective strategies (Vandergrift, 2008). Previous studies have stressed the importance of metacognitive strategies as the chief listening strategies used by language learners to successfully comprehend L2 speech (Namaziandost et al., 2019). Another study has shown a significantly heavier use of metacognitive strategies by L2 listeners with higher proficiency than novice L2 listeners (Vandergrift, 1997). Therefore, metacognitive strategies seem to play a pivotal role in L2 listening. Since metacognitive strategies can be regarded as the behavioral output of cognitive control (Jansiewicz, 2008), cognitive control might play an important role in top-down listening processing. Cognitive inhibition and cognitive flexibility are two different yet correlated abilities of cognitive control (Miyake et al., 2000). They have also been considered as two critical abilities for L2 processing (e.g., Kieffer et al., 2013; Nouwens et al., 2016; Chang, 2020).

Cognitive inhibition is an active process of resisting extraneous or unwanted information that competes for neural resources due to the lack of sufficient capacity (Harnishfeger, 1995). During the process of SWS, irrelevantly activated lexical candidates must be inhibited by listeners to resolve lexical competition, thereby achieving correct and accurate listening processing (Norris et al., 1995). Bilingualism refers to the state of commanding two languages (Wilson and Mihalicek, 2011). According to the definition of bilingualism, bilinguals can be generally distinguished into balanced and unbalanced ones. The balanced bilinguals are those who acquired two languages simultaneously in their early childhood and can use both of their languages fluently. Unbalanced bilinguals are those who acquired their second language in their late childhood or adulthood without reaching the native-like level of proficiency (Vega-Mendoza et al., 2015). Studies have shown that no matter the types of bilingualism, both of the bilinguals' languages would be activated during lexical processing (Sunderman and Kroll, 2006). Therefore, bilinguals have to inhibit lexical competition both within- and cross-language when segmenting L2 speech. And this continuous practice may enhance L2 learners' ability of cognitive inhibition (Carlson and Meltzoff, 2008). These findings suggest that SWS may play a positive role in cognitive inhibition for L2 learners. In addition, previous literature has found that inhibitory control has a direct positive impact on L2 listening comprehension (Kim and Phillips, 2014), indicating that the ability to suppress irrelevant and competing stimuli is necessary for language learners to achieve the success of L2 listening processing.

Cognitive flexibility refers to the ability to shift perspectives, attention, and thinking flexibly based on changed circumstances (Diamond, 2013). During the process of SWS, listeners may exploit cognitive control to revise their miscomprehension of sentences (Novick et al., 2005), which would be caused by the activation of multiple and conflicting candidate representations (Weiss et al., 2009). Therefore, the demand for recurrent conflict monitoring and resolution in bilingual language processing is considered the likely source of bilinguals' cognitive advantage (e.g., Bialystok and Shapero, 2005; Kroll et al., 2012). These cognitive advantages of bilinguals may reflect increased cognitive flexibility (Teubner-Rhodes et al., 2016). These findings altogether imply a positive influence of SWS on cognitive flexibility for L2 learners. Moreover, in the study of L2 learners' behavioral strategy use in L2 listening, Murphy (1985) pointed out that the more proficient L2 listeners are more open and flexible, proved by their greater amount and more significant flexibility of strategy use in L2 listening. In line with the previous finding, Bacon (1992) found that L2 listeners' success in L2 listening appears to be related to the total use of various strategies and the flexibility in changing strategies. On the other hand, studies in L2 learners' underlying cognitive mechanisms have shown that cognitive flexibility and cognitive inhibition are, to some extent, correlated constructs (Miyake et al., 2000). Therefore, it is possible that cognitive flexibility may also be positively correlated with L2 listening processing.

It has been well-established that cognitive inhibition and cognitive flexibility function concurrently within L2 processing (e.g., Kieffer et al., 2013; Nouwens et al., 2016; Chang, 2020). During L2 listening processing, L2 learners may form multiple representations to predict the meaning of the utterance. However, the early prediction of the utterance may sometimes conflict with the later-arriving new information (Teubner-Rhodes et al., 2016). Therefore, it calls for the L2 learners to inhibit the prepotent and irrelevant representations and then revise their misinterpretations flexibly. Evidence has shown that cognitive flexibility is built on cognitive inhibition (Diamond, 2013).

The top-down model deciphers the mystery of listeners with a high level of L2 proficiency who comprehend utterances by flexibly applying heterogeneous metacognitive and cognitive strategies (Vandergrift, 2003), whereas falling short of providing sufficient evidence to unravel the approaches frequently used by less skilled listeners. Similar results have shown that top-down processes are more important for L2 learners with high proficiency than those with low proficiency (e.g., Li et al., 2020; Yang et al., 2021). However, L2 listening comprehension may fail if only the top-down process is initiated (Carrell and Eisterhold, 1983).

Therefore, the nature and defects of the aforementioned L2 listening models call for the systematic integration of both the bottom-up and the top-down models, i.e., the interactive model of L2 listening processing (Oxford, 1993; Rubin, 1994; Lynch, 1998, 2002; Mendelsohn, 1998). Vandergrift (2011) has claimed that the bottom-up and the top-down processes of L2 listening come into play together with each other and function independently. Orii-Akita (2014) found that the interactive model of L2 listening processing was more efficient than the pure bottom-up or top-down model in Japanese EFL university students. Even though very few studies have provided empirical evidence for the interactive model of L2 listening processing, the underlying cognitive mechanisms of this model remain largely unknown. Especially, there has been no empirical research investigating the roles of both cognitive inhibition and cognitive flexibility as top-down processes within the interactive model of L2 listening processing.

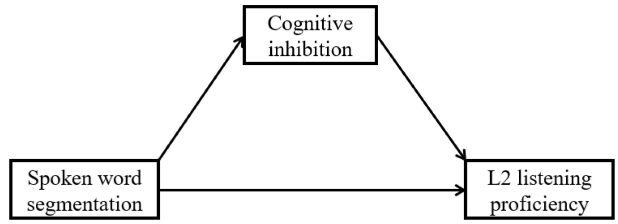


FIGURE 1 | The hypothesized mediation role of cognitive inhibition.

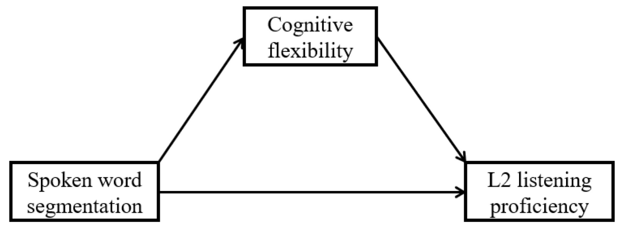


FIGURE 2 | The hypothesized mediation role of cognitive flexibility.

The Current Study

From the literature mentioned above, it can be concluded that SWS may play a positive role in L2 listening, cognitive inhibition, and cognitive flexibility. Additionally, previous studies have suggested that cognitive inhibition and cognitive flexibility may positively predict L2 listening as two different yet correlated variables. Furthermore, studies have highlighted an interactive model that integrates both the bottom-up and top-down processing during L2 listening. The present study aims to provide further empirical evidence for the interactive model of L2 listening processing. We hypothesized the following:

- **H1**: Cognitive inhibition mediates the relationship between L2 learners' SWS and L2 listening proficiency (**Figure 1**).
- **H2**: Cognitive flexibility mediates the relationship between L2 learners' SWS and L2 listening proficiency (**Figure 2**).

METHODS

Participants

One hundred and seventeen healthy volunteers joined the current study (26 males, 91 females, mean age: 19.38 ± 0.69 years). The participants were all Mandarin Chinese (L1) university students. English was reported as their L2. They were all unbalanced bilinguals who had acquired their L2 in their late childhood or adulthood. They all passed CET-4 (College English Test-Band 4), a national standardized English proficiency test for college students in China. The CET-4 lasts 125 min and measures test takers' comprehensive English abilities. The test vocabulary covers about 4,500 English words. Participants' language level of CET-4 ensures that they had qualified English language proficiency to do the language experiments in the current study, i.e., SWS and IELTS listening test. According to participants' self-report, they were all right-handed and had no neurological and psychiatric disorders or substance abuse. They had the normal or corrected-to-normal vision and normal hearing.

Measures

Spoken Word Segmentation

We adopted Cutler and Norris (1988)'s word-spotting paradigm to assess participants' performance in English SWS using E-Prime 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA, United States). We used a total of 96 stimuli of multisyllabic word strings (i.e., word plus nonsense syllable, such as *westej*, *lencool*) that consisted of real words (target word) and nonsense strings from Cutler and Shanley (2010) and Farrell (2015).

In a word-spotting task, participants would see a fixation cross on display for 8 s (**Figure 3**). They then would hear an audio stimulus (approximate duration 800-1,200 ms) played by two loudspeakers. Next, the participants had 3 s to identify the target word they had just heard in the audio stimulus. If they recognized the target word, they were required to make a keyboard response so that an additional 2 s would be given to them to speak out the target word (i.e., verbal response), e.g., speaking out English word *food* and *arm* in response to *foodeeb* and *armlek*. However, a new trial would start if the participant did not identify the target word in the audio stimulus within the time limit. We

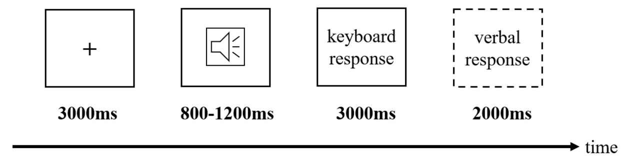


FIGURE 3 | Experiment procedure of spoken word segmentation.

used a digital audio recording pen to record participants' verbal responses, which were assessed after the experiment.

Several manipulations were performed to ensure the validity and reliability of the SWS test in the current participant sample. First, stimuli with target words in CET-4 test vocabulary were chosen to ensure that participants were familiar with target words. Second, target words of stimuli were 69 monosyllabic words and 27 multisyllabic words since Gitt (2006) suggested that modern English has 71.5% monosyllabic and 28.5% multisyllabic words. Third, the position of target words in stimuli was balanced. Half of the target words were in the initial position (e.g., *west* in *westej*). The other half were in the final position of the stimuli (e.g., *cool* in *lencool*). Fourth, we set two syllable boundary conditions (i.e., easy and difficult task conditions) following Cutler and Shanley (2010) and Farrell (2015). An easy task condition has unambiguous and easy to be identified word boundary, e.g., *dog* in *fubdog* and *arm* in *armlek*. In contrast, a difficult task condition has ambiguous and liaison word boundary, e.g., *agree* in *veamagree* and *food* in *foodeeb*. Fifth, we invited a female native speaker of American English to record the audio stimuli using a digital audio recording pen (44.1 kHz, 16 bit, mono). She was required to read in a continuous sequence and at a normal speed. We further used Cool Edit Pro 2.1 (Adobe Inc., San Jose, CA, United States) and Praat 6.1.04 (Boersma and Weenink, 2019) to process the audio stimuli.

Cognitive Inhibition

A Stroop color-word-interference task (hereafter Stroop task) was used to measure participants' cognitive inhibition in E-Prime 2.0 (e.g., Stroop, 1935; Miyake et al., 2000). The paradigm is a widely accepted and classical experimental test (e.g., Vendrell et al., 1995; Heidlmayr et al., 2014). In each trial, one of the four color words was displayed on the screen in its Chinese character of "red," "yellow," "blue," or "green" during the process of the task. These words were randomly presented either in a congruent or incongruent form with the colors red, yellow, blue, or green. Therefore, the color of the words was not matching the meaning of the words in some trials. At the beginning of each trial, a fixation cross was shown in the middle of the screen for 500 ms. Participants were then asked to identify the right color (instead of the meaning) of the stimulus words presented to them and to give their response by pressing the corresponding keys within a 2-second time limit. If they failed to respond within the time limit, a new trial would automatically follow. The Stroop task was conducted in Mandarin Chinese (i.e., participants' L1) to avoid the influence of participants' varied L2 competence and performance. Previous research has shown that the Chinese version of the Stroop task shares the same validity and reliability as its original version (Lee and Chan, 2000). In the present study, the Stroop task has 60 trials.

Cognitive Flexibility

Participants' cognitive flexibility was measured by the cognitive flexibility inventory (CFI; Dennis and Vander Wal, 2010). It was developed as a brief self-report measure of the cognitive flexibility necessary for individuals to successfully challenge and replace inappropriate thoughts with more balanced and appropriate thinking. The CFI consists of 20 items and is distributed on a 5-point Likert scale ranging from 1 ("never") to 5 ("very often"). The performance of participants' cognitive flexibility is indicated by the sum of the 20 items of the CFI. It has been well established in the literature that the CFI has sufficient reliability and validity for measuring cognitive flexibility (Cronbach's α = 0.91; Dennis and Vander Wal, 2010). It is also confirmed that the Chinese version of the CFI has satisfying reliability and validity as its English version (Cronbach's α = 0.83; Wang et al., 2016). Therefore, the scale's Chinese version was used in the present study.

Second Language Listening Proficiency

A listening test from *Cambridge English IELTS 9* (2013) assessed participants' L2 listening proficiency. IELTS (the International English Language Testing System) is an authentic and highly recognized English proficiency test. The IELTS listening proficiency test consists of 40 questions that are distributed in 4 sections. The questions ask the test takers to either choose the correct answers or fill in the blanks with no more than three words after listening to the test audio. The entire test took approximately 40 minutes. It was administered and scored by an experienced associate professor who strictly followed the test instructions and answer keys of the IELTS listening test (full score = 40). A question before the test showed that none of the participants had ever taken this test before.

Procedure and Statistical Analysis

The participants completed a demographic survey, an SWS task, a cognitive inhibition test, a cognitive flexibility questionnaire, and an L2 listening proficiency test. The demographic survey and cognitive flexibility questionnaire were distributed via an online survey platform[1]. The SWS task and cognitive inhibition test were performed using E-Prime 2.0. The L2 listening proficiency test was a paper-based test that was completed in a quiet room.

IBM Statistical Package for Social Sciences (SPSS) version 25.0 (SPSS Inc., Chicago, United States) was used for the descriptive statistics and correlation analyses. The PROCESS (v. 3.5) macro for SPSS was used to test our hypotheses (Preacher and Hayes, 2004; Hayes, 2013). Model 6 was used to test the hypothesized mediation effects, with a bootstrapping sample size of 5,000 and 95% confidence intervals (CIs). We set SWS as the independent variable, cognitive inhibition and cognitive flexibility as the mediation variables, and L2 listening proficiency as the dependent variable.

[1] http://www.wjx.cn

RESULTS

Descriptive Statistics and Correlation Analysis

Participants' performance in SWS (52.872 ± 6.670), cognitive inhibition (46.359 ± 5.255), cognitive flexibility (64.641 ± 7.588), and L2 listening proficiency (27.410 ± 5.238) were tested. **Table 1** shows the means of, standard deviations of, and correlations among those variables. SWS was in positive correlations with cognitive inhibition (r = 0.329), cognitive flexibility (r = 0.405), and L2 listening proficiency (r = 0.361). These results suggest that participants who performed better in SWS also had better performance in the cognitive inhibition, cognitive flexibility, and L2 listening tasks. Cognitive inhibition was in positive correlations with cognitive flexibility (r = 0.335) and L2 listening proficiency (r = 0.362), showing that participants who had better performance in cognitive inhibition tasks also showed greater ability of cognitive flexibility and higher L2 listening proficiency than their counterparts. Cognitive flexibility was positively correlated with L2 listening proficiency (r = 0.372), suggesting that participants who reported more advanced ability of cognitive flexibility also had higher L2 listening proficiency.

Mediation Analysis

We used the PROCESS (v. 3.5) extension for SPSS version 25.0 for mediation analyses. The multiple mediation analysis was performed to test the role of cognitive inhibition and cognitive flexibility in the association between SWS and L2 listening proficiency. The mediation model was significant and accounted for a significant proportion of the variance in explaining the relationship between SWS and L2 listening proficiency [R^2 = 0.233, F (3, 113) = 11.444, $p < 0.001$]. SWS had a positive influence on cognitive inhibition (β = 0.260, SE = 0.069, $p < 0.001$) and L2 listening proficiency (β = 0.157, SE = 0.073, $p < 0.05$). Meanwhile, cognitive inhibition had a positive influence on cognitive flexibility (β = 0.326, SE = 0.127, $p < 0.05$) and L2 listening proficiency (β = 0.223, SE = 0.089, $p < 0.05$). These results suggest that cognitive inhibition mediates the relationship between participants' performance in SWS and L2 listening proficiency. Moreover, SWS also had a positive impact on cognitive flexibility (β = 0.377, SE = 0.100, $p < 0.001$). Additionally, cognitive flexibility had a positive influence on L2 listening proficiency (β = 0.149, SE = 0.064, $p < 0.05$). These results suggest that cognitive flexibility mediated the relationship between participants' SWS performance and L2 listening proficiency. The results are shown in **Figure 4**.

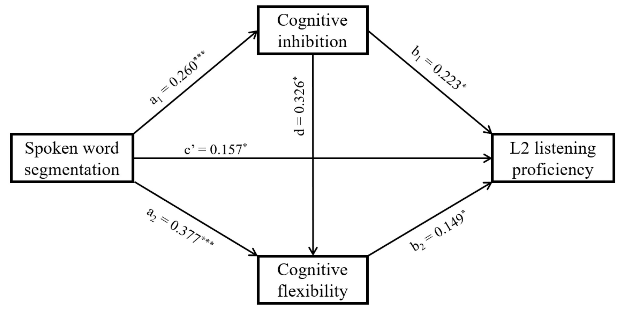


FIGURE 4 | Mediation roles of cognitive inhibition and cognitive flexibility in the association between SWS and L2 listening proficiency. The depicted is the path diagram of the multiple mediation analysis using the PROCESS (Model 6) macro for SPSS. *$p < 0.05$, ***$p < 0.001$.

TABLE 2 | Direct and indirect effects of SWS on L2 listening proficiency.

	95% CI	Effect
Direct effect		
SWS → L2 listening proficiency	[0.013, 0.300]	0.157
Indirect effect		
SWS → cognitive inhibition → L2 listening proficiency	[0.010, 0.117]	0.058
SWS → cognitive flexibility → L2 listening proficiency	[0.009, 0.121]	0.056
SWS → cognitive inhibition → cognitive flexibility → L2 listening proficiency	[−0.001, 0.030]	0.013

SWS, spoken word segmentation.

Furthermore, SWS had two paths of indirect influence on L2 listening proficiency, i.e., through cognitive inhibition (β = 0.058, SE = 0.028, 95% CI = [0.010, 0.117]) and cognitive flexibility (β = 0.056, SE = 0.029, 95% CI = [0.009, 0.121]) respectively. However, the influence of SWS on L2 listening proficiency through sequential effects of cognitive inhibition and cognitive flexibility is not significant (β = 0.013, SE = 0.008, 95% CI = [−0.001, 0.030]), suggesting that the chain mediation model is invalid. These results are shown in **Table 2**.

TABLE 1 | Means, standard deviations, and correlations between variables.

Variables	*M* ± *SD*	1	2	3	4
1. SWS	52.872 ± 6.670	1			
2. Cognitive inhibition	46.359 ± 5.255	0.329	1		
3. Cognitive flexibility	64.641 ± 7.588	0.405	0.335	1	
4. L2 listening proficiency	27.410 ± 5.238	0.361	0.362	0.372	1

SWS, spoken word segmentation.
All ps < 0.05.

DISCUSSION

Prior studies have suggested that non-native speakers are capable of using heterogenous segmentation cues in L2 listening (e.g., Sanders et al., 2002) and have emphasized the positive impact of SWS on the success of L2 listening processing (e.g., Field, 2003; Cutler and Shanley, 2010). Even though listening processing, as a complex cognitive activity, requires bottom-up linguistic skills, such as vocabulary, spoken word segmentation, and recognition, they cannot sufficiently account for the success of listening comprehension (Kim and Phillips, 2014). The current study provided empirical evidence for the interactive model of L2 listening processing by investigating the potential roles of cognitive inhibition and cognitive flexibility in the relationship between L2 learners' English SWS and L2 listening proficiency. The results of our study showed that SWS was in positive correlations with cognitive inhibition and L2 listening proficiency. In addition, SWS also had a significantly positive influence on cognitive flexibility. Further mediation analyses revealed that both cognitive inhibition and cognitive flexibility mediated the relationship between participants' SWS and L2 listening proficiency.

Spoken Word Segmentation and Second Language Listening Proficiency

The current study showed that L2 listeners' SWS might have a direct influence on bilinguals' L2 listening proficiency. SWS is a crucial and challenging bottom-up process exploited by language learners during L2 listening (Vandergrift, 2011). Previous research has pointed out that language learners might make good use of various linguistic cues to segment speech stream into meaningful linguistic units during the process of L2 listening (Goyet et al., 2010). L2 learners gradually combine the segmented linguistic units into larger chunks of units (e.g., from the phoneme level to the discourse level) to achieve successful L2 listening (Vandergrift, 2011). Therefore, the result of the current study is within our expectation that L2 learners who have better performance in SWS also exhibit higher L2 listening proficiency.

The Mediation Role of Cognitive Inhibition

The current study found that bilinguals' SWS might have a positive effect on their cognitive inhibition. Previous studies have revealed that during the process of SWS, both of the two languages of bilinguals would be activated simultaneously for spoken language processing (e.g., Marian et al., 2003; Marian and Spivey, 2003a,b; Blumenfeld and Marian, 2007; Canseco-Gonzalez et al., 2010). This process would lead to the activation of multiple candidate words that compete for recognition (McQueen et al., 1994). Bilinguals thus may need to adopt a certain mechanism of language control to inhibit the lexical competition of both within-language and cross-language to deal with the irrelevantly activated candidates to achieve the success of L2 listening processing (Green, 1998).

The result of the current study is in line with previous findings that participants who perform better in SWS also have better performance in the tasks of cognitive inhibition (e.g., Blumenfeld and Marian, 2011; Mercier et al., 2014). A possible explanation is that bilinguals might have to deal with high degrees of lexical competition in their daily communication, presumably caused by the co-activation of their two languages, especially if the communication is in their L2 (e.g., Weber and Cutler, 2004; Canseco-Gonzalez et al., 2010). With the continuous practice of segmenting L2 speech stream and inhibiting distracting candidate words in their daily life, bilinguals might have better performance in cognitive inhibition, thereby exhibiting top-down cognitive advantages over their monolingual counterparts.

We found that cognitive inhibition was in positive correlation with L2 listening proficiency. Cognitive inhibition is one of the top-down cognitive control mechanisms. It was confirmed to be conducive to academic achievements, such as math (e.g., Best et al., 2011) and reading (e.g., Best et al., 2011; Cartwright, 2012; Nouwens et al., 2016, 2021). The finding of our study was convergent with the literature mentioned above in other disciplines and extended previous research by showing that cognitive inhibition would also predict the success of L2 listening. It suggests that L2 learners' relatively more advanced cognitive inhibition ability than monolinguals would let them better resist the interference of irrelevant competing mental representation and candidate words. Moreover, L2 learners can inhibit the inclination to unconsciously apply their L1 segmenting procedures during L2 listening (Cutler, 2000; Cross, 2009). Therefore, another possible explanation would be that bilinguals who possess a greater ability of cognitive inhibition could perform better in suppressing the natural tendency of utilizing their L1 segmenting procedures for L2 listening processing. Such continuous inhibition might, in turn, lead to higher L2 listening proficiency.

The result of the current study indicated that in addition to the significant and direct impact of SWS on L2 listening, bilinguals' SWS also had an indirect influence on their L2 listening proficiency through cognitive inhibition. Although previous literature has suggested that SWS is not the only factor for the success of L2 listening and cognitive inhibition may play a part in this process (e.g., McQueen et al., 1994; Norris, 1994; Norris et al., 1995; Luce and Cluff, 1998; Luce and Lyons, 1999), to the best of our knowledge, only a few studies have provided direct empirical evidence. Therefore, we extended this line of research by expanding the participant sample of research to bilinguals whose L2 is in Indo-European languages whereas their L1 is in Sino-Tibetan languages.

In accordance with our first hypothesis, the result of mediation analysis in the current study showed that cognitive inhibition mediated the relationship between participants' SWS and L2 listening proficiency. In other words, L2 learners' SWS might have a positive influence on their L2 listening proficiency through cognitive inhibition. Our results support previous findings that bilinguals' language experience may enhance their ability of cognitive control (e.g., Xie and Dong, 2017). Therefore, by the continuous practice of segmenting L2 speech during language learners' daily communication, they would possess a greater ability to inhibit linguistic competition both within-language and cross-language. Consequently, this more advanced ability of cognitive inhibition would then contribute to better performance in L2 listening processing. The results suggested that besides the bottom-up listening processing (i.e., SWS), L2 listeners also recruited top-down cognitive control (i.e., cognitive inhibition) to achieve successful L2 listening. Such results provided robust evidence for the interactive model of L2 listening processing.

The Mediation Role of Cognitive Flexibility

As mentioned earlier, the literature has pointed out that the constructs of cognitive inhibition and cognitive flexibility are strongly intertwined and interdependent (Miyake et al., 2000; Chevalier and Blaye, 2008). Therefore, we expected that cognitive flexibility should also play a similar role in the relationship between SWS and L2 listening proficiency as cognitive inhibition does. However, cognitive flexibility and cognitive inhibition have still, to some extent, been considered as two distinct mechanisms (Miyake et al., 2000), and cognitive flexibility is thought to be built on cognitive inhibition (Diamond, 2013). Therefore, it is evident that the enhancement of one aspect of the two abilities does not necessarily mean the facilitation of the other. The exact role played by cognitive flexibility in the relationship between SWS and L2 listening proficiency remains to be confirmed.

The present study extended previous research by finding that bilinguals' SWS might have a positive influence on their cognitive flexibility. Previous research has found that even though it might be inconducive for successful L2 listening, language learners are seemingly reluctant to abandon the inappropriate SWS procedure they have built for L2 listening comprehension (Field, 2008). This finding stressed the importance of cognitive flexibility during the process of SWS and could serve as indirect evidence for our result. In addition, prior research has shown that when confronted with L2 input, language learners would have multiple interpretations of the utterance (Weiss et al., 2009). Given the ephemeral and multifaceted nature of real-world communication, L2 listeners must process the speech that they have just heard while simultaneously receiving new upcoming utterances by the interlocutors during the bottom-up listening processing (i.e., SWS). This exact nature of communication calls for the L2 listeners' ability of cognitive flexibility to actively shift their focus and mental state not only between the meaning of the words they have just heard and the new input yet to come (Vandergrift, 2011) but also the co-activated irrelevant candidate words and interpretations between the two languages that they know (Dong and Xie, 2014). The result of our study is consistent with previous findings by showing that L2 listeners may use cognitive flexibility to consciously shift between languages and revise misinterpretations triggered by competing alternatives (e.g., Novick et al., 2005; Ye and Zhou, 2009). Therefore, this continuous demand of segmenting L2 speech and flexibly switching between two languages and multiple words during communication reflects better cognitive flexibility for language learners as an outcome.

As mentioned above, cognitive flexibility is considered to be built on cognitive inhibition (Diamond, 2013) since the shifting of perspectives, thinking, and attention requires the suppression of irrelevant or prepotent cues (Bialystok, 2015). Therefore, the results identified in the current study also suggested that cognitive flexibility was positively associated with L2 listening proficiency, similar to the role of cognitive inhibition as we expected before the present study. Previous studies have shown that cognitive flexibility has a significant influence on the decoding and language skills of children (e.g., Cartwright, 2012; Kieffer et al., 2013), young adolescents (e.g., Ober et al., 2019), and adults (e.g., Follmer and Sperling, 2019). The current study's finding is partially consistent with preceding research by suggesting that the L2 listeners who perform better in switching among alternative words and interpretations would exhibit greater language skills in the tasks of L2 listening, thereby showing higher L2 listening proficiency.

In line with the second hypothesis, the result of the current study showed that cognitive flexibility mediated the relationship between participants' SWS and L2 listening proficiency. The result provided empirical evidence to further verify that cognitive inhibition is not the only top-down control mechanism that would contribute to L2 listening. We found in our study that, besides the direct impact, there was an indirect influence of L2 learners' SWS on their L2 listening proficiency through cognitive flexibility. A possible explanation for this finding is that by consistently resolving linguistic competitions during the process of SWS in bilinguals' everyday life, their ability of cognitive flexibility may be facilitated (Bialystok, 2005) and contribute to the performance in L2 listening.

The Interactive Model of Second Language Listening

The current study further extended the research by examining the interaction of cognitive inhibition and cognitive flexibility in the relationship between SWS and L2 listening proficiency. Studies have previously found that the demands of resolving and processing linguistic competitions for L2 learners are the likely source of bilinguals' cognitive advantages over monolinguals (e.g., Bialystok, 2005; Kroll et al., 2012). In addition, such an increased cognitive control mechanism was positively related to L2 listening comprehension (e.g., Kim and Phillips, 2014). A reasonable interpretation is that bilinguals might possess better overall cognitive control due to their continuous engagement in competitive solutions among candidate words both within-language and cross-language during the process of SWS.

Moreover, our study also extended preceding research by showing that, besides the bottom-up listening processing of using linguistic cues (i.e., SWS), non-linguistic top-down cognitive processes such as cognitive inhibition and cognitive flexibility are also recruited by L2 listeners in L2 listening processing. At least on the lexical level, such findings provided further empirical evidence for the interactive model of L2 listening processing. Studies have shown that participants who use top-down listening processes can manage their attentional resources to achieve better L2 listening processing (e.g., Oh and Lee, 2014). In line with previous findings, the result of our study suggests that successful L2 listening not only requires the listeners to segment aural texts into meaningful and proper words during communication but also requires the co-activation of top-down cognitive control mechanisms to inhibit irrelevant competing candidate words actively and flexibly shift among words and interpretations.

Limitations

Several limitations in the present study, as well as suggestions for future studies, should be noted. First, the participants were all healthy Chinese young adults of similar ages, and most of them were females. Future studies should be conducted with a more balanced gender distribution and a more diverse participant group. Second, given that the components of cognitive control are still under heated debate, the current study conducted behavioral experiments only on cognitive inhibition and cognitive flexibility. Future studies should consider taking the influence of other top-down cognitive control mechanisms on L2 listening into account, such as working memory. Third, it is worth notifying that the relationship between bottom-up and top-down listening processing was not fully explored. Future studies should further investigate in what proportion listeners preferentially recruit the bottom-up and top-down processing of L2 listening. Fourth, the data collected in the current study was synchronic. The data may not fully account for the causal relationships among the variables. Future studies should conduct longitudinal investigations to establish causal relationships among bilinguals' SWS, cognitive control, and L2 listening proficiency.

CONCLUSION

The current study showed that L2 listening proficiency was in positive correlations with SWS, cognitive inhibition, and cognitive flexibility ability. The current study also investigated cognitive mechanisms underlying the process of SWS. The mediation analyses revealed that both cognitive inhibition and cognitive flexibility mediate the relationship between L2 learners' SWS and L2 listening proficiency. The results suggested that, along with the bottom-up listening processing of SWS, L2 listeners also exploited top-down cognitive control mechanisms to inhibit irrelevant competing candidate words and shift among words and interpretations to achieve successful L2 listening processing. The findings provided further empirical evidence for the interactive model of L2 listening processing.

The findings of our study may have important implications for future research on L2 listening to investigate the influence of non-linguistic cognitive control mechanisms during L2 listening processing. Furthermore, the findings of our study may also have implications for future L2 teaching and learning. L2 learners should attach importance to the inseparable contribution of both bottom-up and top-down processes in L2 listening and actively put them into practice. For future instructional methodologies, L2 teachers should also consider the interactive model of L2 listening processing to design more comprehensive curricula for language learners from both linguistic and cognitive perspectives.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Academic Committee of the Ministry of Education Key Laboratory of Modern Teaching Technology of Shaanxi Normal University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

YY designed the study, finished the experiments and surveys, performed the data analysis, and drafted the manuscript. GY drafted and revised the manuscript. YL collected data, reviewed and revised the manuscript, and supervised the study. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported and funded by the Shaanxi Provincial Research Project on Major Theoretical and Practical Issues in Philosophy and Social Sciences (2021ND0307); the Sub-discipline Research Team of Foreign Linguistics and Applied Linguistics (XWXKTD2021001); the China Postdoctoral Science Foundation (2017M623099 and 2018T111009); the Shaanxi Post-doctoral Science Foundation (2017BSHEDZZ128); the Research Program Funds of the Collaborative Innovation Center of Assessment toward Basic Education Quality at Beijing Normal University (2021-05-044-BZPK01); the Natural Science Basic Research Plan in Shaanxi Province of China (2022JQ-156); the Innovation Capability Support Program of Shaanxi Province, China (2020TD-037); and the Fundamental Research Funds of Innovation Team for the Central Universities (GK201901006).

REFERENCES

Bacon, S. M. (1992). Phases of listening to authentic input in Spanish: a descriptive study. *Foreign Lang. Ann.* 25, 317–333. doi: 10.1111/j.1944-9720.1992.tb00552.x

Best, J. R., Miller, P. H., and Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learn. Individ. Differ.* 21, 327–336.

Bialystok, E. (2005). "Consequences of bilingualism for cognitive development," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. de Groot (Oxford: Oxford University Press), 417–432.

Bialystok, E. (2015). Bilingualism and the development of executive function: the role of attention. *Child Dev. Perspect.* 9, 117–121. doi: 10.1111/cdep.12116

Bialystok, E., and Shapero, D. (2005). Ambiguous benefits: the effect of bilingualism on reversing ambiguous figures. *Dev. Sci.* 8, 595–604. doi: 10.1111/j.1467-7687.2005.00451.x

Blumenfeld, H. K., and Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: examining proficiency and lexical status using eye-tracking. *Lang. Cogn. Process.* 22, 633–660. doi: 10.1080/01690960601000746

Blumenfeld, H. K., and Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition* 118, 245–257. doi: 10.1016/j.cognition.2010.10.012

Boersma, P., and Weenink, D. (2019). *Praat: Doing Phonetics by Computer In (Version 6.1.04)*. Available online at: https://eric.ed.gov/?id=ED278275 (accessed August 14, 2019).

Canseco-Gonzalez, E., Brehm, L., Brick, C. A., Brown-Schmidt, S., Fischer, K., and Wagner, K. (2010). Carpet or Cárcel: the effect of age of acquisition and language mode on bilingual lexical access. *Lang. Cogn. Process.* 25, 669–705. doi: 10.1080/01690960903474912

Carlson, S. M., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298. doi: 10.1111/j.1467-7687.2008.00675.x

Carrell, P. L., and Eisterhold, J. C. (1983). Schema theory and ESL reading pedagogy. *TESOL Q.* 17, 553–573. doi: 10.2307/3586613

Cartwright, K. B. (2012). Insights from cognitive neuroscience: the importance of executive function for early reading development and education. *Early Educ. Dev.* 23, 24–36. doi: 10.1080/10409289.2011.615025

Chang, I. (2020). Influences of executive function, language comprehension, and fluency on young children's reading comprehension. *J. Early Childh. Res.* 18, 44–57. doi: 10.1177/1476718X19875768

Chevalier, N., and Blaye, A. (2008). Cognitive flexibility in preschoolers: the role of representation activation and maintenance. *Dev. Sci.* 11, 339–353. doi: 10.1111/j.1467-7687.2008.00679.x

Cole, R. A., and Jakimik, J. (1980). How are syllables used to recognize words? *J. Acoust. Soc. Am.* 67, 965–970. doi: 10.1121/1.383939

Cross, J. (2009). Diagnosing the process, text, and intrusion problems responsible for L2 listeners' decoding errors. *Asian EFL J.* 11, 31–53.

Cutler, A. (2000). Listening to a second language through the ears of a first. *Interpreting* 5, 1–23. doi: 10.1075/intp.5.1.02cut

Cutler, A., and Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 113–121. doi: 10.1037/0096-1523.14.1.113

Cutler, A., and Shanley, J. (2010). "Validation of a training method for L2 continuous-speech segmentation," in *Proceedings of the 11th Annual Conference of the International Speech Communication Association (Interspeech 2010)*, Makuhari.

Dennis, J. P., and Vander Wal, J. S. (2010). The cognitive flexibility inventory: instrument development and estimates of reliability and validity. *Cogn. Ther. Res.* 34, 241–253. doi: 10.1007/s10608-009-9276-4

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Dong, Y., and Xie, Z. (2014). Contributions of second language proficiency and interpreting experience to cognitive control differences among young adult bilinguals. *J. Cogn. Psychol.* 26, 506–519. doi: 10.1080/20445911.2014.924951

Farrell, J. (2015). *Training L2 Speech Segmentation with Word-Spotting*. Sydney, NSW: Western Sydney University.

Feyten, C. M. (1991). The power of listening ability: an overlooked dimension in language acquisition. *Mod. Lang. J.* 75, 173–180. doi: 10.2307/328825

Field, J. (2003). Promoting perception: lexical segmentation in L2 listening. *ELT J.* 57, 325–334. doi: 10.1093/elt/57.4.325

Field, J. (2008). Revising segmentation hypotheses in first and second language listening. *System* 36, 35–51. doi: 10.1016/j.system.2007.10.003

Flowerdew, J., and Miller, L. (2005). *Second Language Listening: Theory and Practice*. Stuttgart: Ernst Klett Sprachen.

Follmer, D. J., and Sperling, R. A. (2019). A latent variable analysis of the contribution of executive function to adult readers' comprehension of science text: the roles of vocabulary ability and level of comprehension. *Read. Writ.* 32, 377–403. doi: 10.1007/s11145-018-9872-3

Fotos, S. (2001). "Cognitive approaches to grammar instruction," in *Teaching English as A Second or Foreign Language*, Vol. 3, ed. M. Celce-Murcia (Boston, MA: Heinle and Heinle), 267–283.

Gitt, W. (2006). *In the Beginning was Information: A Scientist Explains the Incredible Design in Nature*. Green Forest, AR: New Leaf Publishing Group.

Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System* 28, 55–75. doi: 10.1016/s0346-251x(99)00060-3

Goyet, L., de Schonen, S., and Nazzi, T. (2010). Words and syllables in fluent speech segmentation by French-learning infants: an ERP study. *Brain Res.* 1332, 75–89. doi: 10.1016/j.brainres.2010.03.047

Graham, S. (2006). Listening comprehension: the learners' perspective. *System* 34, 165–182. doi: 10.1016/j.system.2005.11.001

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81. doi: 10.1017/S1366728998000133

Harnishfeger, K. K. (1995). "The development of cognitive inhibition: theories, definitions, and research evidence," in *Interference and Inhibition in Cognition*, eds F. N. Dempster and C. J. Brainerd (Amsterdam: Elsevier), 175–204.

Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York, NY: The Guilford Press, doi: 10.1111/jedm.12050

Heidlmayr, K., Moutier, S., Hemforth, B., Courtin, C., Tanzmeister, R., and Isel, F. (2014). Successive bilingualism and executive functions: the effect of second language use on inhibitory control in a behavioural Stroop Colour Word task. *Biling. Lang. Cogn.* 17, 630–645. doi: 10.1017/S1366728913000539

Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Q.* 40, 109–131. doi: 10.2307/40264513

Jansiewicz, E. M. (2008). *The relationship Between Executive Functions and Metacognitive Strategy Learning and Application*. Doctoral dissertation. Atlanta, GA: Georgia State University.

Kieffer, M. J., Vukovic, R. K., and Berry, D. (2013). Roles of attention shifting and inhibitory control in fourth-grade reading comprehension. *Read. Res. Q.* 48, 333–348. doi: 10.1002/rrq.54

Kim, Y.-S., and Phillips, B. (2014). Cognitive correlates of listening comprehension. *Read. Res. Q.* 49, 269–281. doi: 10.1002/rrq.74

Kroll, J. F., Dussias, P. E., Bogulski, C. A., and Kroff, J. R. V. (2012). "Juggling two languages in one mind: what bilinguals tell us about language processing and its consequences for cognition," in *Psychology of Learning and Motivation*, Vol. 56, ed. B. H. Ross (Amsterdam: Elsevier), 229–262.

Lee, T. M., and Chan, C. C. (2000). Stroop interference in Chinese and English. *J. Clin. Exp. Neuropsychol.* 22, 465–471. doi: 10.1076/1380-3395(200008)22:4;1-0;FT465

Li, Y., Yang, Y., Tang, A. C., Liu, N., Wang, X., Du, Y., et al. (2020). English spoken word segmentation activates the prefrontal cortex and temporo-parietal junction in Chinese ESL learners: a functional near-infrared spectroscopy (fNIRS) study. *Brain Res.* 1733:146693. doi: 10.1016/j.brainres.2020.146693

Luce, P. A., and Cluff, M. S. (1998). Delayed commitment in spoken word recognition: evidence from cross-modal priming. *Percept. Psychophys.* 60, 484–490. doi: 10.3758/BF03206868

Luce, P. A., and Lyons, E. A. (1999). Processing lexically embedded spoken words. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 174–183. doi: 10.1037//0096-1523.25.1.174

Lynch, T. (1998). Theoretical perspectives on listening. *Annu. Rev. Appl. Linguist.* 18, 3–19. doi: 10.1017/S0267190500003457

Lynch, T. (2002). *Listening: Questions of level*. Oxford: Oxford University Press.

Marian, V., and Spivey, M. (2003a). Bilingual and monolingual processing of competing lexical items. *Appl. Psycholinguist.* 24, 173–193. doi: 10.1017/S0142716403000092

Marian, V., and Spivey, M. (2003b). Competing activation in bilingual language processing: within-and between-language competition. *Biling. Lang. Cogn.* 6, 97–115. doi: 10.1017/S1366728903001068

Marian, V., Spivey, M., and Hirsch, J. (2003). Shared and separate systems in bilingual language processing: converging evidence from eyetracking and brain imaging. *Brain Lang.* 86, 70–82. doi: 10.1016/S0093-934X(02)00535-7

McQueen, J. M., Norris, D., and Cutler, A. (1994). Competition in spoken word recognition: spotting words in other words. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 621–638. doi: 10.1037/0278-7393.20.3.621

Mendelsohn, D. J. (1998). Teaching listening. *Annu. Rev. Appl. Linguist.* 18, 81–101. doi: 10.1017/S0267190500003494

Mercier, J., Pivneva, I., and Titone, D. (2014). Individual differences in inhibitory control relate to bilingual spoken word processing. *Biling. Lang. Cogn.* 17, 89–117. doi: 10.1017/S1366728913000084

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex :"Frontal Lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Murphy, J. M. (1985). *An Investigation into the Listening Strategies of ESL College Students*. Available online at: https://eric.ed.gov/?id=ED278275 (accessed November 10, 2021).

Murphy, J. M. (1991). Oral communication in TESOL: integrating speaking, listening, and pronunciation. *TESOL Q.* 25, 51–75. doi: 10.2307/3587028

Namaziandost, E., Neisi, L., Mahdavirad, F., and Nasri, M. (2019). The relationship between listening comprehension problems and strategy usage among advance EFL learners. *Cogent Psychol.* 6:1691338. doi: 10.1080/23311908.2019.1691338

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition* 52, 189–234. doi: 10.1016/0010-0277(94)90043-4

Norris, D., McQueen, J. M., and Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Exp. Psychol. Learn. Mem. Cogn.* 21, 1209–1228. doi: 10.1037/0278-7393.21.5.1209

Nouwens, S., Groen, M. A., and Verhoeven, L. (2016). How storage and executive functions contribute to children's reading comprehension. *Learn. Individ. Differ.* 47, 96–102. doi: 10.1016/j.lindif.2015.12.008

Nouwens, S., Groen, M. A., Kleemans, T., and Verhoeven, L. (2021). How executive functions contribute to reading comprehension. *Br. J. Educ. Psychol.* 91, 169–192. doi: 10.1111/bjep.12355

Novick, J. M., Trueswell, J. C., and Thompson-Schill, S. L. (2005). Cognitive control and parsing: reexamining the role of Broca's area in sentence comprehension. *Cogn. Affect. Behav. Neurosci.* 5, 263–281. doi: 10.3758/CABN.5.3.263

Ober, T. M., Brooks, P. J., Plass, J. L., and Homer, B. D. (2019). Distinguishing direct and indirect effects of executive functions on reading comprehension in adolescents. *Read. Psychol.* 40, 551–581. doi: 10.1080/02702711.2019.1635239

Oh, E., and Lee, C. M. (2014). The role of linguistic knowledge and listening strategies in bottom-up and top-down processing of L2 listening. *English Teach.* 69, 149–173. doi: 10.15858/engtea.69.2.201406.149

Orii-Akita, M. (2014). "The effectiveness of interactive teaching methods in EFL classrooms: a comparison with bottom-up and top-down methods," in *Proceedings of the International Symposium on the Acquisition of Second Language Speech, Concordia working Papers in Applied Linguistics*, Vol. 5, 464–477. Available online at: doe.concordia.ca/copal/documents32_OriiAkita_Vol5.pdf (accessed November 8, 2021).

Oxford, R. L. (1993). Research update on teaching L2 listening. *System* 21, 205–211. doi: 10.1016/0346-251X(93)90042-F

Preacher, K. J., and Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Methods Instrum. Comput.* 36, 717–731. doi: 10.3758/BF03206553

Rubin, J. (1994). A review of second language listening comprehension research. *Mod. Lang. J.* 78, 199–221. doi: 10.2307/329010

Sanders, L. D., Neville, H. J., and Woldorff, M. G. (2002). Speech segmentation by native and non-native speakers. *J. Speech Lang. Hear. Res.* 45, 519–530. doi: 10.1044/1092-4388(2002/041)

Snow, M. A. (2005). "A model of academic literacy for integrated language and content instruction," in *Handbook of Research in Second Language Teaching and Learning*, ed. E. Hinkel (London: Routledge), 717–736. doi: 10.4324/9781410612700-52

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651

Sunderman, G., and Kroll, J. F. (2006). First language activation during second language lexical processing: an investigation of lexical form, meaning, and grammatical class. *Stud. Sec. Lang. Acquis.* 28, 387–422. doi: 10.1017/S0272263106060177

Teubner-Rhodes, S. E., Mishler, A., Corbett, R., Andreu, L., Sanz-Torrent, M., Trueswell, J. C., et al. (2016). The effects of bilingualism on conflict monitoring, cognitive control, and garden-path recovery. *Cognition* 150, 213–231. doi: 10.1016/j.cognition.2016.02.011

Vandergrift, L. (1997). The comprehension strategies of second language (French) listeners: a descriptive study. *Foreign Lang. Ann.* 30, 387–409. doi: 10.1111/j.1944-9720.1997.tb02362.x

Vandergrift, L. (2003). Orchestrating strategy use: toward a model of the skilled second language listener. *Lang. Learn.* 53, 463–496. doi: 10.1111/1467-9922.00232

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Lang. Teach.* 40, 191–210. doi: 10.1017/S0261444807004338

Vandergrift, L. (2008). "Learning strategies for listening comprehension," in *Language Learning Strategies in Independent Settings*, eds S. Hurd and T. Lewis (Bristol: Multilingual Matters), 84–102. doi: 10.21832/9781847690999-007

Vandergrift, L. (2011). "Second language listening," in *Handbook of Research in Second Language Teaching and Learning*, Vol. 2, ed. E. Hinkel (London: Routledge), 455–471.

Vega-Mendoza, M., West, H., Sorace, A., and Bak, T. H. (2015). The impact of late, non-balanced bilingualism on cognitive performance. *Cognition* 137, 40–46. doi: 10.1016/j.cognition.2014.12.008

Vendrell, P., Junqué, C., Pujol, J., Jurado, M. A., Molet, J., and Grafman, J. (1995). The role of prefrontal regions in the Stroop task. *Neuropsychologia* 33, 341–352. doi: 10.1016/0028-3932(94)00116-7

Wang, Y., Yang, Y., Xiao, W. T., and Su, Q. (2016). Validity and reliability of the Chinese version of the cognitive flexibility inventory in college students. *Chin. Ment. Health J.* 30, 58–63. doi: 10.3969/j.issn.1000-6729.2016.01.012

Weber, A., and Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *J. Mem. Lang.* 50, 1–25. doi: 10.1016/S0749-596X(03)00105-0

Weiss, D. J., Gerfen, C., and Mitchel, A. D. (2009). Speech segmentation in a simulated bilingual environment: a challenge for statistical learning? *Lang. Learn. Dev.* 5, 30–49. doi: 10.1080/15475440802340101

Wilson, C., and Mihalicek, V. (2011). *Language Files: Materials for an Introduction to Language and Linguistics*. Columbus, OH: Ohio State University Press.

Xie, Z., and Dong, Y. (2017). Contributions of bilingualism and public speaking training to cognitive control differences among young adults. *Biling. Lang. Cogn.* 20, 55–68. doi: 10.1017/S1366728915000474

Yang, Y., Li, Y., Wang, X., Liu, N., Jiang, K., Zhang, S., et al. (2021). Cognitive inhibition mediates the relationship between ESL listening proficiency and English spoken word segmentation in Chinese learners: a functional near-infrared spectroscopy (fNIRS) study. *J. Neurolinguistics* 59:100987. doi: 10.1016/j.jneuroling.2021.100987

Ye, Z., and Zhou, X. (2009). Executive control in language processing. *Neurosci. Biobehav. Rev.* 33, 1168–1177. doi: 10.1016/j.neubiorev.2009.03.003

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Psychology

REVIEW
published: 06 May 2022
doi: 10.3389/fpsyg.2022.872145

Mandarin and English Event Cognitive Alignment From Corpus-Based Semantic Fusion Model Perspective

*Xiangling Li**

Institute of Foreign Linguistics and Applied Linguistics, Henan University, Kaifeng, China

The study explores the fusion of semantic roles and the different semantic fusion types, aiming at establishing a semantic fusion model to explain the cognitive alignment of events in Chinese and English simple sentence constructions containing two verbs. In total, 20,280 simple sentence constructions containing two verbs are collected from Chinese literary works, Peking University Chinese Corpus, and English classic literary works. The semantic fusion in the collected simple sentence constructions containing two verbs is classified into five major semantic fusion categories, which appear with different occurrence frequencies in the two languages. The semantic fusion model of event alignment is comprehensively supported by linguistic research in Chinese and English. From a cognitive linguistic perspective, it is found that the double semantic profiles of the same syntactic element N (noun) make N psychologically activated twice and enable it to enter two processes profiled by the two verbs as a participant. The two processes are combined into one event, which designates a cognitive occurrence of any degree of complexity. N's entry into the two subevents is realized by its double semantic profiles that enable it to fuse two semantic roles into one syntactic element and explain the relationship between N's double syntactic identities and double semantic roles. The semantic fusion model was used to explore event alignment in simple sentence constructions containing two verbs, and it was discovered that the fusion of two semantic roles is universal in languages and is a common psychological and cognitive behavior deeply rooted in the mental conceptualization of language users. The empirical discussion of simple sentence constructions containing two verbs proves that semantic fusion as an important psychological passage in event alignment has solid psychological reality and verifies the applicability of the semantic fusion model in the explanation of event alignment.

Keywords: semantic role, semantic fusion, event structure, semantic profile, event alignment

OPEN ACCESS

Edited by:
Yan Jing Wu,
Ningbo University, China

Reviewed by:
Ying Zhang,
University of Maryland, College Park, United States
Jie Xu,
University of Macau, China

***Correspondence:**
Xiangling Li
10070040@vip.henu.edu.cn

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *09 February 2022*
Accepted: *16 March 2022*
Published: *06 May 2022*

Citation:
Li X (2022) Mandarin and English Event Cognitive Alignment From Corpus-Based Semantic Fusion Model Perspective. Front. Psychol. 13:872145. doi: 10.3389/fpsyg.2022.872145

INTRODUCTION

Semantic roles have a long-standing presence in theories of philosophy, cognitive science, and linguistics. The semantic roles such as agent, patient, goal, and instrument are cross-culturally universal (Fillmore, 1968) and are regarded as part of innate core language knowledge (Carey, 2011; Strickland, 2017). For a long time, semantic roles are routinely involved in the studies of language production, language acquisition, the interface between syntax and semantics, and cognitive science. The verb "eat" encodes a semantic relation between someone who eats and something that gets eaten, and the participants involved in this relation are given the role labels "agent" and "patient," respectively (Rissman and Majid, 2019). Semantic roles are very common in that they are fundamental to how people represent the world and how these representations are expressed in language. As a common semantic relationship, semantic fusion refers to the merging of two or more semantic roles and is an important means of event cognitive alignment in Chinese and English. Semantic fusion makes for the succinctness of language expressions; different actions within a simple sentence construction containing two verbs are integrated into one complete event through the fusion of two semantic roles. In the sentence "The boss made Tom do the work all day," the actions "make" and "do" are integrated into an event by way of the shared participant "Tom," with "Tom" being the patient of the action "make" and the agent of the action "do." The fusion of two semantic roles is common in English simple sentence constructions containing two verbs, such as resultative constructions and caused-motion constructions. Similarly, semantic fusion is very pervasive in Chinese simple sentence constructions containing two verbs, such as Chinese pivotal constructions, Chinese constructions with serial verbs, and verb-complement constructions. The semantic fusion in a simple sentence construction containing two verbs is the research focus of this study.

LITERATURE REVIEW

The research of events has been a classic topic of concern in the field of philosophy. Events are divided into actuality and movement, which is regarded as the earliest study of events. Vendler (1957) further divides events into state, activity, achievement, and accomplishment. Davidson (1967) puts forward the concepts of event argument and event individualization and establishes the ontological position of events. In the field of psychology, events are also extensively studied, especially the psychological authenticity of events (Shipley and Zacks, 2008), the causative relationship between an event and the speaker, a causal relationship between the event and the state, the psychological relationship with the action event, and the simultaneous relationship between the state and the event (Kistler, 2006; Chen, 2021). Event-related brain potential (ERP) studies provided evidence in support of parallel lexical access during bilingual language production (Wu and Thierry, 2011).

In the field of linguistics, the study of events is also a key topic for half a century (Jackendoff, 1976, 1990; Talmy, 1985; Pinker, 1989; Rappaport, 2008; Viswanatha et al., 2018; Li F. Y., 2019). Talmy (1985) puts forward a mobile event conceptual framework and the theory of macro-events, defining events as macro-event = motion-event + co-event, with secondary events indicating the way the main event moves or the reason for the movement of the main event (Talmy, 2000a). Pinker (1989) examines the relationship between syntax and semantics through the decomposition of predicate meaning and points out that the meaning of verbs is decomposed into major events and minor events, which are represented by a tree diagram method. In essence, although Pinker and Talmy's macro-event theories are expressed in different ways, they are somewhat similar. Jackendoff (1976) emphasizes that any event such as motion and spatial location is based on the basic predicate verbs and the interrelationship between causative verbs. Rappaport (2008) proposes semantic decomposition and believes that the internal semantics of verbs include root meaning and structural meaning. In short, the syntactic representation of semantic roles in an event and the analysis of the semantic structure of an event gradually arouse great interest in the field of linguistics.

Within the framework of Chinese traditional grammar, Zhang (1999,2001), by exploring Chinese pivotal constructions in the oracle bone script and Chinese sentences with serial verbs in the inscriptions of the Western Zhou Dynasty, point out the double syntactic identity of the same linguistic form in some special sentences, which gradually becomes the focus of debate among scholars. The psychological processing of squeezing two semantic components into one syntactic form is a common sentence-making method in Chinese (Lv, 1979).

The syntactic representation of simple sentence constructions containing two verbs also attracts scholars' observation from the perspective of structural linguistics. Under the influence of behaviorism, which holds that meaning is the situation expressed by a linguistic form and the response aroused in the listeners (Bloomfield, 2002), the structural research attaches great importance to linguistic form and proposes immediate constituent (IC) analysis to analyze the double syntactic identity of one syntactic element and puts forward dividing-one-word–into-two hypothesis to explain why the same syntactic element can merge or fuse two semantic roles in Chinese (Xing, 1986; Wu and Liang, 1992).

Logical analysis of the event structure in simple sentence constructions containing two verbs is carried out by scholars within the theoretical framework of transformational and generative linguistics. In accordance with the Thematic Criterion of the Governing & Binding theory, a theme can only be assigned one and at most one thematic role, and each thematic role can only be assigned to one theme (Huang, 1982; Boeckx and Horstein, 2004; Chomsky, 2010). The previous studies from the perspective of generative grammar believe that the syntactic element with two semantic roles violates the Thematic Criterion at the syntactic level, cannot have two syntactic identities at the same time, and puts forward an empty category (abbreviated to e) to explain the fusion of two semantic roles within one linguistic element. In the deep structure of a simple sentence construction including two verbs, there is an empty category behind the syntactic element that plays two semantic roles. In the English sentence "Tom persuaded Janie (e_i) to go to a picnic," there is an empty category e_i behind Janie, and the empty category e_i also refers to Janie. The proposal of empty category gives a satisfactory answer to the fusion of two semantic roles in a linguistic form and probes deeper into the logical structure, which helps to make clear the semantic structure of the event. The invisibility of the empty category at the syntactic level and its appearance at the semantic level touch upon the psychological representation of event structure in a simple sentence construction containing two verbs (Xing, 2004; Feng and Feng, 2018). However, why is there an empty category hidden behind the syntactic element, why does it turn up in semantic structure, and why it is shaded in the syntactic structure are still some doubts that need further explanation.

The studies on semantic fusion from the perspective of cognitive linguistics and cognitive psychology gradually arouse more attention. Goldberg (1995) pointed out that role merging occurs in reflexive constructions, with one participant role merging with another. The merged participant roles are squeezed with a single argument role and linked with a single grammatical function. Two actions in a sentence are integrated by merging two participant roles into one single argument role to form a composite event that is linked with a single grammatical function. The event participant categories are not as self-evident as categories provided by nouns and verbs (Rissman and Majid, 2019), and a variety of event-specific knowledge is activated during sentence comprehension (Bicknell et al., 2010; Metusalema et al., 2012). Talmy (2000b) discovered that a simple sentence representing an event is universal; it is not language-specific. The typical characteristics isomorphism between the event cognition and linguistic representation and the isomorphism between semantic fusion and syntactic fusion of an event were examined (Davidson, 1967; Parsons, 1990; Croft, 1991; Talmy, 2000b; Givón, 2001; Imbert, 2012; Li, 2020). The previous studies of events from a cognitive perspective are roughly divided into two categories. The first category takes verbs as the core, which focuses on the event structure and the realization of arguments (Jackendoff, 1990; Levin, 1993; Croft, 2012). The second category focuses on the difference in the linguistic representation of event components. Event integration through the fusion of semantic roles arouses scholars' great interest (Talmy, 1991; Fauconnier and Turner, 1996; Givón, 2001).

From the perspective of cognitive psychology, the natural language sentence matching method is proposed to combine high-level and low-level semantic information, using a heuristic fusion function to merge low-level semantic information with high-level semantic information to get the final semantic representation (Jiang et al., 2021). With regard to the mapping between syntactic relations and semantic cases, Van (2005); Gruber (1965), and Fillmore (1968) discovered that there exists a correspondence between semantic roles and syntactic locations. Jackendoff (1983) advocated that the argument structure should be described using complex and clear semantic structure, which is mapped to the syntactic structure. Croft (2012) analyzes the direct mappings between specific event structures and syntactic positions (e.g., subject and object) (Rissman and Majid, 2019). Fuzzy semantic overlapping allows a member to belong to more than one community (Sato et al., 2020). Similarly, in language, semantic fusion helps a participant to enter two actions and link them into a composite event (Langacker, 2012). The research on simple sentence constructions with one subject and two verbs in Chinese and English discovers one interesting fact that the verbs in them share at least one argument role that plays two participant roles. Squeezing the two roles into one word is crucial in the psychological alignment of an event. Semantic fusion seems to be a basic way for people to copy and combine different scenes into a complete human scene in the objective world (Li, 2015; Li X. L., 2019; Liu, 2017; Wen and Yin, 2018; Zhang and Pan, 2019). The cognitive mechanism that enables bilinguals to keep their languages functionally operating has not yet been elucidated (Wu and Thierry, 2017). So, in order to reveal the psychological and cognitive mechanism of this kind of language phenomenon, an in-depth research is needed.

In summary, previous studies on events in Chinese and English are done from different linguistic approaches. The existence of double syntactic identity is the biggest discovery in the previous studies from the perspective of traditional grammar and has aroused heated discussion, but the studies from the traditional grammar cannot explain the reason why the same linguistic form possesses two syntactic identities. The studies within the framework of structural linguistics put forward the dividing-one-word–into-two hypothesis to expound the double identity of the same linguistic form but still cannot explain why one word can be divided into two words at the syntactic level. The studies from the approach of generative linguistics put forward an empty category to offer a very convincing explanation to the double syntactic identity of the same linguistic form with the help of thematic role theory, pushing forward the studies of the events grounded in language, but at the same time leave one doubt why there exists an empty category behind one syntactic form. The cognitive studies of the event in simple sentence constructions attracted more and more attention, and many scholars try to explain the event structure and event integration from a cognitive perspective. But why one syntactic element can play two semantic roles needs to be probed further in order to reveal the cognitive mechanism of event integration in a simple sentence construction containing two verbs.

THEORETICAL GUIDANCE

Research is carried out with the guidance of Gestalt psychology and cognitive linguistics. Gestalt psychology emphasizes the integrity of experience and behavior and studies objects as a whole, which is not equal to the sum of the parts. The whole precedes the parts and determines the nature and meaning of each part. According to the principle of good Gestalt, the parts that belong to each other are easy to combine into a whole; on the contrary, the parts that do not belong to each other are easy to be isolated (Blackburn, 1940). Simplicity is one of the perception principles. When people perceive things, they often grasp the overall objective object through specific characteristics of certain parts and tend to summarize complex things into concise shapes by combining inherent experience and cognition (Chen, 2021). According to the shortest distance principle or proximity factor, some parts that are close to each other are easy to form a whole (Koffka, 1935).

Cognitive linguistics regards "Language is an integral part of human cognition" (Langacker, 1987:12). Composition is the starting point of cognitive linguistics. Different from the traditional valence theory, the cognitive valence theory believes that the valence relationship refers to the composition relationship between two or more linguistic units (Langacker, 1987). The valence relationship between linguistic units is established through the corresponding relationship between the semantic profiles of two semantic components. In accordance with the valence theory of cognitive grammar, a noun profiles a thing, a verb profiles a process, and an adjective or adverb profiles an atemporal relation. Based on this valence theory, Niu (2008) proposes a cognitive analysis model, which provides access to the combination of semantic components within a composite linguistic unit. The semantics of a grammatical structure not only contains objective and real conceptual information but also entails language users' perception, cognition, construal, and reasoning of objective reality. Talmy (2000a) believes that sentences can represent a series of cross-event relations, including time, cause and effect, concessions, and attachments. The typical feature of event fusion is that people's cognition and linguistic representation of events appear in isomorphism (Talmy, 2000b; Givón, 2001). The fusion of cause and effect is an important feature of human cognition. People have widely accepted that the cause event and the effect event are conceptualized as macro-event (Michotte, 1946; Talmy, 1991, 2000b; Wolff, 2003). The sub-events containing causal relationships can be merged into a single macro-event and represented by a single sentence. In this study, Gestalt psychology and cognitive linguistics are used to investigate the cognitive alignment of events in English and Chinese simple sentence constructions containing two verbs by way of semantic fusion.

MATERIALS AND METHODS

In the study, a large number of data are collected in order to analyze and classify the different categories of semantic fusion types. A total of 20,280 Chinese and English simple sentence constructions containing two verbs are collected from the Chinese classical literary works, Peking University Chinese Corpus, and English classic literary works. All the collected simple sentence constructions containing two verbs are classified according to the semantic fusion types, and the occurrence frequency of different semantic fusion types is counted. Based on the analysis of the semantic roles and the fusion types, a quantitative method is used to classify semantic fusion into different categories. Based on the qualitative analysis of semantic fusion realization and the noun's function in the integration of the two actions, an event semantic fusion model is established. The qualitative method is used in the explanation of how the two actions (sub-events) are integrated to form a composite event from a psychological and cognitive approach. Qualitative analysis is also adopted to expound what is the universally applicable cognitive mechanism for the combination of sub-events in a construction. The corpus-based quantitative method and the introspective qualitative method make the semantic fusion model possible and reasonable. Finally, an empirical discussion is done by carrying out a test. A total of 48 participants who are linguistic postgraduates took part in the test. The test is composed of 20 simple sentence constructions in Chinese and English and five types of semantic fusion. In total, 15 sentences are in the form of SV_1NV_2, two sentences are in the form of SV_1V_2N, and three sentences are in the form of SV_1NA(Adj.). The 20 sentences are put together in order to check whether the participants can accurately differentiate the three forms of constructions. In the test, the participants are asked to finish the following three tasks: (1) to recognize the N that plays two semantic roles, (2) to make a judgment whether the constructions in the test are SV_1NV_2 constructions, and (3) to determine whether the two verbs V_1 and V_2 are combined into a complete event through the syntactic element N. The test is used to testify N's function in the event integration of the two actions V_1 and V_2 and to prove that semantic fusion model is feasible and applicable in explaining the event integration.

LINGUISTIC REPRESENTATION OF SEMANTIC FUSION IN SIMPLE SENTENCE CONSTRUCTIONS CONTAINING TWO VERBS

In this section, the linguistic representation of simple sentence constructions including two verbs is discussed in detail. Simple sentence constructions containing two verbs have two typical syntactic features: (i) There are four indispensable elements in the construction: two nouns, one of which is the subject, and two verbs, usually in the form of SV_1NV_2 or SV_1V_2N. (ii) N usually appears before, between, or behind the two verbs.

(1) (a). *tamen pao kafei he.*
They pour coffee drink
'They make coffee to drink.'

(b). *zhangsan qing lisi zuo baogao.*
Zhang San invite Li Si deliver report
'Zhang San invite Li Si to deliver a report'.

(c). *mama qu chaoshi mai cai.*
Mom go supermarket buy vegetables
'Mom went to the supermarket to buy vegetables.'

(d). *ta* jiao *yisheng lai.*
He ask *doctor come*
'He asked a doctor to come.'

(e). *Timu pei nvpengyou gouwu.*
Tim accompany girlfriend go shopping
'Tim accompanied his girlfriend to go shopping.'

(f). *xiaogou beijiu huole.*
dog be saved come to life
'The little dog was saved and came to life.'

(g). tamen *da* *baile* *daishou.*
they *beat* *lose* *opponents*
'They made the opponents lose.'

(h). I saw Tom come out of the house.
(i). Mary wanted her mother to buy that coat.
(j). I heard Jenny crying in the room.
(k). Mary persuaded me to give up.
(l). My mom told me to finish my homework.

In (1), a–g are the simple sentence constructions with two verbs in Chinese. (1) h–l are simple sentence constructions with two verbs in English. In these constructions, the noun (N) is very crucial, and it is usually between the two verbs as it is seen in (1) a, b, d, e, i, and k. In (1) a, *kafei* "coffee" is put between the two verbs *pao* "make" and *he* "drink," and in (1) b, *lisi* "Li Si" is between the verbs *qing* "invite" and *zuobaogao* "make a report." In some of these constructions, the noun (N) is behind the two verbs. In (1) g, the noun *daishou* "opponent" is put behind the verbs *da* "beat" and *bai* "lose." The complexity of (1) e needs a detailed analysis. In (1) e, although there are two verbs *pei* "accompany" and *gouwu* "go shopping," there are two nouns *timu* "Tim" and *nvpengyou* "his girlfriend," and both of the two nouns have a syntactic relationship with the two verbs.

CLASSIFICATION OF SEMANTIC FUSION IN SIMPLE SENTENCE CONSTRUCTIONS CONTAINING TWO VERBS

The semantic role generally refers to the role of the participant in the event or activity described by the predicate. In the study of syntax and semantics, this participant role has been given many different names, such as deep case (Fillmore, 1968), thematic roles (Gruber, 1965; Jackendoff, 1972; Dowty, 1986; Carlson, 1998), participant roles (Allan, 1986), semantic roles (Givón, 1990), and argument roles (Goldberg, 1995). The deep structure of a sentence includes a predicate and one or more noun phrases, and each noun phrase establishes a specific case relationship with the predicate (Fillmore, 1968). Agent, experiencer, patient, theme (undergoer), fractive, and locative are six basic semantic roles.

Semantic fusion is very complex and pervasive, appearing in different combinations of semantic roles. Based on the observation and analysis of 20,820 simple sentence constructions (15,715 in Chinese and 5,105 in English) collected from the corpus and sources mentioned above, it is discovered that there are mainly five semantic fusion categories (refer to **Table 1**): agent-agent fusion, agent-patient fusion, agent-experiencer fusion, patient-patient fusion, and patient-experiencer fusion.

Type I agent/agent fusion is the type of semantic fusion with the highest occurrence frequency and accounts for 31.27% of the data collected. *mama* "mom" is the agent of the action *quchaoshi* "go to supermarket" and the agent of the action *maicai* "buy vegetables" in (1) c. In (1) e, the two actions *pei* "accompany" and *gouwu* "go shopping" share the same agent *xiaowang* "Xiaowang."

Type II agent/patient fusion occurs with relatively high occurrence frequency, *occupying* 28.19% of the data collected. This type of fusion is discovered in (1) b, d, and f. In (1) b, lisi "Li Si" is the patient of the action *qing* "invite" and the agent of the action *zuobaogao* "deliver a report"; in (1) d, *yisheng* "doctor" is the patient of the action expressed by the verb *jiao* "ask" and the agent of the action expressed by the verb *lai* "come"; in (1) f, *xiaogou* "little dog" is the patient of the action *beijiu* "saved" and the agent of the action *huo* "come to life."

The frequency of type III agent/experiencer fusion is slightly low and takes up 17.51% of the data collected. In (1) e, *nvpengyou* "his girlfriend" is the experiencer of the action expressed by the verb *pei* "accompany" and the agent of the action indicated by the verb *gouwu* "go shopping"; in (1) h, "Tom" is the experiencer of the action expressed by the verb "see" and the agent of the action performed by the verb "come"; in (1) j, "Jenny" is the experiencer of the action performed by the verb "hear" and the agent of the action expressed by "cry."

Similar to the first three types of semantic fusion type discussed earlier, the fusion type of patient/patient and patient/experiencer is also pervasive in language, but when compared to the first three fusion types, the occurrence frequency of these two types is slightly lower with 11.29 and 11.74%, respectively. Patient/patient fusion appears in (1) a, where *kafei* "coffee" is the patient of the action expressed by the verb *pao* "pour" and the patient of the action expressed by the verb *he* "drink." Patient/experiencer fusion is discovered in (1) g, where *daishou* "opponent" is the patient of the action *da* "beat" and the experiencer of the action *bai* "lose."

SEMANTIC REPRESENTATION OF THE EVENT IN SIMPLE SENTENCE CONSTRUCTIONS CONTAINING TWO VERBS

The semantic representation of these constructions involves semantic decomposition. By decomposing the meaning of a word into various aspects (components, means, participants, location, etc.), what is latent in the meaning of a word is made apparent. In simple sentence constructions containing two verbs, the understanding of the meaning of V_1 and V_2 is closely related to their arguments. The semantics of simple sentence constructions consists of two basic parts: (1) the representation of semantic components and (2) the representation of the event logical structure. In simple sentence constructions containing two verbs, the predicate is represented by an activity logical structure that has three arguments. In these constructions, the predicate V_1 takes three arguments, with V_2 being one of them. Semantic roles are the roles that arguments of a predicate take. Consider the sentence "Joe squeezed the rubber ball inside the jar," "squeezed" is the predicate. "Joe, rubber ball and jar" gets the semantic roles of squeezer (agent), squeezee (patient), and location. This motion event is described by a verb (squeeze), a proper noun

TABLE 1 | Categories of semantic fusion in Chinese and English.

Corpus Type	Chinese classic literary works		Peking University Chinese Corpus		English classic literary works		Total	
	Quantity amount	percentage	quantity	percentage	quantity	percentage	quantity	percentage
Type I agent /agent fusion	2980	30.21%	1880	32.14%	1650	32.32%	6510	31.27%
Type II agent /patient fusion	2660	26.96%	1660	28.38%	1550	30.36%	5870	28.19%
Type III agent /experiencer fusion	1805	18.30%	980	16.75%	860	16.85%	3645	17.51%
Type IV patient /patient fusion	1255	12.72%	670	11.45%	520	10.19%	2445	11.74%
Type V patient /experiencer fusion	1165	11.81%	660	11.28%	525	10.28%	2350	11.29%
Total	9865	100%	5850	100%	5105	100%	20820	100%

(Joe), a noun phrase (the rubber ball), and a preposition phrase (inside the jar). A neo-Davidsonian event representation of this motion event is as follows: $\exists$ e,x,y Squeezing(e)$\wedge$Squeezer(e, Joe)$\wedge$Squeezed Thing(e,y)$\wedge$Rubber Ball(y).

Similarly, in simple sentence constructions containing two verbs, the semantic roles express the roles that arguments of V_1 and V_2 take. The semantic representation of the event structure is as follows: $\exists$e, S, N, (X) V_1-ing(e1)$\wedge$V1-er (e1, S)$\wedge$$V_1$-ed Thing(e1, N)$\wedge$$V_2$-ing(e2)$\wedge$$V_2$-er (e2, N)$\wedge$V2-ed Thing(e2, x).

This formula encodes an event, and the participants are S, N, and X. S and N are two indispensable participants, and X is not the necessary participant. S and N are the participants of sub-event 1 expressed by V_1. N and X are participants of the sub-event 2 expressed by V_2. The event in (1) k is as follows: $\exists$e, Mary, me, x Persuading (e1)$\wedge$Persuader (e1, Mary)$\wedge$Persuaded Thing (e1, me)$\wedge$Giving up (e2)$\wedge$give-up-er (e2, me)$\wedge$given up Thing (e2, x).

SEMANTIC FUSION MODEL OF EVENT ALIGNMENT

Through the analysis of semantic fusion types, it is discovered that the semantic components in a construction express a complete meaning and are regarded as a whole. Here the cognitive alignment of the semantic components in a sentence is expounded with the guidance of cognitive linguistics. According to the valence theory of cognitive grammar, the three semantic components are aligned into a complete event by the way of semantic profiling, and the alignment process is shown in **Figure 1**.

In **Figure 1**, N is an autonomous element; V_1 and V_2 are dependent elements. N profiles an entity with one substructure elevated to a special level of prominence. The bold circle in the boxes V_1, N, and V_2 stands for the profiled substructure. The box stands for the base, which refers to the basic cognitive domain used to perceive the profiled substructure. The left and right boxes stand for the process or relation profiled by V_1 and V_2. The dotted circle in the right box stands for the possible existence of profiled substructures of V_2. The dotted line stands for the correspondence between the profiled structure of the dependent element and the profiled structure of the autonomous element. The arrow stands for the elaboration relation in which one element provides an elaboration site that is elaborated by the profiled structure of another element in construction.

In simple sentence constructions containing two verbs, the two verbs (V_1 and V_2) are dependent elements that each profiles a process that includes one or two participants. The noun (N) semantically profiles an entity. In the composition between V_1 and N, V_1 provides an elaboration site, and one semantic profile of N (N_1) is psychologically activated to elaborate the site and helps N get its entry into the process profiled by V_1. N becomes one of the participants of the process. In the composition between N and V_2, V_2 profiles a schematic trajector, and another semantic substructure of N (N_2) is mentally activated and elaborates the schematic trajector profiled by V_2. Distinct and related predications are obtained by imposing alternate profiles on a given base (Langacker, 1987). Alternate profiles of N are psychologically activated and enable it to enter two processes as a participant, and the two processes are combined into one event, which designates a cognitive occurrence of any degree of complexity. N's entry into two processes is realized by its double semantic profiles, which explains why the two semantic roles of N are fused into one participant. N becomes a psychological passage in the combination of two processes with the help of its double semantic profiles.

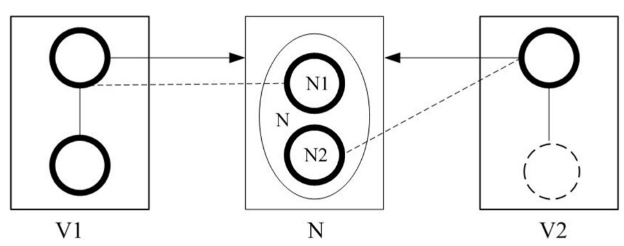


FIGURE 1 | Semantic fusion model of event cognitive alignment.

DISCUSSION OF EVENT ALIGNMENT THROUGH SEMANTIC FUSION MODEL IN SIMPLE SENTENCE CONSTRUCTIONS CONTAINING TWO VERBS

In this section, semantic fusion in simple sentence constructions containing two verbs is discussed in detail to check the operability and rationality of the semantic fusion model in explaining event alignment. Through the application of the event semantic fusion model, how the semantic components in simple sentence constructions containing two verbs are aligned is presented. The alignment clearly reveals that the realization of syntactic overlap is the result of the double semantic profiling of the same syntactic element and explains the correspondence between the syntactic overlap and the semantic overlap.

(i) Patient-agent semantic fusion

The realization of the patient-agent semantic fusion is illustrated through the event alignment in (2) in **Figure 2**.

(2) Tom made Jane cry.

(2) Is patient-agent semantic fusion, with "Jane" being the patient of V_1 (make) and the agent of V_2 (cry). V_1 (make) and V_2 (cry) are two conceptually dependent components and each of the two verbs semantically profiles a process. The conceptualization of the two verbs needs such components as who performs the action, who is affected in the action, where and when the action happens, etc. "Tom" and "Jane" are two nouns, which are two conceptually autonomous components. Each of the two nouns semantically profiles an entity, making one or more aspects of the entity elevated to a special level of prominence. The semantic profiles of the four elements are shown at the bottom of **Figure 2**.

In (2), the cognitive alignment of the event involves two sub-events: sub-event 1 "Tome made Jane" and sub-event 2 "Jane cry." Correspondingly, the event alignment includes two parts: one is the combination of sub-event 1, including S (Tom), V_1 make, and N (Jane), and the other is the combination of sub-event 2, including N (Jane) and V_2 (cry). In the combination of S (Tom), V_1 (make), and N (Jane), the dependent element V_1 (make) profiles a process including a schematic trajector and a schematic landmark as shown at the bottom of **Figure 3**. S (Tom) profiles an entity capable of performing an action, and N (Jane) profiles an entity that is able to accept action. The semantic profile of S (Tom) elaborates the trajector profiled by V_1 (make), and "Tom" enters the process and becomes a participant to perform the action expressed by V_1 (make). The semantic profile of N (Jane) elaborates the landmark profiled by V_1 (make) and enters the process as a participant who is affected by the action expressed by V_1 (make). The three elements are aligned into a composite semantic structure [TOM MAKE JANE] that means "Tom did something unpleasant to Jane". In the combination of N(Jane) and V_2 (cry), the dependent element V_2 (cry) profiles a process with a schematic trajector. The semantic profile of N(Jane) elaborates the trajector profiled by V_2 (cry) and becomes a participant in the process. The two elements are combined into a composite semantic structure [JANE CRY], which means "Jane performs the action of crying." From the alignments of the two sub-events [TOM MAKE JANE] and [JANE CRY], it is clear that the double semantic profiles of N (Jane) help Jane have the ability to play two roles in the two processes V_1 (make) and V_2 (cry) as a participant and that the alignment of a composite semantic structure [TOM MAKE JANE CRY] is realized by the double semantic profiles of the same syntactic element N (Jane), which is regarded as the psychological passage of the two processes.

(ii) Patient-patient semantic fusion

(3) yuehan zuo fan chi.
John cook dinner eat
'John cooked a dinner to eat.'

In (3), N (*fan* "dinner") is the patient of both V_1 (*zuo* "cook") and V_2 (*chi* "eat"). *fan* "dinner" is the patient-patient semantic fusion, with it being the patients of V_1 (*zuo* "cook") and V_2 (*chi* "eat"). In (3), among the four syntactic elements, *Yuehan* "John," *zuo* "cook," *fan* "dinner," and *chi* "eat"; *Yuehan* "John" and *fan* "dinner" are two autonomous elements. *Yuehan* "John" profiles a person capable of performing an action and *fan* "dinner" profiles a thing that can be cooked and eaten. V_1 (*zuo* "cook") and V_2 (*chi* "eat") are two dependent elements, each profiling a process with a schematic trajector and a schematic landmark. The semantic profiles of the four elements are shown in **Figure 3**.

The psychological cognition of the event in (3) involves two sub-events "John cooked a dinner" and "Dinner was eaten." In the alignment of the sub-event "John cooked a dinner," S (*yuehan* "John"), V_1 (*zuo* "cook"), and N (*fan* "dinner"), the semantic profiles of S (*yuehan* "John") and N (*fan* "dinner") elaborate the schematic trajector and landmark provided by V_1 (*zuo* "cook") and the elaboration site into S (*yuehan* "John") and the landmark into N (*fan* "dinner"). S (*yuehan* "John") and N (*fan* "dinner") enter the process profiled by V_1 (*zuo* "cook") as participants, and the three elements are aligned into a composite structure [JOHN COOK DINNER], which means "John cooked a dinner." In the alignment of the sub-event "Dinner was eaten," the semantic profile of N (*fan* "dinner") elaborates the schematic landmark provided by V_2 (*chi* "eat"). Therefore, N (*fan* "dinner") gets a participant membership and enters the process profiled by V_2 (*chi* "eat"), and the two elements are aligned into the composite structure [DINNER EATEN], which means "The dinner was eaten." From the composition of two sub-events, it is discovered that N (*fan* "dinner") is activated in the two processes and the fusion of two processes into one complete event [JOHN COOK DINNER TO EAT] is realized by the double semantic profiles of N (*fan* "dinner").

(iii) Other three types of semantic fusion

Through the detailed illustration of patient/agent and patient/patient fusion based on the semantic fusion model, it is discovered that the model works very well in explaining the alignment of events in simple sentence constructions containing

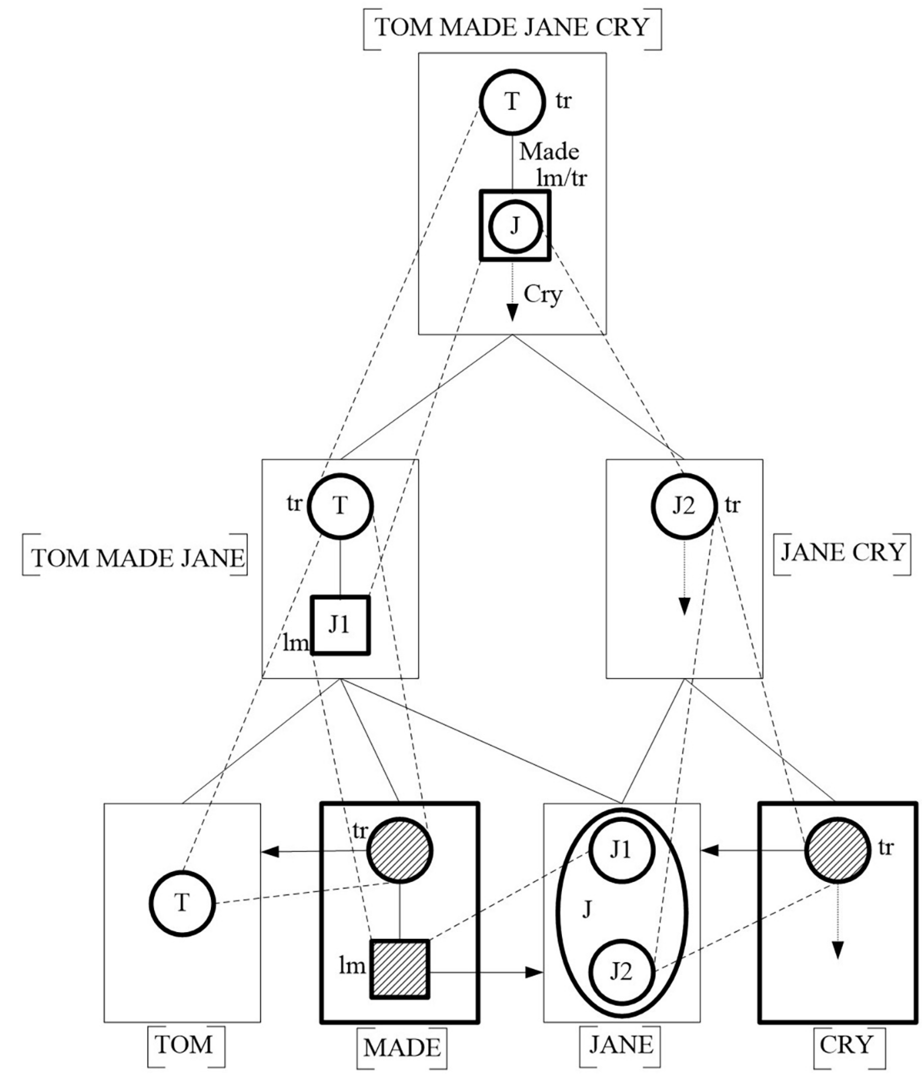


FIGURE 2 | Event alignment in (2) through semantic fusion model.

two verbs. Through the verification of the event alignment in agent/agent, agent/experiencer, and patient/experiencer semantic fusion types by using the semantic fusion model, the model is found to work in the same way as the event alignment in patient/agent and patient/patient semantic fusion types. So, here the detailed alignment process is not provided with more examples and figures. But one point is very apparent; in agent/agent, agent/experiencer, and patient/experiencer semantic fusion types, N's two semantic profiles sanction its entry into two processes profiled by V_1 and V_2 and the two processes are aligned by N's simultaneous participation. The simultaneous participation makes N a psychological passage for V_1 and V_2 to align into a composite event.

In this section, the realization of semantic fusion is discussed by giving an exact account of the event alignment in Chinese and English simple sentence constructions containing two verbs. From the illustration of the event alignment in simple sentence constructions, it is clear that the double semantic profiles of the same syntactic element offer a convincing explanation for the correspondence between double syntactic identities and the double semantic roles. The alignment process of events in these constructions reveals that the two processes are combined into a composite event through the psychological passage N.

(iv) Discussion of the test results

The 20 constructions in the test are in three different forms: SV_1NV_2, SV_1V_2N, and SV_1NAdj. The results of the three tasks in the test are analyzed using R software (R Core Team, 2021). Here, a few sample test constructions are listed in **Table 2**.

The positive results of the three tasks are revealed in **Figure 4** (the blue, orange, and gray columns in the figure stand for the three tasks of the test, which are numbered as ①, ②, and ③, respectively). About task ① in the test, among the 48

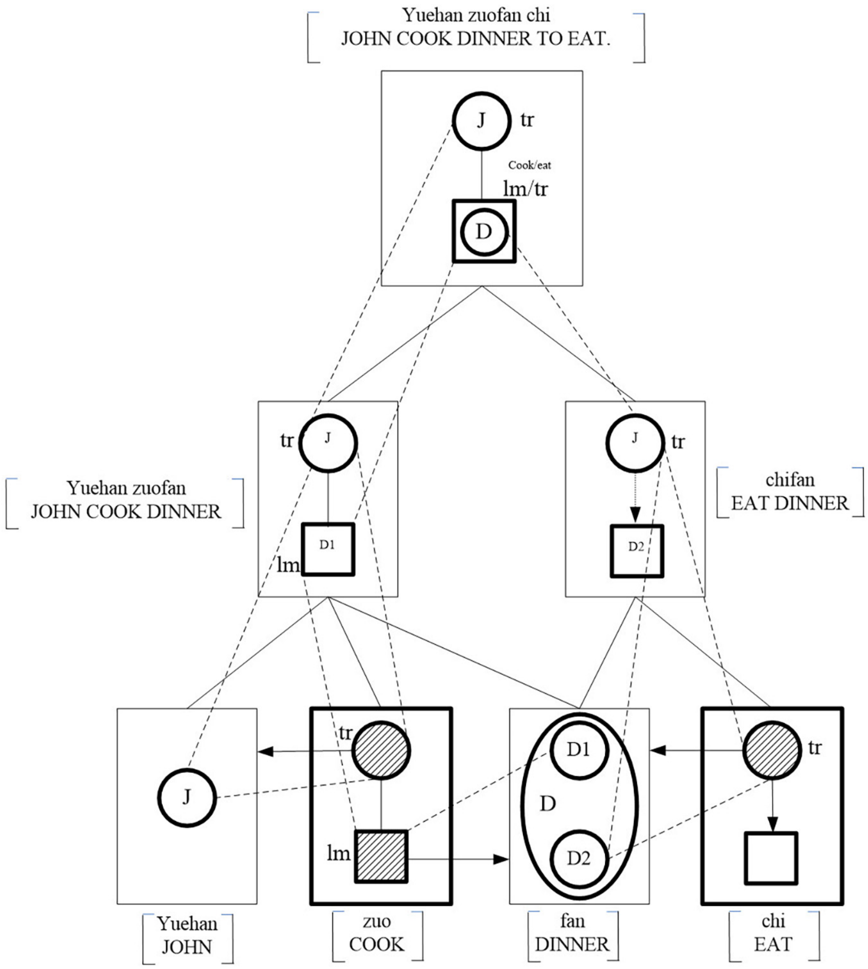


FIGURE 3 | Event alignment in (3) through semantic fusion model.

TABLE 2 | Three sample test constructions.

Forms of the construction	**Construction in Chinese/English**	**Three tasks for each construction**
SV_1NV_2	*zhangsan qing lisi zuo baogao.* Zhang San invite Li Si make report Zhang San invites Li Si to make a report. I saw Tom come out of the house.	① The noun underlined has two semantic roles Yes() No () ② construction judgment: SV_1NV_2 () SV_1V_2N () SV_1NAdj. () ③ V_1 and V_2 are aligned into an event through lisi' Li Si' Agree () Basically agree () Disagree () Basically disagree ()
SV_1V_2N	*women da yingle daishou.* we beat succeed opponent	
SV_1NAdj.	*dajia taoyan xiaoliu xuwei.* everyone detest Xiaoliu hypocritical Everyone dislike Xiaoliu because he is hypocritical.	

participants, on average, 46 participants agree that N possesses two semantic roles in SV_1NV_2, 35 participants maintain that N has two semantic roles in SV_1V_2N, and 42 participants point out the N's two semantic roles in AV_1NAdj. About task ②, on average, 46 participants can clearly differentiate the three linguistic structures, namely, a (SV_1NV_2), b (SV_1V_2N), and c (SV_1NAdj.). About task ③, on average, 45 participants choose "agree" or "basically agree," which means they agree that V_1 and V_2 are aligned into a complete event through N in SV_1NV_2 constructions. Averagely, two participants choose "disagree" that

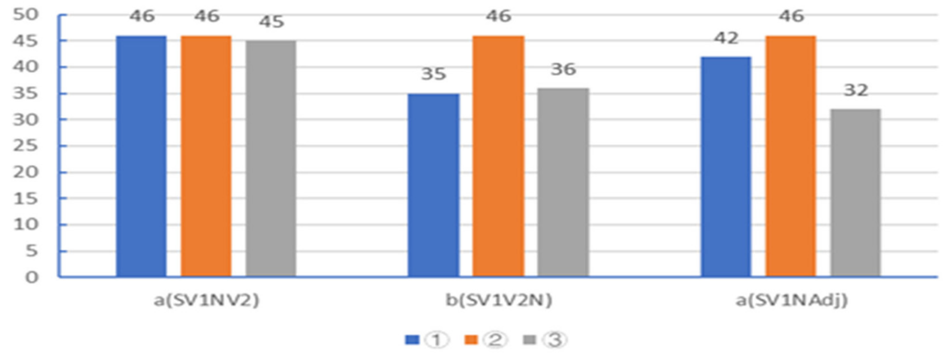


FIGURE 4 | The positive distribution of event semantic fusion in the test.

N plays a bond function in the event formation in SV_1NV_2 constructions. On average, 36 participants choose "agree" and nine participants choose "disagree" about the bond function of N in the formation in SV_1V_2N constructions. About SV_1NAdj. constructions, the results of the test are as follows: 32 participants, on average, agree or basically agree that N is the bond in combining V_1 and V_2 into an event. From the results of the test, it is clear that the syntactic element N is very crucial in the alignment of the event in SV_1NV_2 constructions. N becomes the psychological passage in the cognitive alignment of the event in SV_1NV_2 constructions. Through the verification of the semantic fusion model in Chinese and English constructions, it is found that the semantic fusion model is effective in the explanation of event cognitive alignment and that semantic fusion has solid psychological reality and is, in essence, a basic cognitive ability for people to perceive and process events in the objective world. The results of the test clearly reveal that semantic roles (semantic role overlap/fusion) have a strong psychological reality. The bond function of the syntactic element N in connecting two actions into an event in language has a strong psychological reality.

MAJOR FINDINGS

This study, based on the observation and investigation of a large collection of data and the empirical testing of semantic fusion model in Chinese and English simple sentence constructions containing two verbs, establishes a cross-lingual, cognitive model of event alignment by means of semantic fusion and provides a new perspective for the study of event integration in language. Through the analysis of semantic fusion in both Chinese and English simple sentence constructions containing two verbs, it is found that semantic fusion as the event integration and construal is not language-specific and that semantic fusion is the necessary psychological condition for the alignment of semantic components in construction.

LIMITATIONS AND FUTURE CONCERNS

This study has provided a cognitive explanation for event alignment based on the theory of psychology and cognitive linguistics. However, there are some limitations that are insightful for future concerns. First, the data coverage is relatively small. A large-scale corpus from more languages will make a more comprehensive picture of the cognitive mechanism of event alignment in language. Second, the study is an empirical analysis, which needs to be supported by complicated ERP experiments to prove the effectiveness of the semantic fusion model in the event alignment. Third, the role that the conceptualizer plays in the understanding of event alignment is also a future concern.

CONCLUSION

This study explores the cognitive alignment of events in Chinese and English simple sentence constructions containing two verbs within the framework of psychology and cognitive linguistics and finds that semantic fusion is rooted in the mental conceptualization of language users and is a common psychological and cognitive behavior. Through the double semantic profiling of an autonomous element, the different attributes of an entity are doubly activated, allowing it to enter different processes as a participant. The double semantic profiles of the same autonomous element make it have the ability to fuse two semantic roles into one participant, establishing a psychological passage by which speakers or conceptualizers process events in constructions.

AUTHOR CONTRIBUTIONS

XL has made a direct and intellectual contribution to the work and got it ready for its publication.

FUNDING

This research was granted by the National Social Science Fund of China (19BYY112), entitled Research on Semantic Overlap in Chinese and English, which is hosted by the author.

REFERENCES

Allan, K. (1986). *Linguistic meaning: Vols. 2*. London: Routledge.

Blackburn, J. M. (1940). A review of gestalt psychology. *J. Ment. Sci.*. 360, 1–28. doi: 10.1192/bjp.86.360.1

Bicknell, K., Elman, J. L., Hare, M., McRae, K., and Kutas, M. (2010). Effects of event knowledge in processing verbal arguments. *J. Memor. Lang.* 63, 489–505. doi: 10.1016/j.jml.2010.08.004

Boeckx, C., and Horstein, N. (2004). Movement under control. *Linguist. Inquir.* 2, 431–452. doi: 10.1162/0024389041402625

Bloomfield, L. (2002). *Language*. Beijing: Foreign Language Teaching and Research Press.

Carey, S. (2011). Précis of the Origin of Concepts. *Behav. Brain Sci.* 34, 113–124. doi: 10.1017/S0140525X10000919

Carlson, G. (1998). "Thematic Roles and the Individuation of Events," in *Events and Grammar, vol 70*, ed. S. Rothstein (Dordrecht: Springer), 35–51. doi: 10.1007/978-94-011-3969-4_3

Chen, J. (2021). "Research and Construction of the Holistic View of Book Binding Design from the Perspective of Gestalt Psychology," in *Proceedings of 2nd International Conference on Humanities, Arts, and Social Science* (HASS 2021),(Beijing: Boya Century Publishing) 78–84. doi: 10.26914/c.cnkihy.2021.030073

Chomsky, N. (2010). *Lectures on Government and Binding: The Pisa Lectures*. Berlin: De Gruyter Mouton, doi: 10.1515/9783110884166

Croft, W. (1991). *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press, doi: 10.2307/416423

Croft, W. (2012). *Verbs: aspect and Causal Structure*. Oxford: Oxford University Press, doi: 10.1093/acprof:oso/9780199248582.001.0001

Davidson, D. (1967). *The Logical Form of Action Sentences*. Oxford: Clarendon Press, doi: 10.1093/0199246270.003.0006

Dowty, D. (1986). "Thematic roles and semantics," in *Proceedings of the twelfth annual meeting of the Berkeley linguistics society*, (Berkeley: Berkeley Linguistics Society).

Fauconnier, G., and Turner, M. (1996). "Blending as a central process of grammar," in *Conceptual Structure,Discourse,and Language*, ed. A. Goldberg (Stanford: CSLI), 113-130.

Feng, L. J., and Feng, L. P. (2018). Restriction of the event structure to the syntactic structure in verb-copying VR construction. *J. Yunnan Norm. Univ.* 3, 40–47. doi: 10.16802/j.cnki.ynsddw.2018.03.005

Fillmore, C. J. (1968). "The case for case," in *Universals in linguistic theory*, eds E. Bach and R. T. Harms (New York: Rinehart and Winston), 1–88.

Givón, T. (1990). *Syntax: a Functional-Typological Introduction:Vol. 2*. Amsterdam: John Benjamins, doi: 10.1075/z.17

Givón, T. (2001). *Syntax: An Introduction Volume II*. Amsterdam: John Benjamins Publishing Company.

Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press, doi: 10.1016/S0378-2166(97)81937-6

Gruber, J. S. (1965). *Studies in Lexical Relations*.[Ph.D thesis] Boston: MIT Press.

Huang, C. (1982). *Logical Relations in Chinese and the Theory of Grammar*. New York: Garland Publishing, Inc.

Imbert, C. (2012). Path: Ways typology has walked through it. *Lang.Linguist. Compass* 6, 236–258. doi: 10.1002/lnc3.329

Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.

Jackendoff, R. (1976). Toward an Explanatory Semantic Representation. *Linguist. Inquir.* 7, 89–150. http://www.jstor.org/stable/4177913 doi: 10.1016/j.neuropsychologia.2015.10.023

Jackendoff, R. S. (1983). *Semantics and Cognition*. Cambridge: MIT Press.

Jackendoff, R. (1990). *Semantic Structures*. Cambridge: MIT Press.

Jiang, K. X., Zhao, Y. H., and Cui, R. Y. (2021). Natural language sentence matching method fusion of high-level and low-level semantic information. *Appl. Res. Comput.* 4:6. doi: 10.19734/j.issn.1001-3695.2021.09.0397

Kistler, M. (2006). *Causation and Laws of Nature*. London: Routledge.

Koffka, F. (1935). *Principles of Gestalt Psychology*. London: Kegan Paul.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar: Theoretical Prerequisites. Vol. 1*. Stanford: Stanford University Press.

Langacker, R. W. (2012). Access, Activation, and Overlap: Focusing on the Differential. *J.Foreig. Lang.* 1, 1–24.

Levin, B. (1993). *English Verb Classes and Alternations*. Chicago: The University of Chicago Press.

Li, F. Y. (2019). Evolutionary order of macro-events in Mandarin. *Rev.Cogn. Ling.* 17, 155–186. doi: 10.1075/rcl.00030.li

Li, F. Y. (2020). Macro-event Hypothesis and its empirical studies in Mandarin. *Foreign Lang. Teach.Res.* 3, 349–360. doi: 10.19923/j.cnki.fltr.2020.03.003

Li, X. L. (2015). *A Cognitive Grammar Approach to Semantic Overlapping and Discourse Funciton in Jianyu Construction in Chinese*. Beijing: Science Press.

Li, X. L. (2019). On the cognitive mechanism of syntactic alignment through semantic overlap-ping in Chinese and English. *Stud. Foreign Lang.* 3, 11–18.

Liu, D. Q. (2017). Split between Verb-complement Constructions and Serial Verb Constructions in Chinese Syntactic Inventory. *Lang. Teach. Linguist. Stud.* 2, 1–16.

Lv, S. X. (1979). *Grammatical analysis of Chinese*. Beijing: The Commercial Press.

Metusalema, R., Kutas, M., Urbacha, T. P., Hare, M., McRae, K., and Elmana, J. L. (2012). Generalized event knowledge activation during online sentence comprehension. *J. Mem. Lang.* 4, 545–567. doi: 10.1016/j.jml.2012.01.001

Michotte, A. E. (1946). *The Perception of Causality*. New York: Basic Books.

Niu, B. Y. (2008). A cognitive analytical model of Autonomy/Dependence Alignment. *Foreign Lang. Their Teach.* 1, 1–5.

Parsons, T. (1990). *Events in Semantics of English: A Study in Subatomic Semantics*. Cambridge: MIT Press.

Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge: The MIT Press.

R Core Team (2021). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

Rappaport, H. M. (2008). "Lexicalized meaning and the internal temporal structure of events," in *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect*,ed S. Rothstein (Rothstein Amsterdam: John Benjamins), 13–42. ed. S. doi: 10.1075/la.110.03hov

Rissman, L., and Majid, A. (2019). Thematic roles: Core knowledge or linguistic construct? *Psychon. Bull. Rev.* 26, 1850–1869. doi: 10.3758/s13423-019-01634-5

Sato, M, Niikuni, K, Schafer, A. J., and Koizumi, M. (2020). Agentive versus non-agentive motions immediately influence event apprehension and description: an eye-tracking study in a VOS language. *J. East Asian Linguist.* 4, 211–236. doi: 10.1007/s10831-020-09205-9

Shipley, T., and Zacks, J. (2008). *Understanding Events*. Oxford: Oxford University Press.

Strickland, B. (2017). Language Reflects "Core" Cognition: A New Theory About the Origin of Cross-Linguistic Regularities. *Cognitive Sci.* 1, 70–101. doi: 10.1111/cogs.12332

Talmy, L. (1985). "Lexicalization patterns: Semantic structure in lexical forms," in *Language Typology and Syntactic Description. Vol. III: Grammatical Categories and the Lexicon*, ed. T. Shopen (Cambridge: CUP).

Talmy, L. (1991). "Path to realization: A typology of event conflation," in *Proceedings of the17th Annual Meeting of the Berkeley Linguistics Society*, (Berkeley: Berkeley Linguistics Society), 480-519.

Talmy, L. (2000a). *Toward a Cognitive Semantics. Vol.1*. Cambridge: the MIT Press.

Talmy, L. (2000b). *Toward a Cognitive Semantics. Volume II: Typology and Process in Concept Structuring*. Cambridge: the MIT Press.

Van, V. R. D. (2005). *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.

Vendler, Z. (1957). Verbs and times. *Philos. Rev.* 66, 143–160. doi: 10.2307/2182371

Viswanatha, N., Zlatev, J, Duggirala, V., Weijer, J, V, D. (2018). Holistic s patial semantics and post-Talmian motion event typology: A case study of Thai and Telugu. *Cognitive Semiotics* 2, 1–27.

Wen, X., and Yin, H. (2018). Event Structure Analysis of the Constructionalization of De Emotion-caused Constructions. *J. Northwest Norm. Univ.* 4, 50–55. 2018.04.006 doi: 10.16783/j.cnki.nwnus

Wolff, P. (2003). Direct causation in the linguistic coding and individuation of causal events. *Cognition* 88, 1–48. doi: 10.1016/S0010-0277(03)00004-0

Wu, J. C., and Liang, B. S. (1992). *Sentence Structure of Modern Chinese and Its Analysis*. Beijing: Language and Culture Press.

Wu, Y. J., and Thierry, G. (2011). Event-related brain potential investigation of preparation for speech production in late bilinguals. *Front. Psychol.* 2:114. doi: 10.3389/fpsyg.2011.00114

Wu, Y. J., and Thierry, G. (2017). Brain potentials predict language selection before speech onset in bilinguals. *Brain Lang.* 171, 23–30. doi: 10.1016/j.bandl.2017.04.002

Xing, X. (2004). *Jianyu Construction in Modern Chinese*. Beijing: Communication University of China Press.

Xing, F. Y. (1986). *Modern Chinese*. Beijing: Higher Education Press.

Zhang, H. L., and Pan, H. H. (2019). Inflectional rhyme change and the grammaticalization of event structure. *Stud.Chinese Lang.* 1, 74–88.

Zhang, Y. F. (1999). Pragmatic studies of "having sentence". *J. Henan Univ. Soc. Sci.* 3, 19–23.

Zhang, Y. J. (2001). *Grammatics of Inscriptions on Tortoise Shells of the Shang Dynasty*. Shanghai: Academic Press.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 19 May 2022
doi: 10.3389/fpsyg.2022.892822

Speaker Accent Modulates the Effects of Orthographic and Phonological Similarity on Auditory Processing by Learners of English

Candice Frances[1,2,3]*, *Eugenia Navarra-Barindelli*[1,2] *and Clara D. Martin*[1,4]

[1] Basque Center on Cognition, Brain and Language (BCBL), Donostia, Spain, [2] Department of Social Sciences and Law, The University of the Basque Country (UPV/EHU), Donostia, Spain, [3] Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands, [4] Ikerbasque, Basque Foundation for Science, Bilbao, Spain

The cognate effect refers to translation equivalents with similar form between languages—i.e., cognates, such as "band" (English) and "banda" (Spanish)—being processed faster than words with dissimilar forms—such as, "cloud" and "nube." Substantive literature supports this claim, but is mostly based on orthographic similarity and tested in the visual modality. In a previous study, we found an inhibitory orthographic similarity effect in the auditory modality—i.e., greater orthographic similarity led to slower response times and reduced accuracy. The aim of the present study is to explain this effect. In doing so, we explore the role of the speaker's accent in auditory word recognition and whether native accents lead to a mismatch between the participants' phonological representation and the stimulus. Participants carried out a lexical decision task and a typing task in which they spelled out the word they heard. Words were produced by two speakers: one with a native English accent (Standard American) and the other with a non-native accent matching that of the participants (native Spanish speaker from Spain). We manipulated orthographic and phonological similarity orthogonally and found that accent did have some effect on both response time and accuracy as well as modulating the effects of similarity. Overall, the non-native accent improved performance, but it did not fully explain why high orthographic similarity items show an inhibitory effect in the auditory modality. Theoretical implications and future directions are discussed.

Keywords: bilingualism, auditory processing, cognates, phonology, orthography, lexical decision, typing

OPEN ACCESS

Edited by:
Haoyun Zhang,
The Pennsylvania State University (PSU), United States

Reviewed by:
Paz Suárez-Coalla,
University of Oviedo, Spain
Laura M. Morett,
University of Alabama, United States

****Correspondence:***
Candice Frances
candice.frances@ncf.edu

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *09 March 2022*
Accepted: *19 April 2022*
Published: *19 May 2022*

Citation:
Frances C, Navarra-Barindelli E and Martin CD (2022) Speaker Accent Modulates the Effects of Orthographic and Phonological Similarity on Auditory Processing by Learners of English. Front. Psychol. 13:892822. doi: 10.3389/fpsyg.2022.892822

INTRODUCTION

As has been stated repeatedly in the literature, bilinguals are not "two monolinguals in one" (Grosjean, 1989, 1997, 1998; Grosjean and Nicol, 2007). There is ample evidence that a bilinguals' two languages interact in many ways (Caramazza and Brones, 1979; Grosjean, 2001; Lagrou et al., 2011; Blumenfeld and Marian, 2013). One common evidence of this interaction is the cognate effect (Caramazza and Brones, 1979; Cristoffanini et al., 1986; de Groot and Nas, 1991; Sanchez-Casas et al., 1992; Dijkstra et al., 1998, 1999; Schwartz et al., 2007; Voga and Grainger, 2007). The cognate effect refers to words that are similar in form and meaning between a bilingual's languages, activating the non-target language and thus having a processing advantage over words that only share meaning but not form—i.e., non-cognates or low similarity items. For example, "band" and "banda" are considered cognates between English and Spanish and would thus have an advantage over "cloud" and "nube"—non-cognates or low similarity items. One important distinction to make is that between orthographic/phonological cognates and "false cognates," meaning interlanguage homographs and homophones that do not align in meaning. For example, <once> means "eleven" in Spanish, but in English it means "one time." False cognates, given their semantic misalignment, do not share the same facilitatory effects of cognates in recognition (Dijkstra and van Heuven, 2002) or translation (Janke and Kolokonte, 2015). The cognate effect is quite well described with respect to orthographic cognates—namely, words that are orthographically similar between languages—leading to several processing benefits (van Orden, 1987; Duyck et al., 2007; Van Assche et al., 2011, 2012; Poort and Rodd, 2017). Given the strong cognate effect observed in bilingual speech and the common assumption that the two languages of a bilingual are co-activated at the phonological level (Costa et al., 2000; Colom, 2001; Colomé and Miozzo, 2010; Sadat et al., 2016), words that are phonologically similar between languages may also influence word processing.

In a prior study on the effect of phonological similarity on lexical processing, we found that there was an inhibitory orthographic similarity effect in the auditory modality (Frances et al., 2021). This study showed that, with greater orthographic similarity between the spoken word in the native (NL) and foreign (FL) languages, response times were slowed and accuracy was reduced. For example, when Spanish-English bilingual participants heard the English word "band," which is both a phonological and an orthographic cognate ("banda" in Spanish), their responses were slower than when they heard "jacket"(/dʒækət/), which is a phonological cognate but an orthographic non-cognate ("chaqueta" pronounced /ʧaketa/ in Spanish). In addition, other studies have found similar cross-modality inhibition [e.g., phonological inhibition in the visual modality (Dijkstra et al., 1999; Lemhöfer and Dijkstra, 2004)]. These results point to an independence but co-activation of representations in both modalities. Not only that, but the cross-modal inhibition suggests that the particular relationship between orthography and phonology (i.e., whether they generally have a one-to-one correspondence or not) in each of the languages of a bilingual can influence the cognate effect.

One possible way of explaining this inhibitory orthographic effect in the auditory modality is through a discrepancy between the listener's (FL) phonological representation of the FL item and the native speaker's production of it. In other words, the listener's NL is likely to affect not only their production of FL words, but also their internal phonological representation of them. In addition, individuals with a transparent—i.e., a language with a one-to-one correspondence between the graphemes and the phonemes—NL are likely to have a stronger reliance on orthography than those with an opaque NL, as orthographic consistency aids the auditory processing of words (Seidenberg and Tanenhaus, 1979; Ziegler and Ferrand, 1998). When the participant's NL is transparent, this distortion is likely to be greater, with more interference in items that are orthographically similar between languages. If that is the case, hearing an orthographic cognate said by a native speaker is likely to mismatch with the FL listener's representation of the item. This would, in turn, slow down the process of verifying that the item is in fact a real word. In language production, there are studies that found increased accentedness in the production of orthographic cognates in FL speakers (Costa et al., 2000; Amengual, 2012; Goldrick et al., 2014). This provides support for the idea that cognates may suffer from a greater phonological influence of the NL. Therefore, for instance, Spanish-English bilinguals would have a stronger Spanish accent when producing "band" as compared to "jacket," with "band" being an orthographic cognate, whereas "jacket" is not. Given that the foreign accent is stronger, it might be the case that the internal representation of "band" is more strongly influenced by grapheme-phoneme correspondence rules in Spanish than that of "jacket." In other words, the native Spanish listener's internal representation of the English word <violin> is /biolin/ because of the orthographic similarity with the Spanish translation <violín>. When they hear /vaɪəˈlɪn/, there is a strong mismatch between the perceived word and its internal representation, making word recognition slower.

Another possible way to explain the inhibitory orthographic effect in the auditory modality is that the auditory stimuli activate incorrect orthographic representations. Speakers often rely on orthography to process phonological items (Seidenberg and Tanenhaus, 1979; Ziegler and Ferrand, 1998), possibly "transcribing" the phonological string into its orthography when doing an auditory lexical decision task (LDT). This is particularly problematic in the case of bilinguals with a transparent NL and an opaque FL, as their NL rules are likely to influence and distort this process. For example, when the participant hears the English word <violin> (/vaɪəˈlɪn/), they transcribe it as <baiolin> using their NL phoneme to grapheme correspondence rules, which is quite different from the correct FL spelling of the word (<violin>). When they hear the word <jacket> (/dʒækət/), they transcribe it into something like <jaket>, which is much closer to the correct orthography for the item in English (<jacket>).

The main aim of the present study is to explain the inhibitory orthographic similarity effect in the auditory modality by testing whether it is due to a mismatch (1) between the internal phonological representation and the aural stimulus or (2) between the constructed and real orthographic representations, or (3) possibly a combination of the two. In addition, we will explore the effects of accent—native vs. foreign—in both detecting and identifying words as well as whether and how this interacts with orthographic similarity between languages.

To test this question, the current study includes the following tasks: (1) an LDT and (2) a typing task, both in the auditory modality. In the auditory LDT, we approximate participants' internal phonological representation (in their FL) by presenting stimuli in the accent that is closest to theirs and the pronunciation they are most accustomed to, produced by a non-native speaker with the same origin (as well as a native speaker as a control condition). For the typing task, participants simply type what they hear (phonological strings including words and pseudowords) when they are presented with the auditory stimuli—produced by both the native and non-native speakers.

Based on the two possible explanations we have presented, there are different expected results. One option is that our first explanation is correct and the inhibitory effect of orthographic

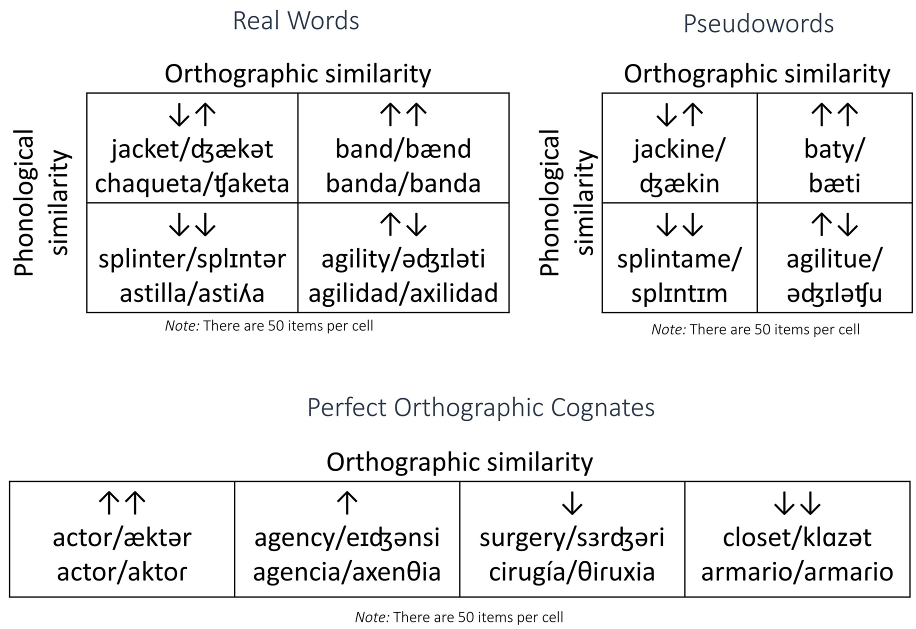


FIGURE 1 | Example stimuli in each similarity condition.

similarity is due to the discrepancy between the internal phonological representation and what the FL listeners are hearing. This would also mean that bilinguals have particularly accented representations of orthographic cognates. If that is the case, (1) participants would show the inhibitory orthographic effect only when they hear the words in the native accent, but not in the non-native accent. In other words, by hearing words in the non-native accent, the stimuli would match their internal representations of the phonological items more closely, thus negating the inhibitory effect of orthographic similarity. On the other hand, in the typing task, participants have extra time to process the stimuli. Therefore, they are unlikely to show the orthographic inhibitory effect in neither the native nor non-native accent. In fact, they are likely to show a facilitatory effect of orthography, as the NL orthography should aid spelling in cases of high similarity. The other option is that our second explanation is correct. In this case, the inhibitory orthographic effect would be due to a mismatch between the constructed and real orthographic representations during aural perception. If so, (2) we would expect to see the inhibitory effect of orthographic similarity in both accents in the LDT. The idea is that, if the phoneme to grapheme correspondence between the NL and FL is the cause of the effect, then the effect should remain unaffected by accent, as this correspondence does not depend on production. We should also see it in the typing task, with an increase of typing errors in orthographically similar words.

In other words, what makes the largest difference between the two hypotheses is that in the first case (1) the inhibitory orthographic similarity effect should disappear in the LDT with the non-native accent and we should see a facilitatory effect of orthographic similarity in spelling in the typing task, whereas in the second case (2) the inhibitory orthographic similarity effect should not disappear in the LDT regardless of accent and there should be an inhibitory orthographic effect in the typing task, as well. In addition, we expect higher performance overall with the non-native accent (see interlanguage speech intelligibility benefit) (Bent and Bradlow, 2003; Xie and Fowler, 2013; Wang and van Heuven, 2015).

METHODS

Participants

Participants were 59 native Spanish speaking adults ($F = 37$, $M_{age} = 27.86$ [SD = 4.42]) from Madrid and Murcia (Spain) with at least an intermediate (B1) level in English. Participants had a minimum score of 40 on the English BEST (de Bruin et al., 2017)—a picture naming task with a maximum score of 65—and of 55% on the English LexTALE (a vocabulary test), which equates to approximately a B2 (upper intermediate) level (Lemhöfer and Broersma, 2012). Participants' average score on the BEST was 60.88 (SD = 4.95) with a range of 42–65. With respect to the LexTALE, their average score was 76.41% (SD = 10.35%) with a range of 55–99%. Their average self-reported age of acquisition of English was 6.61 (SD = 3.08) years old, with a range of 3–19 years of age. All participants provided informed consent before taking part in the experiment, which was conducted in accordance with the Declaration of Helsinki and approved by the Basque Center on Cognition, Brain and Language ethics committee (approval number 12762). Participants were paid for taking part in the experiment.

TABLE 1 | Means, standard deviations, and statistics for variables stimuli were matched on.

Orthographic similarity	Low		High		Identical	
Phonological similarity	**Low**	**High**	**Low**	**High**	**High**	**Statistic**
English frequency	25.21 (21.85)	30.81 (43.55)	42.61 (80.29)	24.63 (21.65)	31.64 (49.27)	$F(4{,}245) = 1.117$, $p = 0.349$, $BF_{01} = 16.175$
English log frequency	1.22 (0.45)	1.12 (0.61)	1.27 (0.57)	1.23 (0.40)	1.17 (0.51)	$F(4{,}245) = 0.667$, $p = 0.616$, $BF_{01} = 32.943$
Spanish frequency	50.37 (56.59)	67.63 (80.91)	79.34 (106.46)	58.89 (58.58)	69.22 (101.77)	$F(4{,}245) = 0.864$, $p = 0.486$, $BF_{01} = 24.147$
Spanish log frequency	1.38 (0.63)	1.47 (0.65)	1.47 (0.69)	1.48 (0.64)	1.38 (0.72)	$F(4{,}245) = 0.304$, $p = 0.875$, $BF_{01} = 58.305$
Number of syllables	2.00 (0.88)	2.08 (0.92)	2.00 (0.90)	1.86 (0.76)	1.88 (0.39)	$F(4{,}245) = 0.670$, $p = 0.613$, $BF_{01} = 32.769$
Number of letters	6.36 (1.96)	6.08 (2.06)	6.52 (1.76)	6.38 (2.00)	5.84 (1.17)	$F(4{,}245) = 1.126$, $p = 0.345$, $BF_{01} = 15.940$
Number of phonemes	5.94 (2.08)	5.98 (2.20)	5.46 (1.76)	5.88 (1.83)	5.56 (1.07)	$F(4{,}245) = 0.839$, $p = 0.502$, $BF_{01} = 25.110$

Values are means with standard deviations in parentheses. The N in all cases is 50.

Stimuli

Stimuli consisted of 300 English words and 300 pseudowords. The words were taken from Frances et al. (2021). These were divided into six categories (see **Figure 1**). Four of them consisted of a Latin square between orthographic and phonological similarity: high orthographic/high phonological similarity, high orthographic/low phonological similarity, low orthographic/high phonological similarity, and low orthographic/low phonological similarity. In our case, the terms "high similarity" and "low similarity" were favored over "cognate" and "non-cognate" because the distribution of similarity we based this distinction on is linear rather than dichotomous. The categories of high and low similarity were determined using a median split of ALINE distance (Kondrak, 1999, 2000).

Another group contained items in the extreme of the orthographic similarity distribution: orthographically identical words or perfect cognates. Finally, we included a group of extreme dissimilarity items in order to balance the number of high and low similarity items (see **Figure 1** for an example of each).

As mentioned above, high and low phonological and orthographic similarity were defined by median split using inverse ALINE distance (Kondrak, 1999, 2000). ALINE distance is a normalized measure of string alignment that provides a value of dissimilarity (with inverse ALINE distance being a measure of similarity) between words. This can be used to compare translations between languages. For phonology, we placed the median split at 0.740. The high similarity range of inverse ALINE distance values was 0.741–0.951 and for the low similarity group, the range was 0.195–0.736. For orthography, we used the median split was at 0.770. The high similarity range of inverse ALINE distance values was 0.771–0.982 and for the low similarity group, the range was 0.360–0.769. ALINE distances were calculated using the alineR package for R (Downey et al., 2017).

For the main manipulation, we focused on the first four groups. The orthographically identical translation group (perfect orthographic cognates) were included to assess the "special status" of those words, also called "perfect cognates." Note that no phonologically identical group was included, since the differences in phonology between English and Spanish made it impossible to find enough items.

All six groups of items were matched on the following variables: word frequency (raw and logarithmic), word frequency of the Spanish translation (raw and logarithmic), number of syllables, number of letters, and number of phonemes (see **Table 1** for means, standard deviations, and statistics), all extracted from CLEARPOND (Marian et al., 2012). Pseudowords were created by exchanging the last two phonemes (2 or 3 letters) between words used in the task (e.g., lens/lɛnz changed to lert/lɛrt and airport/*ɛrpɔrt to airpons*/ɛrpɔnz). This way, the number of letters and phonemes remained constant and all items had to be listened up to the penultimate phoneme in order to differentiate the word from the pseudoword. In other words, we maintained the uniqueness point of target words constant between stimuli and as late as possible.

There was a total of 50 words per group, for a grand total of 300 words. There were also 50 pseudowords per condition—one matched to each word. All words and pseudowords were presented once in each accent condition (see below).

Native accent auditory stimuli were recorded in a quiet recording room by a native speaker of English with a general American accent (Labov et al., 2006) and following the pronunciation reported in the Carnegie Mellon CMU dictionary (Carnegie Mellon, 2020). Foreign accented (non-native speaker) auditory stimuli were similarly recorded by a native speaker of Spanish. Importantly, the non-native speaker did not add or remove phonemes to the words, they simply used the closest Spanish phoneme. For example, for <jacket> they produced /ɟaket/ instead of /dʒækət/ and for <violin> they produced /baiolin/ instead of /vaɪəˈlɪn/. All stimuli were normalized to 1dB and cut with 500 ms of silence before and after, using Audacity (Audacity Team, 2018). They were recorded at a frequency of 44.1 kHz and 32 bits.

Procedure

Participants were pre-selected using a form in which they reported age, language background, and were tested on their level of English [BEST (de Bruin et al., 2017) and LexTALE (Lemhöfer and Broersma, 2012)]. This was completed using LimeSurvey (Schmitz LPT/C, 2019). The stimuli in the testing sessions were presented using Opensesame (Mathôt et al., 2012) through the JATOS platform (Lange et al., 2015).

The data collection consisted of two sessions: one with the native accented stimuli and the other with the non-native accented stimuli. The order of sessions was counterbalanced between participants, and there were at least 2 weeks between the two sessions. In each session, participants first carried out an LDT and then a typing task, with the same stimuli. For the LDT, participants were presented with all 600 items randomly mixed. For each trial, they would see a fixation cross for 500 ms, then hear the word and have 2,500 ms from stimulus onset to respond whether it was a real word or not using the F and J keys on the keyboard (counterbalanced between participants). Participants were provided a self-paced break every 150 words. After the LDT, they carried out the typing task. For the typing task, participants would have a fixation cross for 500 ms, then they heard the item twice. After the first utterance of the word, they were presented with a textbox in order to type in the item. They had unlimited time to type and they could erase and retype freely. The stimulus recordings were the same for the LDT and the typing task; in one session they were both native-accented and in the other they were both non-native-accented.

ANALYSIS

Lexical Decision Task

The duration of each sound file was subtracted from the corresponding response times. Outliers were defined as values two standard deviations from the mean for each condition in each participant. In total, 4.55% of data was removed due to outliers. In all response time analyses, only correct responses were taken into account.

Data was analyzed categorically—low and high similarity as well as low similarity, high similarity, and perfect cognates—using ANOVAs. Accuracy was assessed using A', a measure of signal detection (Zhang and Mueller, 2005), calculated using the Psycho package (Makowski, 2018) for R. This measure was favored as it takes into account participant response tendencies. For A', all analyses were by participant, as we could not carry out by item analyses using A'—participants cannot be paired up as the stimuli were, to provide measures of correct rejections and false alarms.

Analyses were run using JASP (JASP Team, 2020). Additional analyses evaluating orthographic and phonological similarity linearly are provided in the **Supplementary Materials**.

Typing Task

Due to technical errors, two participants had a reduced number of trials: one had 584 trials out of 600 for Day 1 and another had 514 out of 600 for Day 2. Finally, one participant had to be excluded because he was missing all of the typing task data for Day 1, leaving 58 participants—two of which had partial data. For this task, we carried out the same analyses as with the LDT: a three-way ANOVA (accent by orthographic similarity by phonological similarity) and a two-way ANOVA (accent by orthographic similarity including perfect cognates). Accuracy was defined as the number of correctly typed items (i.e., correctly identified and with no typos or spelling errors). All analyses were also run as linear models using ALINE distance (Lange et al., 2015) instead of the binary accuracy. Given that the pattern and significant effects and interaction were strictly identical, we omitted these from the main text. Additional analyses evaluating orthographic and phonological similarity linearly are provided in the **Supplementary Materials**.

RESULTS

Lexical Decision Task

Phonological and Orthographic Similarity Effects on Response Time

We carried out a two accent (native/non-native) by two orthographic similarity (high/low) by two phonological similarity (high/low) repeated measures ANOVA on response times. There was a significant interaction between orthography and phonology [$F_1(1,58) = 14.296$, $p < 0.001$, $\eta^2_p = 0.198$, absent in by item analysis $F_2(1, 196) = 1.597$, $p = 0.202$, $\eta^2_p = 0.008$] as well as a three-way interaction between accent, orthography, and phonology [$F_1(1,58) = 8.834$, $p = 0.004$, $\eta^2_p = 0.132$, absent in by item analysis $F_2(1, 196) = 0.348$, $p = 0.556$, $\eta^2_p = 0.002$]. See **Figure 2** for average response times. We observed a main effect of accent (slower for non-native) only in the by item analysis [$F_1(1,58) = 0.302$, $p = 0.585$, $\eta^2_p = 0.005$; $F_2(1, 196) = 29.484$, p

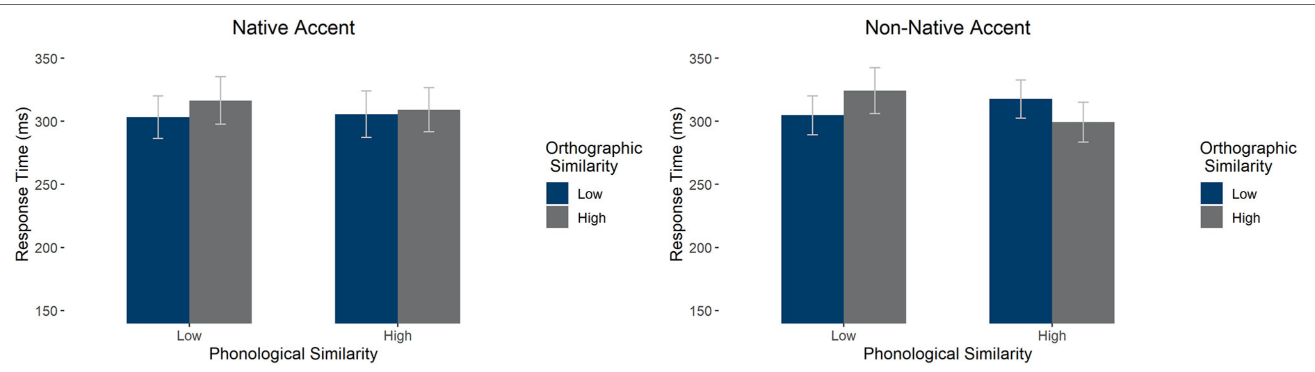


FIGURE 2 | Average response times in the LDT by accent, orthographic similarity, and phonological similarity condition. Error bars mark 95% confidence intervals.

< 0.001, $\eta_p^2 = 0.131$]. There were no other significant effects (all p's > 0.1)

Follow-up two-way ANOVAs exploring phonological and orthographic similarity effects independently for native and non-native accent showed different effects in the two accent conditions. In the native accent, there was a main effect of orthography [$F_1(1,58) = 5.522$, $p = 0.022$, $\eta_p^2 = 0.087$], with low orthographic similarity items being responded to faster than high orthographic similarity items. There was no main effect of phonological similarity and no interaction (p's > 0.1). In contrast, in the non-native accent, there was an interaction between phonology and orthography [$F_1(1,58) = 21.710$, $p < 0.001$, $\eta_p^2 = 0.272$], such that, when orthography and phonology aligned (i.e., high similarity in both orthography and phonology or low similarity in both orthography and phonology), response times were reduced compared to the cases in which the two did not align (i.e., high phonological similarity but low orthographic similarity or vice-versa). There was also a marginal main effect of phonology [$F_1(1,58) = 3.352$, $p = 0.072$, $\eta_p^2 = 0.055$], such that high phonological similarity items were responded to faster than low phonological similarity items. There was no main effect of orthographic similarity ($p > 0.1$). To summarize, in native speech, orthographic similarity led to slower processing of both high and low phonological similarity items. In non-native speech, the pattern was similar for low phonological similarity items—responded to slower in the case of high orthographic similarity, but the pattern was reversed for high phonological similarity—they were responded to faster in the case of higher orthographic similarity.

Phonological and Orthographic Similarity Effects on Signal Detection (A')

We carried out an ANOVA on the effects of accent, phonological similarity, and orthographic similarity on signal detection. We found a main effect of orthography [$F(1, 58) = 51.330$, $p < 0.001$, $\eta_p^2 = 0.469$], qualified by an interaction between orthography and phonology [$F(1, 58) = 55.249$, $p < 0.001$, $\eta_p^2 = 0.488$]: Low phonological similarity items had higher signal detection when orthographic similarity was low, $F(1, 58) = 87.100$, $p < 0.001$, but this was not the case for high phonological similarity items, $F(1, 58) = 0.747$, $p = 0.391$. See **Figure 3** for average signal detection values (A'). There were no other main effects or interactions (p's > 0.1).

Effects of Perfect Orthographic Cognates on Response Time

We also analyzed the effect of perfect orthographic cognates and accent on response time, selecting only high phonological similarity items for the comparison: We compared high phonological similarity items that had low orthographic similarity, high orthographic similarity, or were perfect cognates. We found a main effect of orthographic similarity, $F_1(2, 116) = 36.774$, $p < 0.001$, $\eta_p^2 = 0.388$ (absent by item $F_2(2, 147) = 1.858$, $p = 0.160$, $\eta_p^2 = 0.025$), such that perfect cognates were responded to significantly slower than both high, $t(58) = 8.169$, $pholm < 0.001$, and low similarity items, $t(58) = 6.346$, $pholm < 0.001$, but the last two groups did not differ significantly, $t(58) = 1.823$, $pholm = 0.071$. There was an interaction between accent and orthography [$F_1(2, 116) = 14.818$, $p < 0.001$, $\eta_p^2 = 0.204$; $F_2(1, 147) = 2.568$, $p = 0.080$, $\eta_p^2 = 0.034$], such that there was an effect of accent only for the perfect cognates, $t(58) = 2.954$, $pholm = 0.041$, with participants responding slower in the native than the non-native accent. There was an effect of accent by item [$F_1(1, 58) = 0.886$, $p = 0.351$, $\eta_p^2 = 0.015$; $F_2(1, 116) = 11.408$, $p < 0.001$, $\eta_p^2 = 0.072$]. See **Figure 4** for average response times by condition.

Effects of Perfect Orthographic Cognates on Signal Detection (A')

We also analyzed the effect of perfect orthographic cognates and accent on signal detection, selecting again only high phonological similarity items for the comparison. We found a marginal main effect of accent [$F(1, 58) = 3.960$, $p = 0.051$, $\eta_p^2 = 0.064$], such that there was higher signal detection with the non-native accent, as well as a main effect of orthographic similarity [$F(1, 58) = 8.738$, $p < 0.001$, $\eta_p^2 = 0.131$], such that signal detection was significantly worse for perfect cognates than both high [$t(58) = 3.516$, $pholm = 0.003$] and low similarity items [$t(58) = 3.024$, $pholm = 0.007$], but high and low similarity items did not differ, $p > 0.1$. This effect was qualified by an interaction between accent

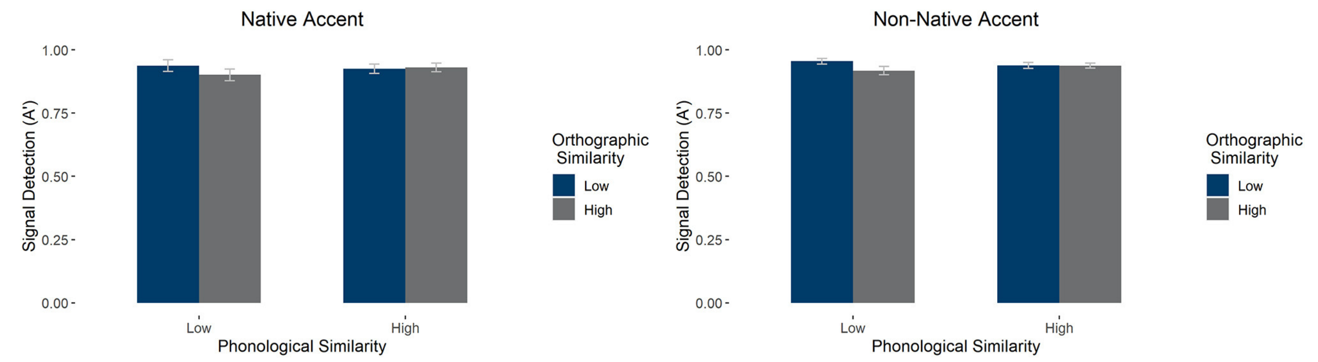


FIGURE 3 | Average signal detection (as measured by A') in the LDT by accent, orthographic similarity, and phonological similarity condition. Error bars mark 95% confidence intervals.

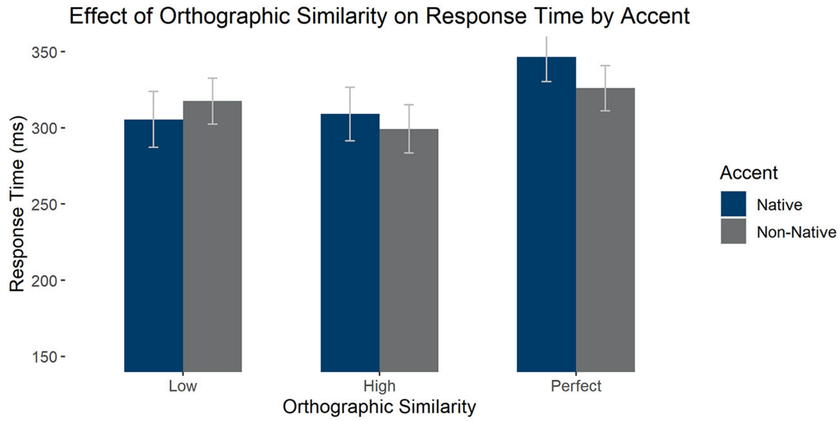


FIGURE 4 | Average response times in the LDT to high phonological similarity words by accent and orthographic similarity condition. Error bars mark 95% confidence intervals.

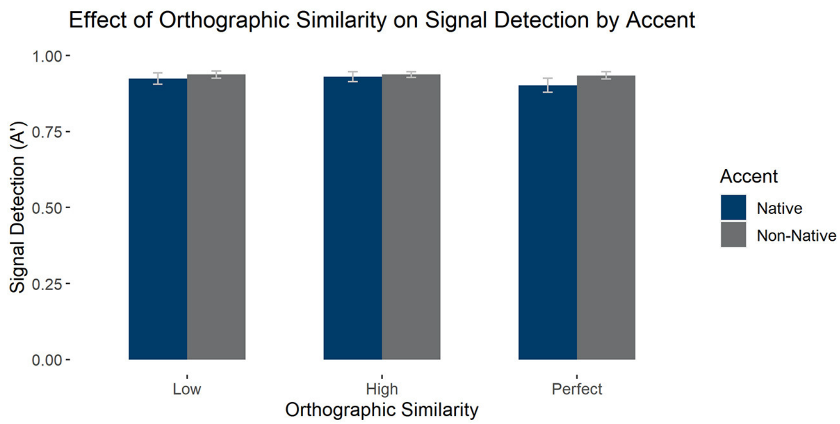


FIGURE 5 | Average signal detection (as measured by A') in the LDT to high phonological similarity words by accent and orthographic similarity condition. Error bars mark 95% confidence intervals.

and orthographic similarity, $F(1, 58) = 6.855$, $p = 0.002$, $\eta_p^2 = 0.106$, such that the effect of orthographic similarity was only present in the native accent [$F(1, 58) = 11.442$, $p < 0.001$, $\eta_p^2 = 0.165$] but not in the non-native accent, $p > 0.1$. See **Figure 5** for average signal detection scores by group.

To summarize, considering response times and signal detection together, there was a strong disadvantage in processing of high phonological similarity items when orthographic similarity was high (slower and less accurate). Importantly, this was predominantly present in the native accent.

Typing Task

Phonological and Orthographic Similarity Effects on Accuracy

There was a main effect of accent—with participants showing higher accuracy in the non-native condition [$F_1(1{,}57) = 15.946$, $p < 0.001$, $\eta_p^2 = 0.219$; $F_2(1, 196) = 9.952$, $p = 0.002$, $\eta_p^2 = 0.048$] and orthographic similarity—with low similarity items leading to higher accuracy [$F_1(1{,}57) = 23.845$, $p < 0.001$, $\eta_p^2 = 0.295$, absent by item $F_2(1, 196) = 0.639$, $p = 0.425$, $\eta_p^2 = 0.003$]. There was a significant interaction between phonological and orthographic similarity [$F_1(1{,}57) = 122.28$, $p < 0.001$, $\eta_p^2 = 0.682$; $F_2(1, 196) = 5.202$, $p = 0.024$, $\eta_p^2 = 0.026$]. This interaction showed that, in both accents, accuracy was greater when orthographic and phonological similarity aligned: Participants were better at typing (i.e., identifying the correct word) when the words were high phonological and orthographic similarity items or low phonological and orthographic similarity items as compared to high phonological and low orthographic similarity items or low phonological and high orthographic similarity items. There was also a marginal interaction between accent and phonological similarity [$F_1(1{,}57) = 3.283$, $p = 0.075$, $\eta_p^2 = 0.055$, absent by item $F_2(1, 196) = 0.408$, $p = 0.524$, $\eta_p^2 = 0.002$] and a three-way interaction [$F_1(1{,}57) = 7.303$, $p = 0.009$, $\eta_p^2 = 0.114$, absent by item $F_2(1, 196) = 1.024$, $p = 0.313$, $\eta_p^2 = 0.005$]. Follow-up simple comparisons on the three-way interaction showed that there was a significant

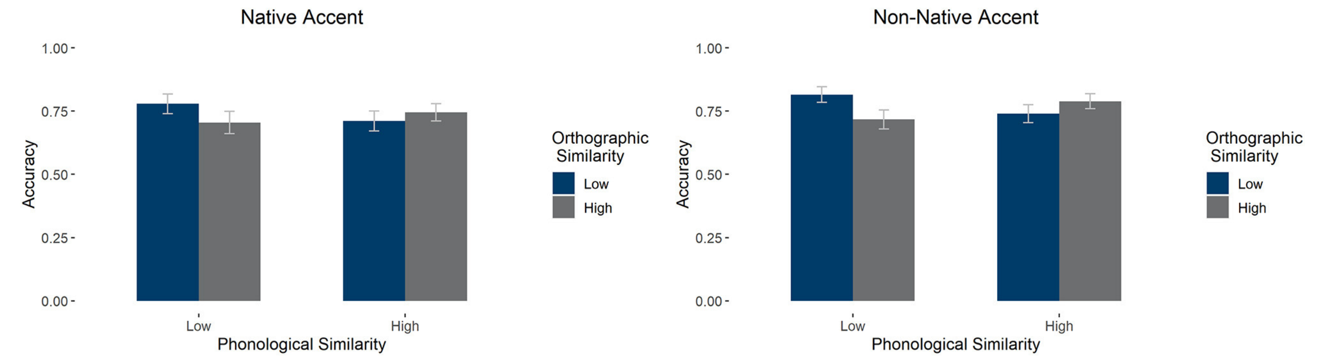


FIGURE 6 | Average accuracy in the typing task by accent, phonological similarity, and orthographic similarity condition. Error bars mark 95% confidence intervals.

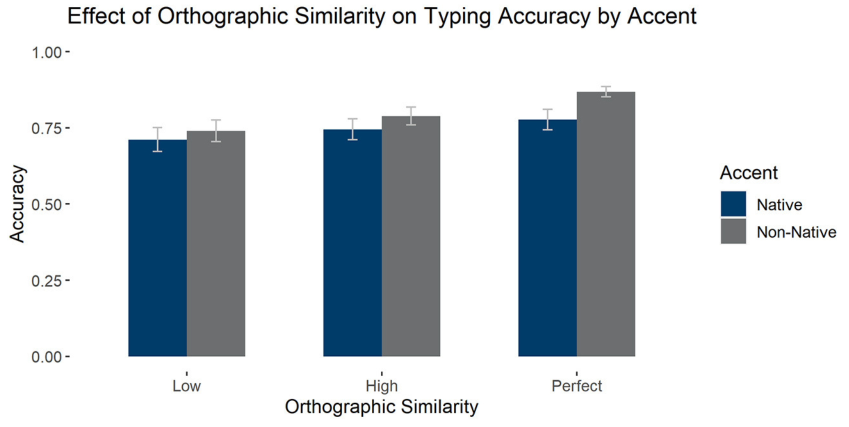


FIGURE 7 | Average accuracy in the typing task for high phonological similarity words by accent and orthographic similarity condition. Error bars mark 95% confidence intervals.

positive effect of non-native accent in all cases [high orthographic and phonological similarity: $t(57) = 4.89$, $pholm < 0.001$; low orthographic and high phonological similarity: $t(57) = 3.13$, $pholm = 0.025$, and low orthographic and phonological similarity: $t(57) = 3.40$, $pholm = 0.014$] except for low phonological and high orthographic similarity items, where there was no effect [$t(57) = 1.22$, $pholm = 1$]. There was no main effect of phonological similarity, nor an interaction between accent and orthographic similarity, p's > 0.1. See **Figure 6** for average accuracy by group.

The Effects of Perfect Orthographic Cognates on Accuracy

We also analyzed the effect of perfect orthographic cognates and accent on accuracy, selecting only high phonological similarity items for the comparison. We found a main effect of accent—such that the non-native accent led to higher accuracy [$F_1(1,57) = 48.673$, $p < 0.001$, $\eta_p^2 = 0.461$; $F_2(1, 147) = 18.151$, $p < 0.001$, $\eta_p^2 = 0.110$] and of orthographic similarity [$F_1(2,114) = 58.019$, $p < 0.001$, $\eta_p^2 = 0.504$; $F_2(2, 147) = 3.071$, $p = 0.049$, $\eta_p^2 = 0.040$], such that perfect cognates were identified more accurately than high [$t(57) = 6.530$, $pholm < 0.001$] and low similarity items [$t(57) = 8.809$, $pholm < 0.001$], while high similarity items were identified better than low similarity items [$t(57) = 5.779$, $pholm < 0.001$]. There was an interaction between accent and orthography [$F_1(2,114) = 19.613$, $p < 0.001$, $\eta_p^2 = 0.256$, absent by item $F_2(2, 147) = 2.144$, $p = 0.121$, $\eta_p^2 = 0.028$] such that the effects of accent were greater in the perfect condition than in the other two. See **Figure 7** for average accuracy by group.

Overall for the typing task looking at orthographic and phonological similarity categorically, we find that the non-native accent aided performance, particularly in the case of perfect cognates. Furthermore, orthographic and phonological similarity interact such that words for which both types of similarity are aligned are typed more accurately than those for which the two are crossed.

DISCUSSION

The current study set out to explain the inhibitory effect of orthographic similarity on auditory word recognition in bilinguals. In a previous study, we showed that orthographic similarity had an inhibitory effect in the auditory modality (Frances et al., 2021). This is in contrast to the facilitatory cognate effect that has been found in the visual modality (Caramazza and Brones, 1979; Cristoffanini et al., 1986; de Groot and Nas, 1991; Sanchez-Casas et al., 1992; Dijkstra et al., 1998, 1999; Schwartz et al., 2007; Voga and Grainger, 2007; but see Schwartz et al., 2007). It should be noted that even though the definition of cognate refers to similarity in form (which should also include the phonological form), they are generally defined orthographically—meaning by similarity in spelling—and studied in the visual modality (Caramazza and Brones, 1979; Sanchez-Casas et al., 1992; Dijkstra et al., 1998; Colom, 2001; Zhang and Mueller, 2005; but see Dijkstra et al., 1999; Schwartz et al., 2007). Our current study finds inhibitory effects of orthographic similarity in the auditory modality and aligns with the results of Frances et al. (2021). Importantly, in our case, we took a step further and tried to understand the origin of this effect—namely, inhibition in processing of orthographically similar words in the auditory modality.

This study explored two alternative hypotheses to explain this effect. One possibility was that the listener's phonological representation of an item in the FL differed from the native production *more* in cases of higher orthographic similarity. Another possibility we explored was that listeners were incorrectly "transcribing" the FL items they heard using their NL orthographic rules (as opposed to their FL rules). Both cases would lead to slower response times and worse identification of items with higher orthographic similarity to their NL when participants were exposed to native accented speech (Frances et al., 2021). Therefore, to test and disambiguate these two possibilities, we had participants carry out both an LDT and a typing task presenting the stimuli in the auditory modality. The items we presented were produced with either a native or a non-native accent (similar to the participants'). If the first hypothesis were true, we would expect that the non-native accent would reduce the discrepancy between the internal representation and the exemplar heard, thus reducing the inhibitory effect of orthographic similarity in the LDT. Furthermore, the orthographic similarity effect should revert and be facilitatory in the typing task. If the second hypothesis were true, we would expect no difference between accents in the LDT, and we would see a detrimental effect of orthographic similarity in the typing task as well as the LDT.

As expected and in replication to Frances et al. (2021), orthographic similarity had a negative effect on signal detection in the auditory LDT, which extended to accuracy in the typing task. For the LDT, we found that using a non-native accent reduced response times and that orthographic similarity was particularly detrimental in cases of high phonological similarity—both in response times and accuracy—specifically in the native accent. We also found that the effect of accent was disproportionately larger for perfect cognates than for high orthographic similarity items. When considering similarity linearly, we found orthographic (inhibitory) and phonological (facilitatory) effects on accuracy. We also found that phonological similarity had the largest effects for words that were more orthographically similar and presented in the non-native accent—matching that of the listener. In the typing task, we also found higher accuracy in the non-native accent condition, particularly in the case of perfect cognates. Furthermore, when similarity was high or low for both orthography and phonology, accuracy was higher than when they were crossed.

The disproportionate effects we found of perfect cognates on both response time (LDT) and word recognition (accuracy in the typing task) align with prior studies that suggest that perfect cognates (or identical cognates) have a "special status" for bilinguals (Dijkstra et al., 1999; Lemhöfer and Dijkstra, 2004). This suggests that sharing the same orthographic item between two languages creates confusion and difficulties in the auditory modality. It is possible that this effect extends to lexical representation in general (i.e., not just orthographic, but also phonological representations), but to establish the extent of this effect, we would need to test perfect phonological cognates. Unfortunately, this was not possible within our study due to the language combination we addressed, which favored a relative dissociation between phonological and orthographic similarity. Therefore, this specific group of words—namely, perfect phonological cognates—should also be tested, as well as languages that share more of their phonemes and phoneme to grapheme correspondences. Finally, we found that accent facilitated word recognition and detection, as has been suggested in other studies (Bent and Bradlow, 2003; Xie and Fowler, 2013; Wang and van Heuven, 2015). This has been referred to as the inter language speech intelligibility benefit (Bent and Bradlow, 2003; Xie and Fowler, 2013; Wang and van Heuven, 2015). Studies so far have mostly focused on overall intelligibility and sentence comprehension, but we were able to extend these results to the word level and show that this effect interacts with other variables, such as orthography.

With respect to our original hypotheses, we found a reduction of the inhibitory orthographic effect in response time with the non-native accent and increased accuracy with the non-native accent, with both results supporting our first hypothesis—namely, a discrepancy in the FL phonological representation. In support of our second hypothesis, accuracy in both the LDT and the typing task showed the same inhibitory effects of orthography for both accent conditions. Overall, we can say our second hypothesis was more strongly supported, but there is evidence that the alignment between the phonological representation of the FL item and the specific auditory stimulus does play a role in auditory processing. In practical terms, hearing words in a non-native accent matching that of the listener seems to help both identify and spell the word more accurately, but the differences in phoneme to grapheme correspondences between the NL and FL make word identification and spelling more difficult. This would also mean that, when hearing words in an FL, participants attempt to "transcribe" these items and orthographic similarity between one's NL and FL becomes confusing.

One important limitation of our study is that we cannot speak to the possible effects of stronger or weaker accents or different regional accents, since we compared only two specific voices. Nevertheless, to our knowledge, there are no prior studies looking at the effects of accent in cognate processing—particularly perception—and the only prior study looking at the auditory effects of cognates is Frances et al. (2021). Furthermore, we were able to observe orthographic and phonological similarity effects separately—which also had not been done before—and tease apart the effects of each. Even though we focused on one pair of languages—namely, English and Spanish—our study highlights that orthographic and phonological similarity do not necessarily have the same effects. Indirectly, this points to the importance of the relationship between the languages of a bilingual when studying the interaction of visual and auditory representations—i.e., orthography and phonology—in processing. In other words, in languages with contradicting phoneme to grapheme conversion rules these factors are likely to have different effects on processing than in language combinations that have different writing systems (e.g., Greek and English or, even more so, Mandarin and English) or very similar phoneme to grapheme conversion rules (e.g., Spanish and Basque or Spanish and Italian). This, as well as the distinction between orthographic and phonological effects, is not contemplated so far in bilingual language processing models (Dijkstra and van Heuven, 2002; Brysbaert and Duyck, 2010). Our results suggest that it is necessary to integrate orthography and phonology as well as the relationship between languages (i.e., similarities and differences between them at various levels) into our current models of bilingual language processing. In the case of the auditory modality, it is also important to integrate "external" or "environmental" factors, so to speak, such as accent (as shown in the present study) or noise (see (Guediche et al., 2021; Navarra-Barindelli et al., 2021) showing a reduction of the cognate effect in noise) that are unique to each specific instance of the auditory input. More specifically in reference to our work, it is important to not only think of the phonology or phonological representations but also of the particular phonetics of the auditory stimulus.

Although this does not diminish the relevance of the effects we found, it is important to test other language combinations in order to fully understand the interactions between the languages of a bilingual. As a whole, our results suggest that the inhibitory orthographic similarity effect in auditory word perception in bilinguals is at least partially due to the relationship between the languages—their orthographies and opacity or transparency—as well as to whether the item is produced more differently or more similarly to the listener's own accent. Our results also call for more complex models of language processing that take into account different modalities and the relationship between the languages of a bilingual. In other words, we cannot expect the same effects when reading or listening and we cannot expect the same effects in Mandarin/English bilinguals as in Spanish/English or Spanish/Basque bilinguals. Future studies should focus on expanding these results to other sets of languages in order to assess the role of the relationship between languages in the effects of orthographic and phonological similarity.

CONCLUSIONS

In the current study, we found that both the accent in which an item is produced and the phoneme to grapheme conversion rules of the FL modulate the effect of orthographic similarity on auditory word processing. In general, detecting whether a phonological string is a word (LDT) was not affected by accent, but spelling out the correct word (typing task) was. Orthographic similarity had a negative effect in both cases and phonological similarity improved accuracy in the typing task, but not the LDT. Although further studies are needed in order to fully elucidate the origin of the inhibitory effect of orthographic similarity in the auditory modality, our results have both theoretical—pointing toward the need to take different modalities and language combinations into account for bilingual language processing models—as well as practical implications—for example, for foreign language learning.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/pr3es/?view_only=affcb60dc99b48868df05bcf68f2e671.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Basque Center on Cognition, Brain and Language Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CF, EN-B, and CM all conceived of the idea. CF and EN-B designed the experiment. EN-B carried out the experiment. CF programmed the experiment, wrote the manuscript, and did the analyses. CM supervised the project. The final manuscript was reviewed and commented by all authors.

FUNDING

This research was supported by the Basque Government through the BERC 2022-2025 program and by the Spanish State Research Agency through BCBL Severo Ochoa excellence accreditation CEX2020-001010-S. CF and EN-B are supported by MINECO predoctoral grants from the Spanish government (BES-2016-077169) and (BES-2016-078896) respectively. CM was further supported by the Spanish Ministry of Economy and Competitiveness [PID2020-113926GB-I00, PSI2017-82941-P, and RED2018-102615-T] and the Basque Government [PIBA18-29] and funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No: 819093 to CM).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.892822/full#supplementary-material

REFERENCES

Amengual, M. (2012). Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism* 15, 517–530. doi: 10.1017/S1366728911000460

Audacity Team (2018). *Audacity(R): Free Audio Editor and Recorder. Version 2.2.2*. Available online at: https://audacityteam.org/

Bent, T., and Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *J. Acoust. Soc. Am*. 114, 1600–1610. doi: 10.1121/1.1603234

Blumenfeld, H. K., and Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *J. Cogn. Psychol*. 25:e812093. doi: 10.1080/20445911.2013.812093

Brysbaert, M., and Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism* 13, 359–371. doi: 10.1017/S1366728909990344

Caramazza, A., and Brones, I. (1979). Lexical access in bilinguals. *Bull. Psychon. Soc*. 13, 212–214. doi: 10.3758/BF03335062

Carnegie Mellon (2020). *The CMU Pronunciation Dictionary*.

Colom,é À. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *J. Mem. Lang*. 45:2793. doi: 10.1006/jmla.2001.2793

Colomé, À., and Miozzo, M. (2010). Which words are activated during bilingual word production? *J. Exp. Psychol. Learn. Mem. Cogn*. 36:a0017677. doi: 10.1037/a0017677

Costa, A., Caramazza, A., and Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *J. Exp. Psychol. Learn. Mem. Cogn*. 26, 1283–1296. doi: 10.1037/0278-7393.26.5.1283

Cristoffanini, P., Kirsner, K., and Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *Q. J. Exp. Psychol. Sect. A* 38, 367–393. doi: 10.1080/14640748608401604

de Bruin, A., Carreiras, M., and Duñabeitia, J. A. (2017). The BEST dataset of language proficiency. *Front Psychol*. 8, e522. doi: 10.3389/fpsyg.2017.00522

de Groot, A. M. B., and Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *J. Mem. Lang*. 30, 90–123. doi: 10.1016/0749-596X(91)90012-9

Dijkstra, T., Grainger, J., and Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *J. Mem. Lang*. 41, 496–518. doi: 10.1006/jmla.1999.2654

Dijkstra, T., and van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Biling Lang Cogn*. 5, 175–197. doi: 10.1017/S1366728902003012

Dijkstra, T., Van Jaarsveld, H., and Ten, B. S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Biling. Lang. Cogn*. 1, 51–66. doi: 10.1017/S1366728998000121

Downey, S. S., Sun, G., and Norquest, P. (2017). alineR: An R package for optimizing feature-weighted alignments and linguistic distances. *R J*. 9, 138–152. doi: 10.32614/RJ-2017-005

Duyck, W., Van Assche, E., Drieghe, D., and Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *J. Exp. Psychol. Learn. Mem. Cogn*. 33, 663–79. doi: 10.1037/0278-7393.33.4.663

Frances, C., Navarra-Barindelli, E., and Martin, C. D. (2021). Inhibitory and facilitatory effects of phonological and orthographic similarity on L2 word recognition across modalities in bilinguals. *Sci. Rep*. 11:e9225. doi: 10.1038/s41598-021-92259-z

Goldrick, M., Runnqvist, E., and Costa, A. (2014). Language switching makes pronunciation less nativelike. *Psychol. Sci*. 25, 1031–1036. doi: 10.1177/0956797613520014

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain Lang*. 36, 3–15. doi: 10.1016/0093-934X(89)90048-5

Grosjean, F. (1997). The bilingual individual. *Interpret. Int. J. Res. Pract. Interpret*. 2, 163–187. doi: 10.1075/intp.2.1-2.07gro

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Biling Lang. Cogn*. 1, 131–149. doi: 10.1017/S136672899800025X

Grosjean, F. (2001). *Processing mixed language: Issues, findings, and models*, in The Bilingualism Reader, ed L. Wei (New York, NY: Routledge), 408–434.

Grosjean, F., and Nicol, J. (2007). "The bilingual's language modes," in *The Bilingualism Reader*, 2nd Edn, ed L. Wei (New York, NY: Routledge), 428–449. doi: 10.4324/9781003060406-39

Guediche, S., Navarra-Barindelli, E., and Martin, C. D. (2021). *Noise Modulates Crosslinguistic Effects on L2 Auditory Word Recognition*.

Janke, V., and Kolokonte, M. (2015). False cognates: The effect of mismatch in morphological complexity on a backward lexical translation task. Second *Lang Res*. 31, 137–156. doi: 10.1177/0267658314545836

JASP Team (2020). *JASP (Version 0.13.1)[Computer software]*. Available online at: https://jasp-stats.org/ (accessed February 01, 2022).

Kondrak, G. (1999). *Alignment of Phonetic Sequences*. Available online at: http://www.cs.utoronto.ca/csri/reports.html (accessed May 7, 2020)

Kondrak, G. (2000). A New Algorithm for the Alignment of Phonetic Sequences. *Proc First Meet North Am Chapter Assoc Comput Linguist*. 2000, 288–295.

Labov, W., Ash, S., and Boberg, C. (2006). *The Atlas of North American English: Phonetics, Phonology and Sound Change*. Berlin: Walter de Gruyter.

Lagrou, E., Hartsuiker, R. J., and Duyck, W. (2011). Knowledge of a second language influences auditory word recognition in the native language. *J. Exp. Psychol. Learn. Mem. Cogn*. 37:e023217. doi: 10.1037/a0023217

Lange, K., Kühn, S., and Filevich, E. (2015). "Just another tool for online studies" (JATOS): An easy solution for setup and management of web servers supporting online studies. *PLoS ONE* 10, e130834. doi: 10.1371/journal.pone.0130834

Lemhöfer, K., and Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behav. Res. Methods*. 44, 325–343. doi: 10.3758/s13428-011-0146-0

Lemhöfer, K., and Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Mem Cogn*. 32, 533–550. doi: 10.3758/BF03195845

Makowski, D. (2018). The Psycho package: An efficient and publishing-oriented workflow for psychological science. *J. Open Source Softw*. 3:470. doi: 10.21105/joss.00470

Marian, V., Bartolotti, J., Chabal, S., and Shook, A. (2012). Clearpond: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLoS ONE* 7, e43230. doi: 10.1371/journal.pone.0043230

Mathôt, S., Schreij, D., and Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behav. Res. Methods*. 44, 314–324. doi: 10.3758/s13428-011-0168-7

Navarra-Barindelli, E., Martin, C. D., and Guediche, S. (2021). *Visually Challenging Input Impacts L1 Effects on L2 Visual Word Recognition*.

Poort, E. D., and Rodd, J. M. (2017). The cognate facilitation effect in bilingual lexical decision is influenced by stimulus list composition. *Acta Psychol*. 180, 52–63. doi: 10.1016/j.actpsy.2017.08.008

Sadat, J., Martin, C. D., Magnuson, J. S., Alario, F. X., and Costa, A. (2016). Breaking down the bilingual cost in speech production. *Cogn. Sci*. 40:e12315. doi: 10.1111/cogs.12315

Sanchez-Casas, R. M., Davis, C. W., and Garcia-Albea, J. E. (1992). Bilingual lexical processing: Exploring the cognate/non-cognate distinction. *Eur. J. Cogn. Psychol*. 4, 293–310. doi: 10.1080/09541449208406189

Schmitz LPT/C (2019). *LimeSurvey: An Open Source Survey Tool*. Hamburg: LimeSurvey Project.

Schwartz, A. I., Kroll, J. F., and Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Lang. Cogn. Process*. 22, 106–129. doi: 10.1080/01690960500463920

Seidenberg, M. S., and Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *J. Exp. Psychol. Hum. Learn. Mem*. 5, 546–554. doi: 10.1037/0278-7393.5.6.546

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., and Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *J. Mem. Lang*. 64:6. doi: 10.1016/j.jml.2010.08.006

Van Assche, E., Duyck, W., and Hartsuiker, R. J. (2012). Bilingual word recognition in a sentence context. *Front. Psychol*. 3, 1–8. doi: 10.3389/fpsyg.2012.00174

van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Mem. Cognit*. 15:e97716. doi: 10.3758/BF03197716

Voga, M., and Grainger, J. (2007). Cognate status and cross-script translation priming. *Mem. Cognit*. 35, 938–952. doi: 10.3758/BF03193467

Wang, H., and van Heuven, V. J. (2015). The interlanguage speech intelligibility benefit as bias toward native-language phonology. *Iperception* 6, 1–13. doi: 10.1177/2041669515613661

Xie, X., and Fowler, C. A. (2013). Listening with a foreign-accent: The interlanguage speech intelligibility benefit in Mandarin speakers of English. *J. Phon*. 41, 1–24. doi: 10.1016/j.wocn.2013.06.003

Zhang, J., and Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika* 70, 203–212. doi: 10.1007/s11336-003-1119-8

Ziegler, J. C., and Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychon. Bull. Rev*. 5, 683–689. doi: 10.3758/BF03208845

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 19 May 2022
doi: 10.3389/fpsyg.2022.870152

Mapping Pitch Accents to Memory Representations in Spoken Discourse Among Chinese Learners of English: Effects of L2 Proficiency and Working Memory

Connie Qun Guan[1,2], Wanjin Meng[3], Laura M. Morett[4] and Scott H. Fraundorf[5]*

[1] School of Foreign Studies, Beijing Language and Culture University, Beijing, China, [2] Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, United States, [3] China National Institute of Education Sciences, Beijing, China, [4] Department of Educational Studies in Psychology, Research Methodology, and Counseling, University of Alabama, Tuscaloosa, AL, United States, [5] Department of Psychology and Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, United States

OPEN ACCESS

Edited by:
Yan Jing Wu,
Ningbo University, China

Reviewed by:
Yao Wang,
Beijing Normal University, China
Erik D. Thiessen,
Carnegie Mellon University, United States

***Correspondence:**
Wanjin Meng
1205747017@qq.com

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *06 February 2022*
Accepted: *09 March 2022*
Published: *19 May 2022*

Citation:
Guan CQ, Meng W, Morett LM and Fraundorf SH (2022) Mapping Pitch Accents to Memory Representations in Spoken Discourse Among Chinese Learners of English: Effects of L2 Proficiency and Working Memory. Front. Psychol. 13:870152. doi: 10.3389/fpsyg.2022.870152

We examined L2 learners' interpretation of pitch accent cues in discourse memory and how these effects vary with proficiency and working memory (WM). One hundred sixty-eight L1-Chinese participants learning L2-English listened to recorded discourses containing pairs of contrastive alternatives and then took a later recognition memory test. Their language proficiency and WM were measured through standard tests and the participants were categorized into low, medium, advanced, and high advanced language proficiency groups. We analyzed recognition memory task performance using signal detection theory to tease apart response bias (an overall tendency to affirm memory probes) from sensitivity (the ability to discern whether a specific probe statement is true). The results showed a benefit of contrastive L + H* pitch accents in rejecting probes referring to items unmentioned in a discourse, but not contrastive alternatives themselves. More proficient participants also showed more accurate memory for the discourses overall, as well as a reduced overall bias to affirm the presented statements as *true*. Meanwhile, that the benefit of L + H* accents in rejecting either contrast probes or unmentioned probes was modulated for people with greater working memory. Participants with higher WM were quite sure that it did not exist in the memory trace as this part of discourse wasn't mentioned. The results support a contrast-uncertainty hypothesis, in which comprehenders recall the contrast set but fail to distinguish which is the correct item. Further, these effects were influenced by proficiency and by working memory, suggesting they reflect incomplete mapping between pitch accent and discourse representation.

Keywords: L2 processing, pitch accent, discourse, memory, working memory

INTRODUCTION

A general challenge for the second language (L2) learners is learning to associate linguistic forms with meaning at different levels of linguistic representation (Perfetti, 1997). Thus, a central scientific issue in the study of L2 learning is whether and to what extent learners can exploit these associations in their L2 well enough to achieve native-like performance in learning and memory. More generally, there is a debate as to whether L2 processing is qualitatively different from L1 processing [e.g., Marinis et al. (2005); Papadopoulou (2005), Clahsen and Felser (2006), and Felser and Roberts (2007)] or if native-like performance can be obtained with increased L2 proficiency or exposure [e.g., French-Mestre (2002), Hopp (2006, 2010), Jackson (2008), and Jackson and Dussias (2009)].

The goal of the current study aims to examine the individual differences effects on the prosodic memory trace when the discourse contains high vs. low PA contrast vs. non-mentioned alternatives. Herein, we examine these questions in the domain of speech prosody, an important cue to sentence and discourse processing (Cutler et al., 1997; Wagner and Watson, 2010). L2 listeners adopt different prosodic processing strategies or mechanisms than L1 listeners (Pennington and Ellis, 2000; Akker and Cutler, 2003; Baker, 2010; Braun and Tagliapietra, 2011). When prosodic cues match the content of speech, they enhance L2 speakers' discourse comprehension and memory (Lee and Fraundorf, 2019, 2021). For L1 speakers, when prosodic cues fail to match the content of speech, they interfere with discourse content and memory (Harrington, 1992; van den Noort et al., 2006; Morett and Fraundorf, 2019; Morett et al., 2020, 2021). At present, it is unclear whether this is the case for L2 speakers, however.

In the present study, we examine how differences in prosodic cues affect how L1 Chinese learners of L2 English map prosodic pitch accents (PAs) to discourse status and how it affects comprehension and memory for spoken discourse. We also consider the contributions of language proficiency and working memory (WM) to the impact of PA on discourse comprehension and memory.

We focus on L2 learners' acquisition of the mapping of different PAs to representation of spoken discourse as we all know that, learning to comprehend PAs in spoken discourse is a practical issue for L1-Chinese learners of L2-English. Nevertheless, the mapping between PAs and semantic integration in discourse has long been neglected in L2 teaching and learning practices (Gut and Milde, 2002; Gut, 2003; Gut et al., 2007; Braun and Tagliapietra, 2011). Therefore, learners with limited L2 proficiency are likely to fail to make use of PA in L2 listening comprehension, leading to poor comprehension and memory. This void in both research and practice requires further scientific investigation on whether and to what extent L2 learners can map PAs to discourse representations in learning and memory.

How Do Pitch Accents Affect Memory in L2 Processing?

Previous work (Lee and Fraundorf, 2017) found that L1 Korean learners of L2 English showed substantial differences from L1 English monolinguals in how they used PA information to encode and remember a discourse. In the present study, we aimed to investigate whether similar L1-L2 differences exist among L1 Chinese learners of L2 English. Below, we review three hypotheses about how contrastive PAs might affect language comprehension and memory: granularity, contrast representation, and contrastive uncertainty. First, the *granularity* account claims that the effect of a contrastive PA, and perhaps focusing devices more generally, is to enhance the representation of the accented word itself. This account was originally proposed to describe the mnemonic benefit of other focus-marking devices, including *it*-cleft constructions (Just and Carpenter, 1987, 1992; Birch and Garnsey, 1995; Sturt et al., 2004) and font emphasis in a written discourse (Sanford et al., 2006), but could also describe effects of PAs (Norberg and Fraundorf, 2021). The granularity account predicts a contrastive PA should enhance a listener's ability to reject all false alternatives because all of those are inconsistent with the correct information.

Alternatively, the *contrast representation* account proposes that the mnemonic benefit of contrastive PAs over presentational PA is that contrastive PAs enhance the representation of specific salient alternatives in the discourse. For example, in discourse (1, 2) above, the *British* scientists are contrasted with the *French* scientists, so a contrastive PA may lead listeners to retain something about the contrastive *French* scientists in particular. This account is supported among L1 listeners by Fraundorf et al. (2010), who found that a contrastive PA facilitated comprehenders' later ability to correctly reject the salient alternative, but not of wholly unmentioned items never part of the contrast set.

Thirdly, the *contrastive uncertainty*[1] account refers to the possibility that comprehenders encountering a contrastive PA may bring to mind the set of contrasting alternatives but fail to encode which is the correct statement and which is the alternative. After all, some degree of uncertainty about the linguistic input is a fundamental characteristic of language processing [e.g., Goodman and Lassiter (2015)]; comprehenders may remember that there was a contrast between two alternatives (such as *British* and *French*) but be uncertain about which member of the set was referred to later. In this case, a contrastive PA may thus confer no benefit on, or even counterintuitively *impair*, the ability to rule out the salient alternative. By comparison, a strong memory representation of the two mentioned alternatives might benefit comprehenders' ability to reject the unmentioned-item probe. That is, the correct and the contrastive alternatives are easy to tease apart among L1 listeners, but not among L2 speakers because both pieces of information exist in L2 listeners' memory with less degree of certainty about which was originally stated. This is what has been found with L1-Korean speakers learning L2-English (Lee and Fraundorf, 2017, 2021), suggesting that L2 learners do not process PA cues the same way as native speakers.

However, given the influence of L1 on L2 processing (Rasier and Hiligsmann, 2009; Mennen, 2015), it is unclear whether this pattern is unique to L1-Korean speakers or whether it characterizes L2 prosodic processing more broadly. Here, we sought to determine whether a similar pattern of L2 prosodic processing emerged among L1 Chinese speakers learning L2 English.

[1] Lee and Fraundorf (2017) called this a shallow representation; here, we introduce the term *contrastive uncertainty* to better distinguish this account of the effect of PAs from the (largely unrelated) Shallow Structure Hypothesis account of L2 syntactic processing.

Pitch Accents as Cues to Discourse Processing

PA represents a point of both commonality and difference between Chinese and English. On the one hand, both Chinese and English have tones anchored on stressed syllables, i.e., lexical tones in Chinese and PAs in English (Duanmu, 2004). On the other, the function of these tones differs drastically. English has no lexical tones, and so its words can take different PAs to express contextual meanings, such as marking a focused entity. By contrast, Chinese tones are lexically contrastive. Instead, focus on Chinese is conveyed by providing lexical cues or by putting stress at the end of the syllable *via* pausing, lengthening, rising, or falling tones (Ouyang and Kaiser, 2012; Yang and Chen, 2014, 2018; Lee et al., 2015, 2016; Wang et al., 2020). Hence, understanding the focus information conveyed by PAs in English could be a great challenge for Chinese learners of L2 English.

Pitch accents are phonological constructs that are placed on particular words and are usually realized acoustically with longer duration, greater intensity, and greater pitch excursion than unaccented words [for review, Ladd (2008)]. Most theories of prosody distinguish multiple types of PAs. Here, we focus on the distinction between contrastive PAs and presentational PAs, denoted as L + H* and H*, respectively, in the ToBI system for intonational transcription of English (Silverman et al., 1992; Beckman and Elam, 1997). Specifically, the L + H* and the H* accents both indicate a salient tonal target (an H*) on stressed syllables, but they are distinct from each other in that, in the L + H* accent, the salient tonal target involves a steep rise from an initial low tone (L), similar to Mandarin Tone 2, whereas in the H* accent, the salient tonal target remains flat, similar to Mandarin Tone 1. The L + H* accent has been considered to correspond to information that specifically contrasts with some other information in the discourse, whereas the H* accent accompanies new information more generally (Pierrehumbert and Hirschberg, 1990).

Thus, understanding PAs can be crucial for interpreting a sentence or discourse at a semantic and pragmatic level. Indeed, PAs contribute to native listeners' initial online processing of spoken discourse. Experiments using the visual-world paradigm suggest that not only are PAs rapidly detected, but they can also be integrated into the discourse representation in the first moment of processing (Dahan et al., 2002; Ito and Speer, 2008; Watson et al., 2008).

Further, PAs contribute to long-term memory for a spoken discourse [e.g., Fraundorf et al. (2010, 2012), Gotzner et al. (2013), and Lee and Snedeker (2016)]. We highlight a particular study by Fraundorf et al. (2010) because it is most relevant to our present design. Fraundorf et al. (2010) examined memory for spoken discourses, such as (1, 2) below. A context passage (1) first established two contrasts, each between a pair of items (e.g., *British* vs. *French* and *Malaysia* vs. *Indonesia*). A continuation passage (2) then picked out one item from each contrasting set. The pitch accent on each of these critical words was manipulated between a presentational or contrastive accent through splicing.

(1) Both the British and the French biologists had been searching Malaysia and Indonesia for endangered monkeys.

(2) Finally, the (British/BRITISH) spotted one of the monkeys in (Malaysia/MALAYSIA) and planted a radio tag on it.

After listening to all the recorded stories, participants completed a recognition memory test for the referent chosen in each continuation. Participants were presented with probe statements, such as (3), and had to indicate whether each statement was true or false. The probe statements could refer either to the correct item (e.g., *British*, a true statement that should be affirmed), to the contrastive alternative from the original discourse (e.g., *French*, a false statement that should be rejected) or a wholly unmentioned item (e.g., *Portuguese*, a false statement that should be rejected).

(3) The (British/French/Portuguese) scientists spotted the endangered monkey and planted a radio tag on it.

Fraundorf et al. (2010) found that when the critical word was originally heard with a contrastive PA, memory was more accurate even a day later. Critically, this effect came about specifically because contrastive PAs facilitated rejection of the contrast item (e.g., *French* in example 3 above); contrastive PAs did not benefit rejections of an unmentioned item that was never part of the contrast set (such as *Portuguese*). Thus, Fraundorf et al. (2010) concluded that contrastive PAs lead comprehenders to encode and remember a salient alternative to the accented item, consistent with linguistic theories that posit the role of contrastive focus is to introduce a set of salient alternatives into the discourse [e.g., Rooth (1992)].

Processing Pitch Accents in L2

However, recent research suggests L1–L2 differences frequently emerge in PA processing even under circumstances otherwise favorable for L2 processing. For instance, Akker and Cutler (2003) found that L1 Dutch listeners of L2 English could only detect focused information, and not contrastive information, despite the similarity of the two languages in their prosodic systems. Similarly, even highly proficient non-native listeners could not distinguish idiomatic from non-idiomatic expressions in English when marked with prosodic cues (Vanlancker-Sidtis, 2003).

With respect to the above memory task probing the effect of prosody on long-term discourse representations, Lee and Fraundorf (2017, 2021) found dramatically different patterns among L1-Korean college students learning L2-English. Low- and mid- proficiency learners (as defined by scores on a cloze task) showed no memory benefit whatsoever from contrastive PAs. High-proficiency learners were sensitive to contrastive PAs, but even they processed them in a non-native-like way: The contrastive PAs did not help them rule out a contrastive alternative and in fact, led participants to falsely affirm the contrastive alternative more often. Rather, contrastive PAs helped the high-proficiency L2 learners to reject items completely unmentioned in the discourse. This pattern suggests that the high-proficiency L2 learners did encode a set of contrasting alternatives in response to the contrastive pitch accent (which would help them reject the unmentioned alternative), but they failed to correctly encode *which* alternative was the correct one.

Considering the similarities in tonal features in Chinese and Korean L1, these differences are likely to be even starker for L1 Chinese speakers learning L2 English because of the dissimilarity of their prosodic systems. Chinese is a tonal language centered on each character and each character matches with the correct tonal information that represents a meaning unit. Different meaning units plus pitch accents will construct a sentence structure. Chinese tones are fixed to each Chinese character, which suggests that sentences with different tones can be spoken accurately as long as each Chinese character and its tone is pronounced, and tones are correct. The same syllable, when it is spoken with four different tones, can have four distinctive meanings (Wang et al., 2008). For example, /shu/1 uncle, /shu/2 ripe, /shu/3 summer, /shu/number. There are two basic types of Chinese tone trends in general: falling tone and rising tone. Chinese tones and intonation work together, but tones do not change with the intonation. Korean is an alphabetic language (Kim et al., 2016), which is similar to English. Each of its vowels and consonants has no tone. Standard Korean also does not have a system of using stress to distinguish the meaning of words. The most basic tone in Korean is like a curve that bulges out in the middle. That is, it starts with a low tone that goes up and then goes down. This is the same as in English. Korean tone is characterized by the number of syllables, which is about 2,000 (Taylor and Taylor, 2014). This is the same as Chinese, but there is a difference in the number of syllables, and Korean has a richer phonological system than Chinese. Chinese has about 400 syllables, which is much less than Korean (Taylor and Taylor, 2014); the first two syllables and the last two syllables have low-high intonation, and when the first one is difficult to pronounce, "low-high-low-high" becomes "high-high-low-high." The intonation of the phrases in the sentence changes according to the intention of the speech, which is also similar to Chinese.

In general, the L1–L2 relationship seems to affect the acquisition of L2 prosody, which in some cases can make L2 prosody more difficult to process. For example, Rasier and Hiligsmann (2009) found a relationship between the typological distance between the learner's L1 and L2 (markedness relationships) and the occurrence of transfer in their use of (de)-accentuation. Specifically, only marked L1 patterns were transferred from L1 to L2, suggesting that markedness is an important factor in L2 prosodic learning and transfer. Additionally, transfer from L1 is particularly persistent in prosody and can explain L2 learners' difficulties adopting a language-appropriate pitch range [e.g., Curtis and D'Esposito (2002); Mennen (2004), and Scharff-Rethfeldt et al. (2008)]. Research suggests that non-target-like prosody in a L2 plays an important and independent role in the perception of foreign accentedness and native-listener judgments of comprehensibility (Magen, 1998; Jilka, 2000; Trofimovich and Baker, 2006). Mennen (2015) argued that the relative difficulty of L2 prosody is influenced by L1 (Willems, 1982; Jilka, 2000; Grabe, 2004; Mennen et al., 2010) and is, to some extent, predictable from universal markedness (Rasier and Hiligsmann, 2007; Zerbian, 2015) and from universal developmental paths in L2 prosodic acquisition in which some segmental learning must occur before learning intonational characteristics, such as stress, rhythm, tone, tempo pauses, loudness and voice quality (Nolan, 2006; Li and Post, 2014; Mennen and De Leeuw, 2014).

The influence of L1 on the acquisition of L2 prosody is supported by findings that stress distinctions are difficult for speakers of non-stress languages to process and—especially—retain in memory (Beckman, 1986; Dupoux et al., 1997; Peperkamp and Dupoux, 2002), as are tone distinctions (Shen, 1989). Listeners whose L1 is a non-stressed, tonal language, such as Chinese, are accustomed to syllabic non-stressed tonal information. Therefore, L1 Chinese listeners of L2 English often do not show sensitivity to English prosodic distinctions, neither online nor in memory retrieval after listening. For instance, Pennington and Ellis (2000) found that L1 Cantonese learners of L2 English had a poor memory for prosodically signaled information, including focused contrasts.

Why L2 Differences?

To the extent that L2 learners do not show native-like prosodic processing (e.g., exhibiting a contrastive-uncertainty pattern rather than a contrast-representation effect), a second question is *why* these differences exist. This issue relates to more general questions about constraints on L2 learning. Here, we consider two potentially relevant constructs: proficiency and working memory.

One possibility is that non-native-like prosodic processing of L2 reflects a lack of knowledge of the L2. The mapping between prosodic cues and semantic and discourse information varies across languages. Thus, non-native listeners of any given language may initially have little knowledge of intonational form and meaning, especially since they are less frequently taught in formal L2 instruction, but may gain them with experience. Furthermore, PA processing in an atonal L2 may implicitly activate tonal language speakers' representations of lexical tones, which may result in interference given that PA and lexical tone serve different linguistic functions. Under this account, as L2 knowledge and proficiency increases, interference from L1 lexical tone may decrease and L2 learners might be expected to become more native-like in their prosodic processing. Supporting this, Lee and Fraundorf (2017) found that low- and moderate-proficiency learners of L2 English showed no effects of contrastive PAs whatsoever, whereas high-proficiency learners did. Nevertheless, even high-proficiency learners did not process PAs the same way that native speakers did, suggesting that proficiency may not be the only relevant factor.

An alternative possibility is that L2 prosodic processing is constrained by more general cognitive resources, such as WM. There are at least two reasons to hypothesize a role for WM in L2 PA processing. First, learners of L2 English may have to rely more on declarative knowledge. Ullman (2001, 2004), Pinker (1999), and Pinker and Ullman (2002) have hypothesized that both declarative and procedural memory contributes to native speakers' language processing. Declarative memory refers to verbalizable knowledge, for instance, in the domain of language, the association of vocabulary items with their respective meanings. In contrast, procedural memory is used to learn and control skills and habits that are not recognized explicitly. In the domain of language, native speakers process structural information (e.g., prosodic structure in the current study) based on procedural memory. But for L2 learners, the declarative-procedural (DP) model claims that their processing of linguistic knowledge might rely more on declarative memory. Thus, L2 learners may have access, and use declarative knowledge about pitch accents, which relies on working memory.

Second, greater WM capacity may help with the general processing demands of L2 comprehension. Difficulty with phonetic distinctions, lower vocabulary size, lesser accumulated lexical familiarity, and unfamiliarity with idiomatic expressions all combine to make non-native comprehension of spoken language less efficient than comprehension by native listeners (Akker and Cutler, 2003). As a result, it can be difficult enough to keep up with lexical and syntactic processing in an L2 language, leaving L2 comprehenders with insufficient processing resources for discourse or prosodic processing. Whereas native listeners can adopt a top-down mechanism in which their prosodic processing assists phonetic identification and lexical access, non-native listeners mainly employ a bottom-up mechanism in that they focus on the lexical and phonetic levels of information before applying prosodic cues (Akker and Cutler, 2003). Lee and Fraundorf (2017) suggested this could create the contrastive-uncertainty effect if L2 comprehenders do not have sufficient processing resources (i.e., WM) to encode which member of the contrast set is the true proposition and which is the contrastive alternative. Supporting this, individual differences in WM constrain PA interpretation even in L1 (Fraundorf et al., 2010).

Lastly, another possibility is that L1–L2 differences in prosodic processing are intrinsic to listening in a second language. This possibility accords with a long theoretical tradition proposing that L2 processing is qualitatively different from L1 processing even with extensive experience (Beckman, 1986; Cutler et al., 1997). In this account, neither greater working memory nor increased proficiency may be sufficient to modulate L2 PA processing.

Research Questions and Hypotheses

In the present study, we tested how PAs influenced L1 Chinese learners' memory for L2 English spoken discourse. We had two primary research questions:

First, we considered whether and how L2 learners' ability to use PAs for encoding relevant contrasts in a discourse differs from that of native speakers. We contrasted the predictions of three hypotheses. The granularity hypothesis predicts that a contrastive PA should help rule out any false information. The contrast representation hypothesis predicts that a contrastive PA should specifically facilitate rejections of the false alternative. Finally, the contrastive uncertainty hypothesis predicts that a contrastive PA leads comprehenders to affirm both the correct item and (erroneously) the contrast item, but it does help them rule out the unmentioned item that is entirely outside the contrast set.

Second, we considered how prosodic processing varies with L2 proficiency and/or WM capacity. The theoretical rationale for exploring the cognitive factors in bilingual processing is crucial. We hope to reveal whether and to what extent the bilinguals with high proficiency or high cognitive control abilities (represented by working memory) could use PAs for encoding relevant contrasts in a discourse differs from that of native speakers. Although we already know that the native L1 speakers could clearly put information about PA in their text memory, we still do not know how text memory would be represented when there is uncertain information in the discourse. Therefore, we would predict again that a contrastive PA might leads bilinguals with more advanced cognitive abilities to affirm both the correct item to a certain extent but not in the confirmed L2 text memory and might cause them to erroneously remember the contrast item, and this effect might help them rule out the unmentioned item that is entirely outside the contrast set.

MATERIALS AND METHODS

Participants

We recruited forty-two participants from each of four different subject populations of native Chinese speakers whom we expected to vary in English proficiency: high school students (Mean Age = 17.1, SD = 0.89), university undergraduates not majoring in English (Mean Age = 18.3, SD = 0.92), undergraduates majoring in English (Mean Age = 18.2, SD = 0.79), and graduate students of English (Mean Age = 21.4, SD = 0.67). The students were recruited from the similar educational backgrounds in the same neighborhood. Although there might be differences between the high school students and the university students in Chinese fluency, the differences were not that salient between these groups ($p > 1$). Their language ability differed only in English L2 fluency, rather than something more like maturation. All told, one hundred sixty-eight native speakers of Chinese (35 males) participated in the study. They were all recruited from the Beijing University of Science and Technology, from which the ethic committee approved the study in 2016 and the study was conducted in 2018 while the first author was sponsored by the Sino-US-Fulbright Scholarship. All resided in China at the time of participation.

To verify the differences in proficiency, all participants completed (a) a demographic survey on their language background reporting their years of formal English education and language proficiency and (b) the Quick Placement Test (Quick Placement Test [QPT], 2001), a test of English language proficiency (Geranpayeh, 2003). These two scores of years of formal English education and self-ratings of proficiency were highly correlated: $\eta^2 = 0.71$, $p < 0.05$, suggesting that we had obtained reliable measures of participants' English proficiency. Further, an ANOVA revealed that the proficiency level varied significantly across the four participant groups, $F_{(1,41)} = 7.794$, $p = 0.003$, $\eta^2 = 0.328$, confirming that we had successfully identified groups that differed in their L2 proficiency.

Table 1 presents the demographic information as well as the average scores on the working memory tasks (discussed below) for each group. The reliability coefficients of the two WM measures were 0.69 and 0.72 respectively.

Materials

The listening materials were the 36 audio-recorded discourses previously used in Experiment 3 of Fraundorf et al. (2010) and

TABLE 1 | Demographic information for the four participant groups.

Group	Age		Gender (M/F)	OSpan		RSpan		QPT		QPT proficiency percentile	Subject population	Years of English education
	M	SD		*M*	SD	*M*	SD	*M*	SD			
Low	16.73	1.2	23/19	13.01	3.6	10.34	3.9	28.45	3.25	<25th	High school students	8
Medium	19.58	2.2	4/38	11.44	3.7	9.08	3.3	37.65	3.84	25th- 50th	Non-English Major Undergraduates	10
Advanced	24.10	2.0	4/38	11.48	4.0	9.55	3.2	41.23	4.13	50th to 75th	English-major Undergraduates	11
High advanced	22.97	1.8	3/38	11.88	4.2	10.26	3.9	47.86	2.97	>75th	English postgraduates	14

OSpan, operational span; RSpan, read span; QPT, quick placement test.

Lee and Fraundorf (2017). Each discourse consisted of a context passage, such as (1a) below, introducing two contrastive sets of items (e.g., *British* and *French* as one set and *Malaysia* and *Indonesia* as the other), followed by a continuation passage that referred to one item in each contrast set, as in (1b). All these discourses were recorded by a native English speaker who was trained to produce the different pitch-accent types.

In a within-participants design, we varied whether the critical word in the continuation passage was produced with a presentational H* accent (indicated by regular text in the example) or a contrastive L + H* accent (indicated in capital letters in the example). The accent on each of the two critical nouns was orthogonally manipulated such that an L + H* accent could be placed on either the first critical word, second critical word, both, or neither. Audio recordings were created using cross-splicing such that only the critical words varied across conditions, and the rest of the recordings were identical.

(1a) Context Passage: Both the British and the French biologists had been searching Malaysia and Indonesia for endangered monkeys.
(1b) Continuation Passage: Finally, the (British/BRITISH) spotted one of the monkeys in (Malaysia/MALAYSIA) and planted a radio tag on it.

Memory for the spoken discourses was tested using probe statements presented *via* text, for which responses consisted of true/false. No prosodic cues were present during the test phase. Each critical word could be tested with one of three probe types: the correct fact, the contrastive alternative, or a wholly unmentioned item. For example, probes (2a) through (2c) test the critical word *British* in discourse (1), and probes (3a) through (3c) test the critical word *Malaysia*. Each critical word was tested in only one probe condition per subject.

(2a) The British scientists spotted the endangered monkey and tagged it.
(2b) The French scientists spotted the endangered monkey and tagged it.
(2c) The Portuguese scientists spotted the endangered monkey and tagged it.
(3a) The endangered monkey was finally spotted in Malaysia.
(3b) The endangered monkey was finally spotted in Indonesia.
(3c) The endangered monkey was finally spotted in the Philippines.

This resulted in a 4 × 2 × 3 factorial design: Proficiency Group (low, medium, advanced, and most advanced) × Pitch Accent Type (H* or L + H*) × Probe Type (correct, contrast, or unmentioned), with the first variable varying between subjects and the others within subjects. Assignment of items to conditions was counterbalanced across six presentation lists using a Latin Square design. The complete lists of stories and test probes are available in Fraundorf et al. (2010).

Procedure

Participants completed the demographic questionnaire and QPT followed by two working-memory tasks before proceeding to the discourse-memory task.

Operation Span

To assess the participants' working memory capacity (WMC), we implemented a version of the operation span (OSpan) task distributed by Redick and Engle (2006). The OSpan task has been shown to correlate with a wide range of higher-order cognitive tasks, such as reading and listening comprehension (Engle, 2010).

In the OSpan task, participants were presented with a series of simple equations one at a time. The left-hand side of each equation consisted of two operations, either (a) two additions or subtractions or (b) one addition/subtraction and one simple multiplication, such as 23–16 + 7 = 14? or 136 + 64 × 2 = 401? Participants had to mentally determine whether the number of the right-hand side of the equation correctly completed the equation and then responded by saying "yes" or "no" aloud. The experimenter pressed a key after the participant gave the response. Immediately afterward, a capital letter was displayed on the screen, which participants were tasked with remembering, and then the next equation appeared. The set of possible letters included only the *letters H*, *J*, *K*, *L*, *N*, *P*, *Q*, *R*, *S*, *T*, and *Y* because those letters were relatively phonologically distinct.

After a certain number of letters had been presented (the *span length*), participants were instructed to write down on paper the presented letters in the order in which they were presented. Two trials were presented at each of the span lengths from 2 to 9 sentence/letter pairs (14 trials total); the order of the trials was random. The task took approximately 15 min to complete.

Scoring was performed using *partial-credit unit scoring*, which in past work has been recommended for producing the most normal distribution of scores (Conway et al., 2005). If participants remembered all the letters in a particular trial, that was scored as 1. If participants remembered some but not all of the letters, they received a credit equal to the proportion of letters they did recall; for example, remembering two letters on a trial of span length six would be scored as 0.33. The maximum score for the task was 14.

Reading Span

We also implemented a modified version of the reading span (RSpan) task in Chinese (Daneman and Carpenter, 1980) modeled after the OSpan task. The reading span task was similar to the OSpan except that, instead of verifying equations, participants read aloud sentences in Chinese (e.g., [translated English version] *On warm sunny afternoons, I like to walk in the park.*?) and verified whether the sentence made sense by saying "yes" (makes sense) or "no" (does not make sense) immediately after they finished reading the sentence. As in the OSpan task, the to-be-remembered stimuli were capital letters presented between the sentences. We use a Chinese version of this task because we wanted this task to capture differences in working memory *per se* rather than L2 English proficiency, which we measured separately.

As in the OSpan task, the span length varied from 2 to 9 sentence/letter pairs, with two trials at each span length; the order of the same lengths was random. At the end of the series, participants wrote down the sequence of capitalized letters in the same order they read in the test. The RSpan task was scored the same way as the OSpan task, with a maximum score of 14.

Discourse Recognition Memory Task

Pilot testing indicated that remembering the discourses was relatively difficult for our L2 English participants. To reduce the memory burden and maximize the chance that performance was not at floor level, we split the materials into six blocks; each contained four recorded discourses followed by eight critical probe questions (two for each discourse) and two filler probes. The filler probes tested memory for other aspects of the stories unrelated to those emphasized by the PAs.

Each block began with the presentation of the auditory stimuli. During each story, the screen was black. There were a 5 s interstimulus interval between stories. After all of the stories in a block had been presented, the recognition memory task began. Each probe statement was displayed on the screen one at a time in different randomized orders from the original presentation, and participants indicated whether they judged the probe as *True* or *False* by pressing one of two keys on the keyboard. Participants were instructed to reject a probe as *False* if any part of it did not match the story they had heard. After participants made their response, there was a 1,000 ms interstimulus interval before the next probe was presented. Once all of the probes in a block had been presented, the procedure repeated with the study phase of the next block.

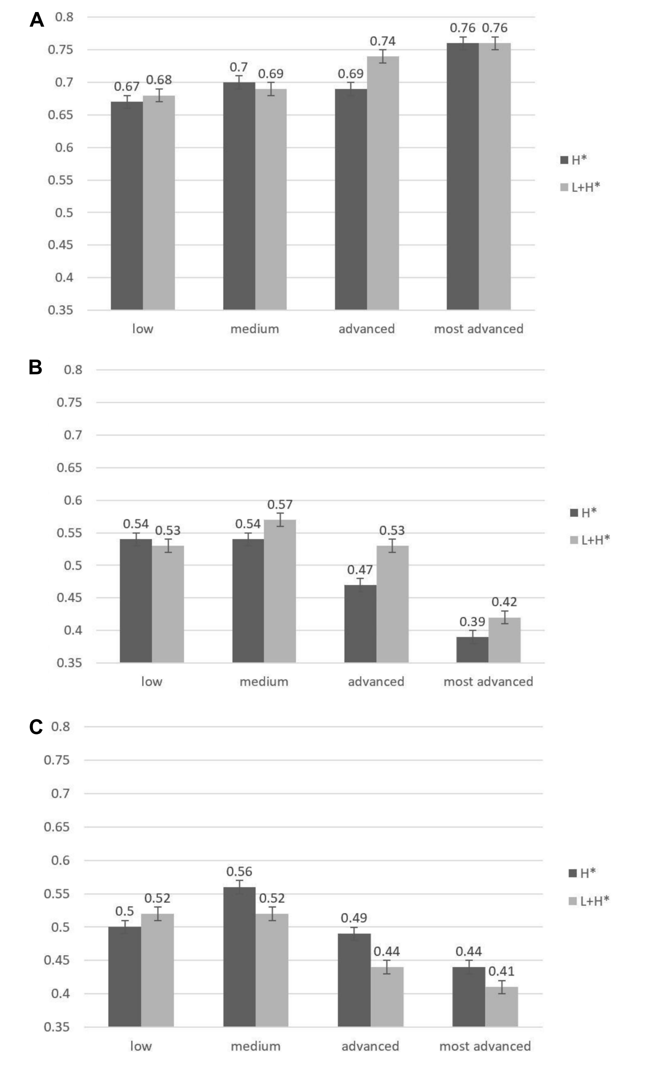


FIGURE 1 | Bar charts of proportions of *true* responses across groups and conditions. (Note that *true* is a correct response to correct probes, but an incorrect response to contrast and unmentioned probes). **(A)** Correct probes. **(B)** Contrast-probes. **(C)** Unmentioned probes. *stands for the abbreviation.

RESULTS

Analytic Strategy

In the recognition memory task, the accurate answer to the correct probes was *true* but *false* to the other probe conditions. Thus, comparing simple accuracy rates across conditions confounds the accuracy of participants' memory with any overall tendency to respond *true* or *false*. To eliminate this confound, we analyzed the memory data based on the theory of signal detection, in which the dependent measure is whether a participant judged a particular probe as *true* [Green and Swets, 1966; Macmillan and Creelman, 2005; for applications to mixed-effects modeling, Wright et al., 2009 and Murayama et al. (2014)]. This analysis allows a theoretical and empirical separation of *response bias* (an overall tendency to respond *true*) from *sensitivity* (the ability

TABLE 2 | Fixed effects estimates from mixed logit model of "*True*" responses with probe type, accent type and proficiency group as fixed effects.

	Estimate	SE	Wald *z*	*p*-value
Main effects across proficiency levels				
Baseline rate of *true* responses (response bias)	0.29	0.06	5.59	<0.001
Contrast probe vs. baseline (sensitivity)	−0.60	0.06	−10.81	<0.001
Unmentioned probe vs. baseline (sensitivity)	−0.70	0.06	−12.62	<0.001
L + H* accent (effect on response bias)	0.02	0.04	0.54	0.59
L + H* accent × contrast probe (effect on sensitivity)	0.17	0.11	1.50	0.13
L + H* accent × unmentioned probe (sensitivity)	−0.23	0.11	−2.05	0.04
Effects of proficiency				
Medium vs. low proficiency (response bias)	0.12	0.07	1.63	0.10
Advanced vs. low/medium proficiency (response bias)	−0.09	0.06	−1.47	0.14
Most advanced vs. low/medium/advanced proficiency (response bias)	−0.17	0.07	−2.46	0.01
Medium vs. low proficiency × contrast probe (sensitivity)	−0.01	0.15	−0.06	0.95
Medium vs. low proficiency × unmentioned (sensitivity)	−0.01	0.15	−0.04	0.96
Advanced vs. low/medium proficiency × contrast probe (sensitivity)	−0.19	0.13	−1.47	0.14
Advanced vs. low/medium proficiency × unmentioned (sensitivity)	−0.34	0.13	−2.64	0.01
Most advanced vs. low/medium/advanced proficiency × contrast probe (sensitivity)	−0.68	0.15	−4.66	<0.001
Most advanced vs. low/medium/advanced proficiency × unmentioned (sensitivity)	−0.34	0.15	−2.24	0.02
Effects of proficiency in comprehension of prosody				
Medium vs. low proficiency × L + H* accent (response bias)	−0.04	0.10	−0.36	0.72
Advanced vs. low/medium proficiency × L + H* accent (response bias)	0.10	0.09	1.08	0.28
Most advanced vs. low/medium/advanced proficiency × L + H* accent (response bias)	−0.04	0.10	−0.43	0.69
Medium vs. low proficiency × L + H* × contrast probe (sensitivity)	0.42	0.29	1.45	0.15
Medium vs. low proficiency × L + H* × unmentioned probe (sensitivity)	−0.36	0.29	1.23	0.22
Advanced vs. low/medium proficiency × L + H* × contrast probe (sensitivity)	0.23	0.26	0.89	0.37
Advanced vs. low/medium proficiency × L + H* × unmentioned probe (sensitivity)	−0.52	0.26	−2.04	0.04
Most advanced vs. low/medium/high proficiency × L + H* × contrast (sensitivity)	0.12	0.29	0.40	0.69
Most advanced vs. low/medium/high proficiency × L + H* × unmentioned (sensitivity)	−0.01	0.29	−0.04	0.97

**stands for the abbreviation.*

TABLE 3 | Fixed effects estimates from mixed logit model of "True" responses with probe type, accent type, and working memory as fixed effects.

	Estimate	SE	Wald *z*	*p*-value
Main effects across proficiency levels				
Baseline rate of *true* responses (response bias)	0.30	0.05	5.66	<0.001
Contrast probe vs. baseline (sensitivity)	−0.55	0.06	−10.01	<0.001
Unmentioned probe vs. baseline (sensitivity)	−0.67	0.06	−12.10	<0.001
L + H* accent (effect on response bias)	0.03	0.04	0.79	0.43
L + H* accent × contrast probe (effect on sensitivity)	0.15	0.11	1.38	0.17
L + H* accent × unmentioned probe (sensitivity)	−0.20	0.11	−1.85	0.06
Effects of working memory				
Working memory (response bias)	−0.03	0.03	−0.76	0.45
Working memory × contrast probe (sensitivity)	−0.12	0.06	−1.85	0.06
Working memory × unmentioned (sensitivity)	−0.14	0.06	−2.18	0.03*
Effects of WM in comprehension of prosody				
Working memory × L + H* accent (response bias)	0.03	0.05	0.58	0.56
WM × L + H* × contrast probe (sensitivity)	−0.01	0.13	−0.12	0.91
WM × L + H* × unmentioned probe (sensitivity)	> −0.01	0.13	> −0.01	0.99

**Significant level = 0.05.*

to discern whether a specific probe is *true*). Specifically, if participants have accurate memory for the discourse, they should respond *true* more often to the correct probes and less often to the contrast and unmentioned probes.

We analyzed participants' *true* responses in a mixed-effects logit model as a function of three fixed effects: pitch accent type, probe type, and proficiency level. All variables were coded with mean-centered contrasts to obtain estimates of main effects analogous to those from an ANOVA. For probe type, we used two effect-coded contrasts: One compared the responses to contrast probes to the mean rate of *true* responses, and one compared responses to unmentioned probes to the mean rate of *true* responses (Fraundorf et al., 2013) (Because our primary interest was in how participants rejected false information about the discourse, we were less interested in responses to correct probes, which essentially constituted fillers). The four ordered proficiency groups were coded using Helmert contrasts, which compares each successive proficiency group to the mean of the less proficient groups (e.g., advanced-proficiency learners versus medium- and low-proficiency learners). **Figure 1** displays the mean rate of *true* responses for each proficiency group in each experimental condition, and **Table 2** displays the results of the mixed-effects model.

Effects of L2 Proficiency

First, we examine overall trends across proficiency groups. The positive intercept term indicates that, overall, participants had a bias to respond *true* rather than *false*, with the odds 1.33 (95% CI: [1.19, 1.50]) in favor of responding *true*. This bias to respond *true* was obtained despite the fact that the majority of probes were false; that is, participants often accepted false statements. Nevertheless, participants responded *true* less frequently to contrast probes and unmentioned probes, indicating that they had at least some veridical memory for the discourse. Specifically, the odds of responding *true* were reduced 1.82 times (95% CI: [1.62, 2.05]) for contrast probes and 2.01 times (95% CI: [1.79, 2.27]) for unmentioned probes.

Effects of Pitch Accents

What about the effects of PAs? Pitch accent type had no main effect on response bias, indicating that contrastive PAs did not simply induce an overall bias to respond *true* or *false*. Rather, pitch accent interacted with probe type: For probes referring to items unmentioned in the discourse, the odds of correct rejection increased by 1.26 times (95% CI: [1.01, 1.56]) when the critical word was originally heard with a contrastive $L + H^*$ pitch accent. By comparison, $L + H^*$ did not significantly facilitate rejection of the salient contrastive alternatives, consistent with the results of Lee and Fraundorf (2017); indeed, the effect was numerically (but non-significantly) in the direction of the $L + H^*$ accent *hindering* correct rejection.

Pitch Accent Effects Qualified by Proficiency

Importantly, however, many of these effects varied across proficiency levels. First, the overall bias to respond *true* was 1.19 times smaller (95% CI: [1.03, 1.36]) for the most advanced learners; that is, at the most advanced proficiency level, participants had less of a tendency to simply accept the presented statements as true. Second, participants with more advanced proficiency were more accurate at judging whether specific probe statements were true. The odds of correctly rejecting an unmentioned probe were 1.40 times greater (95% CI: [1.09, 1.81]) for advanced-proficiency learners than low- or medium-proficiency learners, though advanced-proficiency learners were still no more successful at ruling out the salient contrast items. It was not until the high advanced level of proficiency that learners finally showed greater success in rejecting the contrast probes, with the odds of correct rejection increasing by 1.97 times (95% CI: [1.47, 2.65]) for the most advanced learners as compared to the other groups. The most advanced learners also showed a further 1.40-times increase (95% CI: [1.05, 1.89]) over the advanced learners in the odds of correctly rejecting the unmentioned probes.

Most critically, the effects of prosody were also qualified by proficiency. Specifically, for learners who had attained at least advanced proficiency, the benefit of the contrastive $L + H^*$ accent in rejecting the unmentioned probes was 1.68 times greater than for less proficient learners (95% CI: [1.01, 2.80]); high advanced learners did not further differ in this effect. As noted above, the $L + H^*$ accent did not affect participants' overall tendency to respond *true* or *false*, and this did not interact with proficiency level; that is, at no proficiency level did the PAs affect response bias.

Effects of Working Memory

We also examined whether apparent effects of proficiency reflect proficiency with the language itself or rather the ability to hold more material in working memory. Because we had two working-memory measures (which showed modest agreement, $r = 0.46$, $p < 0.001$), we created a composite measure by averaging each participant's *z*-scores on each of the two tasks; using multiple measures in this way reduces measurement error by reducing the influence of task-specific variance associated with any particular task [e.g., the influence of arithmetic ability on OSpan performance; Cronbach (1957) and Bollen (1989)].

It seems unlikely that our present effects of proficiency can be attributed to working memory: More proficient learners did not necessarily have higher working memory; indeed, the Spearman correlation between proficiency rank and working memory was actually *negative*, rho = −0.31, $p < 0.001$, such that more proficient learners had *lower* working memory. Indeed, as can be seen in **Table 1**, the group that scored highest in working memory was the *low* proficiency group. Thus, it does not seem to be the case that the more proficient groups were more sensitive to contrastive prosody because they had greater working-memory resources.

Nevertheless, as a more direct test of whether working memory accounts for the proficiency effects, we replaced the proficiency variable in the mixed-effects model with the mean-centered working memory score[2] to test whether working memory could be observed to have similar effects as proficiency.

[2]We use each participant's individual working memory score rather than splitting participants into discrete groups because this preserves the full range of variation in working memory and contributes more information and more statistical power to the model (Cohen, 1983).

Table 3 displays the results of the mixed-effects modeling including working memory. There was some evidence that participants with higher working memory scores performed better on the task overall. A 1-standard deviation increase in working memory corresponded to a significant 1.15 times (95% CI: [1.02, 1.29]) increase in the odds of successfully rejecting probes referring to items unmentioned in the discourse, and a marginal 1.13 times (95% CI: [1.00, 1.27]) in the odds of successfully rejecting the contrastive alternatives. Critically, however, working memory and the unmentioned text did not significantly interact with pitch accent type; there was no evidence that the benefit of L + H* accents in rejecting either contrast probes or unmentioned probes was enhanced for people with greater working memory (both *ps* > 0.90).

DISCUSSION

The present study examined how L1 Chinese learners of L2 English processed and remembered spoken L2 discourses containing contrastive PAs (L + H*) or non-contrastive presentational PAs (H*). We tested L2 learners' memory using a recognition task including three types of probes: correct, contrastive alternative, and unmentioned items. We also assessed the readers' language proficiency (using two standardized English proficiency scores) and working memory (RSpan and OSpan).

We contrasted three hypotheses about how contrastive PAs might influence memory for a discourse: the granularity account, the contrast representation account, and the contrast uncertainty account. The granularity account predicts that the salient acoustic or perceptual aspects of a contrastive PA facilitate memory representations of the accented word itself, which should help comprehenders reject any items inconsistent with the true statement (Sanford et al., 2006). Alternatively, the contrast-representation hypothesis proposes that contrastive PAs, relative to presentational PAs, promote representation of a specific salient alternative and should facilitate rejection of only that salient alternative, not a completely unmentioned item; this pattern has been found for L1 English comprehenders (Fraundorf et al., 2010). Lastly, the contrastive uncertainty hypothesis (Lee and Fraundorf, 2017) proposes that because contrastive PAs evoke the salient alternative, they lead to confusion over which was the correct proposition and which was the salient alternative; this should allow comprehenders to easily reject the unmentioned items, but to have difficulty discriminating the correct and contrast items.

Our principal findings are threefold. First, across all proficiency levels, our L2 English learners did not show a native-like contrast-representation effect in which contrastive PAs facilitated rejections of a specific salient alternative in memory. Instead, to the extent PAs influenced L2 comprehenders' memory at all, they showed a contrastive uncertainty pattern. Contrastive PAs helped L2 learners reject the unmentioned item that was never part of the discourse, but they impaired L2 learners' ability to discriminate between the correct item and its salient alternative. This finding replicates previous studies among a different population of L2 English learners whose L1 was Korean (Lee and Fraundorf, 2017, 2021). Second, we found a significant interaction effect of proficiency by contrastive PA. Specifically, the benefit of contrastive PA in rejecting unmentioned items was enhanced for both advanced and high-advanced learners relative to low- or medium-proficiency learners. More proficient participants also showed more accurate memory for the discourses overall, as well as a reduced overall bias to affirm the presented statements as *true*. Third, there was no evidence that the benefit of contrastive PAs in rejecting either contrast probes or unmentioned probes was enhanced for people with greater working memory. We discuss the implications of each of these findings below.

A Contrastive Uncertainty Effect in L2 Pitch Accent Comprehension

The current study indicates that contrastive PAs led L2 listeners to represent salient alternatives differently from L1 English native speakers. For native speakers, emphasizing a word with a contrastive PA helped listeners rule out a specific alternative to that word on a later memory test, suggesting that they had represented that particular alternative in memory. L2 learners did not derive these same memory benefits. Among the low- and mid-proficiency groups, there were no mnemonic benefits of contrastive accents whatsoever. Among more proficient learners, there was a different effect such that contrastive PAs facilitated rejection of items entirely unmentioned in the discourse. This pattern suggests that L2 learners may have represented the *set* of alternatives—which would help reject any item in the set—but failed to distinguish which was the true proposition and which was the salient alternative. Interestingly, the most advanced group revealed a somewhat better ability to rule out the contrastive alternative, but this was not qualified by PA type, suggesting that lexical tone may interfere with representations of English PA even in this group.

These results replicate those of previous studies of L1-Korean L2-English undergraduates (Lee and Fraundorf, 2017, 2021). Notably, however, we replicate them in a population with a different L1 (Chinese) and with a wider age range (from high school to graduate study). This suggests that the prior results were not simply an idiosyncratic effect in L1 Korean learners. Rather, a more general property of second-language processing may be difficulty in distinguishing the members of a set of alternatives.

This difficulty in distinguishing members of a contrast set may have been enhanced by the isolated nature of our stimuli. Theories of memory generally distinguish *episodic memory* for things and events that happened to a person from *semantic memory*, or more general knowledge (James, 1890). Although our materials were semantically coherent and comprehensible, these short, discrete stories were largely distinct from listeners' prior semantic knowledge and primarily tapped episodic memory. This may have made it particularly difficult to distinguish the two members of the contrastive set (such as whether the *British* or *French* scientists found the monkey) because this relied entirely on detailed episodic information. But an irrelevant or unmentioned item, such as *Portuguese* scientists, could be more easily rejected since this piece of semantic memory did not exist in the memory trace.

Proficiency-Driven Pitch Accent Effects on Memory Representation

We found modulation and qualification of PA effect by proficiency. Similar to previous research (Lee and Fraundorf, 2017, 2021), less proficient learners did not show any sensitivity to contrastive PAs whereas more proficient learners did—though even more proficient learners did not show fully native-like comprehension.

Why is proficiency critical to capitalizing on prosodic information in remembering a discourse? We speculate there at least two reasons. First, L2 proficiency can shape language processing and cognition more generally such that proficient L2 learners have better attentional focus, which supports veridical memory encoding and recognition. This attention advantage for bilinguals has been demonstrated consistently in the reading and second language learning literature (Bialystok, 2017). Thus, even though the L2 listeners in our study did not reach the highest level of native-L1-like performance, they have benefited over less proficient L2 learners in focusing attention and storing contrastive information more stably and steadily in memory.

Second, L2 proficiency may be critical to understanding the meaning of PAs themselves and to understanding how interference from L1 lexical tone affects their representation. Although we are unaware of any research examining the influence of a tonal L1 on L2 PA interpretation, the finding that contrastive PA failed to facilitate the most advanced L2 English learners' ability to rule out contrastive alternatives suggests that L1 lexical tone may have interfered with their representations of L2 PA, preventing them from using it to strengthen memory for contrastive information in discourse as native speakers do. Native-like processing of English PAs requires L2 English learners to learn how particular PA types should be mapped to discourse representation. This linguistic knowledge may be acquired only gradually with increasing proficiency, especially since it is rarely taught in formal instructions. Consistent with this claim, Lee and Fraundorf (2019) suggest that L2 learners *can* make use of other cues to focus whose purpose may be more readily apparent, such as font emphasis.

Working Memory Effects

We also observed a significant effect of WM in that WM predicted participants' ability to correctly rule out the probe statements referring to items wholly unmentioned in the original discourse. This effect may be thought of in terms of familiarity (Yonelinas, 2002) or episodic memory traces. In example (1), since both *British* and *French* appeared in the discourse in some capacity, it may have been difficult to distinguish them. By comparison, the unmentioned lure *Portuguese* did not appear in the discourse at all and would have not existed in the memory trace as this part of the discourse wasn't mentioned, so participants could have more confidently rejected it. Thus, performance on the task may be related to participants' ability to retrieve details of the memory traces. This is in line with the argument that deficits in episodic memory are associated with reduced retrieval of episodic details and reduced coherence of discourse (Seixas-Lima et al., 2020). Another reason may be that the trace retrieval strategy of episodic memory promotes the long-term retention of bilingual vocabulary in the mind. This dovetails with the finding of Zhang et al. (2021) that the non-verbal episodic memory ability of highly proficient bilinguals contributes to bilingual vocabulary development. Nonverbal episodic memory skills contribute to lexical competence because participants with them become more proficient at higher levels.

Working memory may be especially important for L2 discourse comprehension (although WM also predicts baseline performance in discourse memory even for L1 participants; Fraundorf et al., 2012). As claimed by Witzel and Forster (2012), L2 words must be stored in working memory; however, L1 words are placed in the semantic systems that store knowledge. The link between episodic memory ability and L2 lexical competence suggests that episodic memory may play a role not only in initial second language lexical acquisition but also possibly in long-term retention and representation of second language vocabulary. These results are also consistent with the episodic L2 hypothesis (e.g., Jiang and Forster, 2001; Witzel and Forster, 2012), which predicts that episodic L2 lexical representations persist even at higher levels of bilinguals' proficiency at later stages. Memory plays an important role in bilinguals' L2 lexical repertoires. Furthermore, these results suggest that individual differences in working memory affect memory for L2 spoken discourse. L2 episodic memory is activated first, due to repetition of words (Witzel and Forster, 2012); second, due to prosodic feature adjustments, e.g., similarity of the talker in voice (Shao et al., 2017), segmental and suprasegmental features (Lengeris, 2012), or conjunction illusions (e.g., list 1 and list 2) (Brainerd et al., 2014).

Critically, however, while working memory predicted overall performance, we did *not* find that working memory moderated or qualified the use of the information conveyed by contrastive pitch accents: The pitch accenting effect was no larger (or smaller) for participants higher in working memory. This suggests that limits in the ability to carry out online cognitive operations were not the reason that L2 learners struggled to make use of the contrastive pitch accents (see also Lee and Fraundorf, 2021 for a similar conclusion).

Significance and Practical Implications

What practical applications does this current study have? We found that L1 Chinese learners of L2 English did have some ability to leverage contrastive PAs in language comprehension, at least when they were relatively advanced in proficiency. This suggests that contrastive PAs can be useful in conveying information even to L2 listeners. Nevertheless, our results suggest that L2 learners may have limited and insufficient knowledge of the meaning of L2 intonation (as we discuss above). Thus, it may be beneficial to teach L2 learners how to attend to and interpret salient intonational information.

Second, methodologically, we capitalized on the use of mixed-effects models to address the problems of a non-normal dependent variable. These models permit the use of link functions, such as the log odds (known as the logit), to relate experimental or observational variables to outcome variables that are not normally distributed, such as binomial outcomes like recognition accuracy (Baayen et al., 2008; Jaeger, 2008). Further, by incorporating information from multiple levels—both trial-level characteristics of the experimental design and subject-level individual differences, such as working memory—and their interaction, we could examine how different types of English L2 learners leverage PA cues.

Limitations and Future Directions

Because the discourses were presented aurally, one question is whether contrastive PA interpretation varied with proficiency effects simply because only more proficient learners could comprehend the lexical and syntactic content of the auditory input. This explanation may apply to some extent: Overall memory accuracy, regardless of PA type, increased with proficiency. Nevertheless, above and beyond these effects, we found effects of proficiency on how contrastive PAs affect memory.

Although our work suggests that comprehenders gradually learn how to map L2 PAs onto particular meanings with increasing proficiency, one question for future work is how, precisely, this mapping is acquired. The contrastive L + H* PA is acoustically more salient than the presentational H* PA, but earlier research among L1 native speakers suggested that the mnemonic benefit of contrastive PA stem from their contrastive interpretation and not merely its audibility or perceptual salience (Cutler et al., 1997; Fraundorf et al., 2010; Wagner and Watson, 2010). Therefore, future work could examine how perceptual features (such as embodied perceptual symbols) are integrated into the ultimate memory representation of the text (Barsalou, 1999; Del Giudice et al., 2004).

The current study is based upon L2 English proficiency, and we did not assess the students' L1 Chinese language proficiency and fluency. Thus, it's possible that a confound between Chinese and English ability may serve as an alternative explanation if the pitch accents that participants heard were easy to relate to their Chinese L2 native language.

Finally, although recent research suggests that L2 learners tend to be more sensitive to online sentence processing as their WM capacity in L1 increases (Coughlin and Tremblay, 2013), we did not observe any such effects. One reason for this may be that WM was inversely related to proficiency within our sample—WM scores were highest among the low-proficiency group—which might have obscured any potential WM effects. Future research could more thoroughly investigate this issue by examining variability in WM within L2 speakers with similar proficiency.

CONCLUSION

We examined that how PAs influenced how L1-Chinese learners of L2 English comprehended and remembered a spoken L2 discourse. We compared four L2 proficiency groups (low, medium, advanced, and high advanced) based on their Quick Placement Text (QPT) levels. Signal detection analysis (implemented via mixed-effects modeling) revealed that L2-English learners were more sensitive to PA as L2 proficiency increased. However, even the most advanced learners showed a pattern of memory effects distinct from native speakers: Rather than discriminating a correct proposition from a salient alternative in the discourse, contrastive PAs facilitated rejection only of items never mentioned in a discourse, suggesting that L2 learners had difficulty discriminating the items within contrast sets. Further, these effects were influenced only by proficiency and not by WM, suggesting they reflect incomplete knowledge of the intonation-to-meaning mapping more than limitations in online processing resources.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Science and Technology. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CG and SF designed the study. CG conducted the research. SF analyzed the data. All authors wrote the manuscript and approved the submitted version.

FUNDING

This work was funded by Fulbright funds (E0582119), and China NSF funds (31571155), (31500915), and (62077011) awarded to CG.

ACKNOWLEDGMENTS

We thank two research assistants, Wei Zhang and Jianrong Zhao, who helped with data collection and pre-processing.

REFERENCES

Akker, E., and Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism* 6, 81–96. doi: 10.1017/s1366728903001056

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412. doi: 10.1080/00273171.2021.1889946

Baker, R. E. (2010). "Non-native perception of native English prominence," in *Proceedings of the 5th International Conference on Speech Prosody*, (Chicago, IL: Cambridge University Press).

Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 637–660.

Beckman, M. E. (1986). *Stress and Non-Stress Accent*. Dordrecht: Foris Publications.

Beckman, M. E., and Elam, G. A. (1997). *Guidelines for ToBI Labelling, Version 3*. Columbus, OH: Ohio State University.

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1037/bul0000099

Birch, S. L., and Garnsey, S. M. (1995). The effect of focus on memory for words in sentences. *J. Mem. Lang.* 34, 232–267. doi: 10.1006/jmla.1995.1011

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociol. Methods Res.* 17, 303–316. doi: 10.1016/j.apmr.2021.06.016

Brainerd, C. J., Holliday, R. E., Nakamura, K., and Reyna, V. F. (2014). Conjunction illusions and conjunction fallacies in episodic memory. *J. Exp. Psychol.* 40:1610. doi: 10.1037/xlm0000017

Braun, B., and Tagliapietra, L. (2011). On-line interpretation of intonational meaning in L2. *Lang. Cogn. Proc.* 26, 224–235. doi: 10.1080/01690965.2010.486209

Clahsen, H., and Felser, C. (2006). How native-like is non-native language processing? *Trends Cogn. Sci.* 10, 564–570. doi: 10.1016/j.tics.2006.10.002

Cohen, J. (1983). The cost of dichotomization. *Appl. Psychol. Measur.* 7, 249–253. doi: 10.1177/014662168300700301

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., and Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychon. Bull. Rev.* 12, 769–786. doi: 10.3758/bf03196772

Coughlin, C. E., and Tremblay, A. (2013). Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. *Appl. Psycholing.* 34, 615–646. doi: 10.1017/s0142716411000890

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *Am. Psychol.* 12, 671–684. doi: 10.1023/a:1014592301729

Curtis, C. E., and D'Esposito, M. (2002). Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.* 7, 415–423. doi: 10.1016/s1364-6613(03)00197-9

Cutler, A., Dahan, D., and Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Lang. Speech* 40, 141–201. doi: 10.1177/002383099704000203

Dahan, D., Tanenhaus, M. K., and Chambers, C. G. (2002). Accent and reference resolution in spoken language comprehension. *J. Mem. Lang.* 47, 292–314. doi: 10.1007/s10936-014-9306-6

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verb. Learn. Verb. Behav.* 19, 450–466.

Del Giudice, P., Fusi, S., and Mattia, M. (2004). Modelling the formation of working memory with networks of integrate-and-fire neurons connected by plastic synapses. *Physiol. Paris* 97, 659–681. doi: 10.1016/j.jphysparis.2004.01.021

Duanmu, S. (2004). Tone and non-tone languages: An alternative to language typology and parameters. *Lang. Linguis.* 5, 891–924.

Dupoux, E., Pallier, C., Sebastian, N., and Mehler, J. (1997). A destressing "deafness" in French? *J. Mem. Lang.* 36, 406–421. doi: 10.1016/j.cognition.2007.04.001

Engle, R. W. (2010). Role of working-memory capacity in cognitive control. *Curr. Anthropol.* 51, S17–S26.

Evans, V. (2006). Lexical concepts, cognitive models and meaning-construction. *Cogn. Ling.* 17, 491–534.

Felser, C., and Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Sec. Lang. Res.* 23, 9–36. doi: 10.1177/0267658307071600

Fraundorf, S. H., Benjamin, A. S., and Watson, D. G. (2013). What happened (and what did not): Discourse constraints on encoding of plausible alternatives. *J. Mem. Lang.* 69, 196–227. doi: 10.1016/j.jml.2013.06.003

Fraundorf, S. H., Watson, D. G., and Benjamin, A. S. (2010). Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *J. Mem. Lang.* 63, 367–386. doi: 10.1016/j.jml.2010.06.004

Fraundorf, S. H., Watson, D. G., and Benjamin, A. S. (2012). The effects of age on the strategic use of pitch accents in memory for discourse: A processing-resource account. *Psychol. Aging* 27, 88–98. doi: 10.1037/a0024138

French-Mestre, C. (2002). "An on-line look at sentence processing in the second language," in *Bilingual Sentence Processing*, eds R. Heredia and J. Altarriba (New York, NY: Elsevier), 217–236. doi: 10.1016/s0166-4115(02)80012-7

Geranpayeh, A. (2003). A quick review of the English Quick Placement Test. *Res. Notes* 12, 8–10.

Goodman, N., and Lassiter, D. (2015). "Probabilistic semantics and pragmatics: Uncertainty in language thought," in *Handbook of Contemporary Semantics*, 2nd Edn, eds S. Lappin and C. Fox (Wiley-Blackwell).

Gotzner, N., Spalek, K., and Wartenburger, I. (2013). "How pitch accents and focus particles affect the recognition of contextual alternatives," in *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, eds M. Knauff, M. Pauen, N. Sebanz, and I. Wachsmuth (Texas, US: Cognitive Science Society), 2434–2439.

Grabe, W. (2004). Research on teaching reading. *Annu. Rev. Appl. Linguist.* 24, 44–69.

Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*, Vol. 1. New York, NY: Wiley.

Gut, U. (2003). Prosody in second language speech production: the role of the native language. *Zeitschrift für Fremdsprachen Lehren Lernen* 32, 133–152. doi: 10.1177/0023830918773118

Gut, U., and Milde, J.-T. (2002). "The prosody of Nigerian English," in *Proceedings of the Speech Prosody 2002 Conference*, (France: Aix-en-Provence), 367–370.

Gut, U., Trouvain, J., and Barry, W. J. (2007). "Bridging research on phonetic descriptions with knowledge from teaching practice–The case of prosody in non-native speech," in *Non-Native Prosody: Phonetic Description and Teaching Practice*, eds J. Trouvain and U. Gut (New York, NY: De Gruyter Mouton), 3–21. doi: 10.1515/9783110198751.0.3

Harrington, M. (1992). Working memory capacity as a constraint on L2 development. *Adv. Psychol.* 83, 125–135.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Sec. Lang. Res.* 22, 369–397. doi: 10.1191/0267658306sr272oa

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua* 120, 901–931. doi: 10.1016/j.lingua.2009.06.004

Huang, Y.-T., and Arnold, A. R. (2010). Word learning in linguistic context: Processing and memory effects. *Cognition* 156, 71–87. doi: 10.1016/j.cognition.2016.07.012

Ito, K., and Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *J. Mem. Lang.* 58, 541–573. doi: 10.1016/j.jml.2007.06.013

Jackson, C. N. (2008). Proficiency level and the interaction of lexical and morphosyntactic information during L2 sentence processing. *Lang. Learn.* 58, 875–909. doi: 10.1111/j.1467-9922.2008.00481.x

Jackson, C. N., and Dussias, P. E. (2009). Cross-linguistic differences and their impact on L2 sentence processing. *Bilingualism* 12, 69–82. doi: 10.3389/fpsyg.2021.673761

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *J. Mem. Lang.* 59, 434–446. doi: 10.1016/j.jml.2007.11.007

James, W. (1890). *The Principles of Psychology*. New York, NY: Henry Holt.

Jiang, N., and Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *J. Mem. Lang.* 44, 32–51. doi: 10.1006/jmla.2000.2737

Jilka, M. (2000). "Testing the contribution of prosody to the perception of foreign accent," in *Proceedings of New Sounds: 2000 4th International Symposium on the Acquisition of Second Language Speech*, eds A. James and J. Leather (Amsterdam: University of Klagenfurt), 199–207. doi: 10.1159/000097308

Just, M. A., and Carpenter, P. A. (1987). *The Psychology of Reading and Language Comprehension*. Newton, MA: Allyn & Bacon.

Just, M. A., and Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychol. Rev.* 99, 122–149. doi: 10.1037/0033-295x.99.1.122

Kim, S. Y., Qi, T., Feng, X., Ding, G., Liu, L., and Cao, F. (2016). How does language distance between L1 and L2 affect the L2 brain network? An fMRI study of Korean–Chinese–English trilinguals. *NeuroImage* 129, 25–39. doi: 10.1016/j.neuroimage.2015.11.068

Ladd, D. R. (2008). *Intonational Phonology*. Cambridge, UK: Cambridge University Press.

Lee, E.-K., and Fraundorf, S. H. (2017). Effects of contrastive accents in memory for L2 discourse. *Bilingualism* 20, 1063–1079. doi: 10.1017/s1366728916000638

Lee, E.-K., and Fraundorf, S. H. (2019). Native-like processing of prominence cues in L2 written discourse comprehension: Evidence from font emphasis. *Appl. Psycholing.* 40, 373–398.

Lee, E.-K., and Fraundorf, S. H. (2021). Do L1-L2 differences in discourse processing reflect processing demands or difficulty of form-function mapping?: Evidence from self-paced listening of contrastive prosody. *Studies in Second Language Acquisition*. doi: 10.1017/S0272263121000619

Lee, E.-K., and Snedeker, J. (2016). Effects of contrastive accents on children's discourse comprehension. *Psychon. Bull. Rev.* 23, 1589–1595. doi: 10.3758/s13423-016-1069-7

Lee, Y., Wang, B., Chen, S., Adda-Decker, M., Amelot, A., Nambu, S., et al. (2015). "A crosslinguistic study of prosodic focus," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, (Cambridge, UK: Cambridge University Press), 4754–4758.

Lee, Y. C., Wang, T., and Liberman, M. (2016). Production and perception of tone 3 focus in Mandarin Chinese. *Front. Psychol.* 7:1058. doi: 10.3389/fpsyg.2016.01058

Lengeris, A. (2012). "Prosody and Second Language Teaching: Lessons from L2 Speech Perception and Production Research," in ed J. Romero-Trillo, *Pragmatics and Prosody in English Language Teaching. Educational Linguistics*, (Dordrecht: Springer), doi: 10.1007/978-94-007-3883-6_3.

Li, A., and Post, B. (2014). Acquisition of prosodic properties of speech rhythm: Evidence from L1 Mandarin and German learners of English. *Stud. Sec. Lang. Acquis.* 36, 223–255. doi: 10.1017/s0272263113000752

Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Magen, Y. (1998). Ancient Israel's stone age. *Bibl. Archaeol. Rev.* 24:46.

Marinis, T., Roberts, L., Felser, C., and Clahsen, H. (2005). Gaps in second language processing. *Stud. Sec. Lang. Acquis.* 27, 53–78.

Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *J. Phon.* 32, 543–563. doi: 10.1016/j.wocn.2004.02.002

Mennen, I. (2015). "Beyond segments: Towards a L2 intonation learning theory," in *Prosody and Language in Contact*, eds E. Delais-Roussaire, M. Avanzi, and S. Herment (Berlin: Springer), 171–188. doi: 10.1007/978-3-662-45168-7_9

Mennen, I., Chen, A., and Karlsson, F. (2010). "Characterising the internal structure of learner intonation and its development over time," in *Proceedings of New Sounds 2010: 6th International Symposium on the Acquisition of Second Language Speech*, eds K. Dziubalska-Kołaczyk, M. Wrembel, and M. Kul (Poznan: Adam Mickiewicz University), 319–324.

Mennen, I., and De Leeuw, E. (2014). Beyond segments: Prosody in SLA. *Stud. Sec. Lang. Acquis.* 36, 183–194. doi: 10.1017/s0272263114000138

Morett, L. M., and Fraundorf, S. H. (2019). Listeners consider alternative speaker productions in discourse comprehension and memory: Evidence from beat gesture and pitch accenting. *Mem. Cogn.* 47, 1515–1530. doi: 10.3758/s13421-019-00945-1

Morett, L. M., Fraundorf, S. H., and McPartland, J. C. (2021). Eye see what you're saying: Contrastive use of beat gesture and pitch accent affects online interpretation of spoken discourse. *J. Exp. Psychol.* 47, 1494–1526. doi: 10.1037/xlm0000986

Morett, L. M., Roche, J. M., Fraundorf, S. H., and McPartland, J. C. (2020). Contrast is in the eye of the beholder: Infelicitous beat gesture increases cognitive load during online spoken discourse comprehension. *Cogn. Sci.* 44:e12912. doi: 10.1111/cogs.12912

Murayama, K., Sakaki, M., Yan, V. X., and Smith, G. M. (2014). Type I error inflation in the traditional by-participant analysis to metamemory accuracy: A generalized mixed-effects model perspective. *J. Exp. Psychol.* 40:1287. doi: 10.1037/a0036914

Nolan, F. (2006). "Intonation," in *Handbook of English linguistics*, eds B. Aarts and A. McMahon (Oxford: Blackwell), 433–457.

Norberg, K. A., and Fraundorf, S. H. (2021). Memory benefits from contrastive focus truly require focus: evidence from clefts and connectives. *Lang. Cogn. Neurosci.* 36, 1010–1037. doi: 10.1080/23273798.2021.1901944

Ouyang, I., and Kaiser, E. (2012). Focus-marking in a tone language: Prosodic cues in Mandarin Chinese. *LSA Annu. Meet. Extend. Abstr.* 3, 1–5.

Papadopoulou, D. (2005). Reading-time studies of second language ambiguity resolution. *Sec. Lang. Res.* 21, 98–120. doi: 10.1191/0267658305sr254oa

Pennington, M. C., and Ellis, N. C. (2000). Cantonese speakers' memory for English sentence with prosodic cues. *Modern Lang. J.* 84, 372–389. doi: 10.1111/0026-7902.00075

Peperkamp, S., and Dupoux, E. (2002). A typological study of stress deafness. *Lab. Phonol.* 7, 203–240. doi: 10.1515/9783110197105.1.203

Perfetti, C. A. (1997). "The psycholinguistics of spelling and reading," in *Learning to Spell: Research, Theory, and Practice Across Languages*, eds C. A. Perfetti, L. Rieben, and M. Fayol (England, UK: Routledge), 21–38.

Pierrehumbert, J., and Hirschberg, J. (1990). "The meaning of intonation in the interpretation of discourse," in *Intentions in Communication*, eds P. Cohen, J. Morgan, and M. Pollack (Cambridge, MA: MIT Press), 271–311. doi: 10.1515/9783110261790.271

Pinker, S. (1999). *Words and Rules: The Ingredients of Language*. New York, NY: Basic.

Pinker, S., and Ullman, M. T. (2002). The past and future of the past tense. *Trends Cogn. Sci.* 6, 456–463. doi: 10.1016/s1364-6613(02)01990-3

Quick Placement Test [QPT] (2001). *Paper and Pen Test Edition Quick Placement Test [kit]/Local Examinations Syndicate*. Cambridge, UK: University of Cambridge.

Rasier, L., and Hiligsmann, P. (2007). Prosodic transfer from L1 to L2: Theoretical and methodological issues. *Nouveaux Cahiers de Linguistique Française* 28, 41–66.

Rasier, L., and Hiligsmann, P. (2009). "Exploring the L1-L2 relationship in the L2 acquisition of prosody," in *Online Proceedings of first and Second Languages: Exploring the Relationship in Pedagogy-Related Contexts*, (Oxford, UK: University of Cambridge).

Redick, T. S., and Engle, R. W. (2006). Working memory capacity and attention network test performance. *Appl. Cogn. Psychol.* 20, 713–721. doi: 10.1002/acp.1224

Rooth, M. (1992). A theory of focus interpretation. *Natural Lang. Sem.* 1, 75–116. doi: 10.1007/bf02342617

Sanford, A. J., MuMolle, J., and Emmott, C. (2006). Shallow processing and attention capture in written and spoken discourse. *Discour. Proc.* 42, 109–130. doi: 10.1207/s15326950dp4202_2

Scharff-Rethfeldt, W., Miller, N., and Mennen, I. (2008). Unterschiede in der mittleren Sprechtonhöhe bei Deutsch/Englisch bilingualen Sprechern [Speaking fundamental frequency differences in German-English bilinguals]. *Sprache Stimme Gehör* 32, 123–128. doi: 10.1055/s-0028-1083799

Seixas-Lima, B., Murphy, K., Troyer, A. K., Levine, B., Graham, N. L., Leonard, C., et al. (2020). Episodic memory decline is associated with deficits in coherence of discourse. *Cogn. Neuropsychol.* 37, 511–522. doi: 10.1080/02643294.2020.1770207

Shao, J., Mak, C. Y., and Zhang, C. (2017). The role of talker similarity in the perceptual learning of L2 tone categories. *Cogn. Sci. Soc.* 14:e0221498.

Shen, X. S. (1989). Toward a register approach in teaching Mandarin tones. *J. Chin. Lang. Teach. Assoc.* 24, 27–47.

Silverman, K. E., Beckman, M. E., Pitrelli, J. F., Ostendorf, M., Wightman, C. W., Price, P., et al. (1992). ToBI: A standard for labeling English prosody. *ICSLP* 2, 867–870.

Sturt, P., Sanford, A. J., Stewart, A. J., and Dawydiak, E. (2004). Linguistic focus and good-enough representations: An application of the change-detection paradigm. *Psychon. Bull. Rev.* 11, 882–888. doi: 10.3758/bf03196716

Taylor, I., and Taylor, M. M. (2014). *Writing and literacy in Chinese, Korean and Japanese: Revised edition*, Vol. 14. Amsterdam: John Benjamins Publishing Company.

Trofimovich, P., and Baker, W. (2006). Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Stud. Sec. Lang. Acquis.* 28, 1–30.

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *J. Psycholing. Res.* 30, 37–69. doi: 10.1023/a:1005204207369

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/ procedural model. *Cognition* 92, 231–270. doi: 10.1016/j.cognition.2003.10.008

van den Noort, M. W. M. L., Bosch, P., and Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *Europ. Psychol.* 11, 289–296. doi: 10.1027/1016-9040.11.4.289

Vanlancker-Sidtis, D. (2003). Auditory recognition of idioms by native and nonnative speakers of English: It takes one to know one. *Appl. Psycholing.* 24, 45–57. doi: 10.1017/s0142716403000031

Wagner, M., and Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Lang. Cogn. Proc.* 25, 905–945. doi: 10.1080/01690961003589492

Wang, M., Anderson, A., Cheng, C., Park, Y., and Thomson, J. (2008). General auditory processing, Chinese tone processing, English phonemic processing and English reading skill: A comparison between Chinese-English and Korean-English bilingual children. *Read. Writing* 21, 627–644. doi: 10.1007/s11145-007-9081-y

Wang, T., Liu, J., Lee, Y. H., and Lee, Y. C. (2020). The interaction between tone and prosodic focus in Mandarin Chinese. *Lang. Ling.* 21, 331–350. doi: 10.1016/j.cognition.2020.104311

Watson, D. G., Tanenhaus, M. K., and Gunlogson, C. A. (2008). Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cogn. Sci.* 32, 1232–1244. doi: 10.1080/03640210802138755

Willems, N. (1982). *English Intonation from a Dutch Point of View*. Dordrecht: Foris.

Witzel, N. O., and Forster, K. I. (2012). How L2 words are stored: The episodic L2 hypothesis. *J. Exp. Psychol.* 38:1608. doi: 10.1037/a0028072

Wright, D. B., Horry, R., and Skagerberg, E. M. (2009). Functions for traditional and multilevel approaches to signal detection theory. *Behav. Res. Methods* 41, 257–267. doi: 10.3758/BRM.41.2.257

Yang, A., and Chen, A. (2014). "Prosodic focus-marking in Chinese four- and eightyear-olds," in *Social and Linguistic Speech Prosody*, eds N. Campbell, D. Gibbon, and D. Hirst (Dublin: Trinity College Dublin), 713–717.

Yang, A., and Chen, A. (2018). The developmental path to adult-like prosodic focus-marking in Mandarin Chinese-speaking children. *First Lang.* 38, 26–46. doi: 10.1177/0142723717733920

Yonelinas, A. P. (2002). The nature of recollection and familiarity: a review of 30 years of research. *J. Memory Lang.* 46, 441–517. doi: 10.1006/jmla.2002.2864

Zerbian, S. (2015). "Markedness considerations in L2 prosodic focus and givenness marking," in *Prosody and Language in Contact. Prosody, Phonology and Phonetics*, eds E. Delais-Roussarie, M. Avanzi, and S. Herment (Heidelberg: Springer).

Zhang, Y., Ridchenko, M., Hayashi, A., and Hamrick, P. (2021). Episodic memory contributions to second language lexical development persist at higher proficiencies. *Appl. Cogn. Psychol.* 35, 1356–1361. doi: 10.1002/acp.3865

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 27 May 2022
doi: 10.3389/fpsyg.2022.881778

Speech Disfluencies in Consecutive Interpreting by Student Interpreters: The Role of Language Proficiency, Working Memory, and Anxiety

*Nan Zhao**

Department of Translation, Interpreting and Intercultural Studies, Hong Kong Baptist University, Kowloon, Hong Kong SAR, China

Growing research has revealed that interpreters' individual cognitive differences impact interpreting. In this article, I examined how an interpreter's language proficiency, working memory, and anxiety level impact speech disfluencies in target language delivery. Fifty-three student interpreters took part in three cognitive tests, respectively, of their proficiency in English (their non-native language), working memory, and anxiety level. Then they consecutively interpreted an English speech into Mandarin (their native language); their target language output was coded for different types of disfluencies (pauses, fillers, repetitions, and articulatory disfluency). It was found that anxiety level, but not language proficiency and working memory, impacted the occurrence of disfluencies in general. In particular, more anxious interpreters tended to have more fillers, such as *er* and *um*, and more repetitions of words and phrases. I discuss these findings in terms of how anxiety may impact the cognitive processes of interpreting and how to reduce student interpreters' anxiety level in interpreting teaching and learning.

Keywords: interpreting, language proficiency, working memory, anxiety, disfluency

OPEN ACCESS

Edited by:
Yan Jing Wu,
Ningbo University, China

Reviewed by:
Dechao Li,
Hong Kong Polytechnic University,
Hong Kong SAR, China
Xingcheng Ma,
Southeast University, China

***Correspondence:**
Nan Zhao
nanzhao@hkbu.edu.hk

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *23 February 2022*
Accepted: *07 March 2022*
Published: *27 May 2022*

Citation:
Zhao N (2022) Speech Disfluencies in Consecutive Interpreting by Student Interpreters: The Role of Language Proficiency, Working Memory, and Anxiety. Front. Psychol. 13:881778. doi: 10.3389/fpsyg.2022.881778

INTRODUCTION

Interpreters translate from a source language to a target language. Such a task is often delivered under time pressure, in front of an audience, and requires multitasking. For instance, in consecutive interpreting, a speaker delivers a segment of speech (varying from one to a dozen sentences) and the interpreter needs to quickly transcode the source language (e.g., words, syntactic structure) into the target language, which they keep in their working memory (or on a note); then, when the speaker pauses, they output the target language as fluently and accurately as possible to an audience. Thus, to successfully accomplish an interpreting task, the interpreter needs to, among other things, be fully proficient in both the source and target language, actively keep a large amount of linguistic information in their working memory, and overcome the anxiety and stress of public speaking.

In this article, I focus on one important aspect of target language delivery, namely speech disfluencies in the target language. I examine how speech disfluencies vary as a function of the interpreter's cognitive traits: language proficiency, working memory, and anxiety. Below, I first review how these different cognitive traits may impact interpreting and then review speech disfluencies in interpreting, before reporting my own study.

Language Proficiency in Interpreting

As interpreters often interpret between a native language and a non-native language, proficiency in the non-native language is critical for interpreting. Blasco Mayor (2015) showed that student interpreters' listening comprehension ability predicts their interpreting performance. She argued that it's important to train students' listening skills in interpreting teaching and learning. Jiménez Ivars et al. (2014) showed that both interpreting performance and self-efficacy in student interpreters increased as a function of their non-native language proficiency. Christoffels et al. (2003) also showed that the speed with which interpreters retrieve translation equivalents between languages and the speed with which they name pictures are correlated with their interpreting performance, again highlighting the role of language proficiency in interpreting. Indeed, there is evidence that interpreting training often improves student interpreters' language skills compared to non-interpreting bilingual controls (Tzou et al., 2011).

Working Memory in Interpreting

Working memory has been shown to play a critical role in many aspects of language processing, including language comprehension (e.g., Daneman and Carpenter, 1983; Waters and Caplan, 2005) and language production (e.g., Belke, 2008; Martin and Slevc, 2014). As interpreting involves both the comprehension of the source language and the production of the target language, it is no surprise that working memory has long been assumed to likewise play a critical role in interpreting models (Gerver, 1975; Moser, 1978; Darò and Fabbro, 1994). There is also empirical evidence for the role of working memory in interpreting. First, it has been shown that trained interpreters outperformed bilingual controls on working memory tests (Christoffels et al., 2006; Hodáková, 2009; Tzou et al., 2011), suggesting that cognitive resources (i.e., working memory) are a critical sub-capacity for interpreting. However, other studies have not found a reliable difference in working memory capacities between interpreters and bilinguals controls (Liu et al., 2004; Köpke and Nespoulous, 2006).

Instead of comparing interpreters with bilingual controls, other studies have investigated whether interpreting performance relates to an interpreter's working memory capacity. Timarová (2008) showed that some working memory functions (especially the capacity to inhibit irrelevant information) correlate with simultaneous interpreting performance. Christoffels et al. (2003) showed that working memory makes a contribution to interpreting skills independent of an interpreter's language proficiency. There is also evidence that a sign interpreting performance correlates with working memory span (Van Dijk et al., 2012). These findings thus point to a positive correlation between working memory and interpreting performance.

Anxiety in Interpreting

It goes without saying that interpreting is a very stressful activity because it involves performing a series of complex cognitive and psychomotor for an audience, whether in public or private. Students training to become interpreters have to overcome anxiety and stress about having to speak (interpret) in public. Indeed, researchers have long considered the capacity to control anxiety and stress as an important requisite for a good interpreter (Cooper et al., 1982; Moser-Mercer, 1985; Longley, 1989; Klonowicz, 1994; Gile, 1995; Moser-Mercer et al., 1998) and a predictor of an interpreter's competence (Alexieva, 1997; Dong et al., 2013a). Some researchers have proposed to take the capacity to control anxiety and stress while interpreting into account in interpreting entrance exams (e.g., Moser-Mercer, 1985).

Empirical research has emphasized professional interpreters at work, focusing mainly on physiological responses to stress during interpreting: cardiovascular activity (Klonowicz, 1994), causes of anxiety and stress (Cooper et al., 1982), and chemical and physiological analysis (Moser-Mercer et al., 1998). There is now good evidence that interpreting leads to anxiety and stress for the interpreter (e.g., Cooper et al., 1982; Kurz, 1997, 2003). Thus, it is likely that interpreters, especially inexperienced ones like student interpreters, may experience a high level of anxiety when delivering target language, resulting in speech disfluencies (e.g., Cho and Roger, 2010). Indeed, people tend to stutter more when they are anxious (Craig, 1990; Menzies et al., 1999; Messenger et al., 2004). Because stuttering is an extreme example of disfluency, it is likely that anxiety may also lead to disfluencies in speech. Indeed, anxiety has long been associated with foreign language speaking (e.g., MacIntyre and Noels, 1996), leading to disfluencies in foreign language speech (Arnaiz and Pérez-Luzardo, 2014) and public speaking (Andrade and Williams, 2009).

Speech Disfluencies in Language Production and in Interpreting

Disfluencies are generally described as interruptions of the execution of a speech plan (Postma et al., 1990). As a form of language production, target language output is also filled with speech disfluencies. Some research has looked into speech disfluencies in simultaneous interpreting. Pöchhacker (1995) examined speech repairs (e.g., false starts, lexical blends, and syntactic blends) in conference source speech and corresponding simultaneous interpreting output (between English and German in both directions). More speech repairs were found in the target language output than in the source language output. There were more simple errors and false starts in the output of speakers, whereas in the output of interpreters, the most frequent disfluencies were lexical and structural blends.

Mead (2000) examined the control of pauses when students interpreted into their native or non-native language. Using Gósy's (2007) taxonomy, a series of papers were published on speech disfluencies in the output of simultaneous interpreters working in Hungarian.

Tissi (2000) attempted to come up with a simultaneous interpreting–specific taxonomy of disfluencies and at the same time stressed the communicative value and the strategic use of disfluencies in interpretation. She focused on silent pauses (the two subcategories being grammatical and/or communicative pauses and non-grammatical pauses) and disfluencies (including

TABLE 1 | Classification of disfluencies in interpreting, with examples.

Type of disfluency	Definition and example
Pause (DP)	A silence inside a clause E.g., *And companies like China Mobile… 同样的像中<DP>国移动这样的公司……*
Filler (DF)	The use of speech signals such as “uh,” “mm,” etc., to fill a pause E.g., *I need 100 million units… 需要呃……<DF>, 呃……<DF>一千亿的, 呃…….订购。*
Repetition (DRe)	The repetition of a single Chinese morpheme, a whole word or a phrase (in order to buy time for subsequent lexical access) E.g., *I watch all the time, students made perfectly beautiful programs 我曾经, 曾经<DRe>见过很多学生, 他们可以做出来很好, 的<DRe>他们可以做出来很好的*
Articulatory disfluency (DAr)	The stuttering of a morpheme within a word E.g., 1: *And the reason I'm focused on children is because… 程序。<DAr>, 我之……* E.g., 2: *We want to make… 之所以去进行这一个项目的原因<DAr>, 我<DAr>, 我……*
Other disfluency	Unidentified disfluencies that don't fit into the above categories

In the examples, the English text is the source language and the Chinese text is the target language. Letter strings in brackets (e.g., <DP>) are codes for different disfluency types and are used here to indicate the position of the disfluency.

fillers such as vocalized hesitations, vowel and consonant lengthenings, and interruptions such as repeats, restructuring and false starts). Tissi found large individual variations, and argued that no clear trends can be identified and that the influence of the source speech is not as direct as one would assume. She also found that vowel and consonant lengthenings are much more numerous in the target speech, and false starts occur only in the target speech. She also noticed the communicative, sometimes even strategic, use of disfluencies by the interpreter (e.g., silent or filled pauses before a correction), lengthenings of the tonic vowel, and retrospective repeats.

In this article, following the psycholinguistic literature (e.g., Postma et al., 1990; Fox Tree, 1993), I propose that interpreting disfluencies mainly include pauses, fillers, repetitions, and articulatory disfluencies, among others (see **Table 1** below)[1].

Speech is often disrupted by (silent) pauses and fillers (filled pauses). Pauses are a period of silence in the middle of an utterance, often caused by speech-planning problems. But in interpreting (as in conversation), a pause in speech may be ambiguous to the speaker and the audience, who may take it to signal the end of the interpreting. Thus, interpreters (as speakers) tend to have filled pauses (or fillers) during speech. Fillers specifically refer to *uh* and *um* (and equivalents in other languages), which are very common in speech production (e.g., Clark and Fox Tree, 2002), and have received much attention in recent years. According to Clark and Fox Tree, *uh* and *um* in English are signals that allow the speaker to keep the floor during conversation so that he/she will have more time for language planning (e.g., searching for words or framing the message). Thus, I also assume that interpreters use fillers strategically to hold the floor during interpreting, especially when there is a speech planning problem.

Repetitions occur when an interpreter or a speaker repeats a word or words without any grammatical or apparent semantic purposes (e.g., *she… she likes it*). In natural speech, speakers tend to repeat function words such as articles (e.g., *the, a*), prepositions (e.g., *of*), and auxiliaries (e.g., *do*) more often than content words such as nouns and verbs (e.g., Fox and Jasperson, 1995), probably because function words tend to begin a phrase (e.g., Clark and Wasow, 1998). Furthermore, another type of word that is often repeated is pronouns, especially when they begin a phrase (Clark and Wasow, 1998). For instance, it was shown that the possessive *her* (e.g., *her son*) was repeated more frequently than the accusative pronoun *her* (e.g., *love her*), despite the fact that they have the same form (Clark and Wasow, 1998).

Finally, speech can be disrupted when the speaker experiences articulatory disfluencies such as stuttering. Articulatory disfluencies can be seen in non-stuttering interpreters/speakers, often manifesting as difficulty producing a syllable in the middle of a word (e.g., *sec..secondary*). Articulatory disfluencies thus occur as a result of difficulties during speech programming rather than intentionally repeating a word (as a repetition).

The Current Study

The quality of interpreting depends on, among other things, two important criteria: accurate delivery of content in the source language and fluent delivery of the target language (e.g., Zhao and Dong, 2013). The former can be reflected in the likelihood of erroneous interpreting (see Zhao et al., 2021) and the latter can be reflected in the (dis)fluency of interpreting output. In this article, I focus on speech disfluencies in interpreting. In particular, I examine how different types of interpreting disfluencies relate to a student interpreter's cognitive traits, in particular, to a student interpreter's language proficiency, working memory span, and anxiety.

As I reviewed above, there is much evidence that language proficiency, working memory, and anxiety impact how well the interpreter conducts interpreting. As a specialized form of bilingual language processing, consecutive interpreting involves both the comprehension of a source language and the production of a target language. Therefore, it is critical that interpreters have sufficient proficiency in both the source and target language (Blasco Mayor, 2015). In addition, consecutive interpreting requires the storage of much source language information in working memory before it can be delivered in the target language; hence, working memory capacity is also shown to be critical in

[1] Arguably speech repair is also a form of disfluency. However, I did not include this in the current study because speech repair is better treated as a form of speech monitoring (e.g., Levelt, 1983) and is thus beyond the scope of the current study. Interested readers can refer to Zhao (2015).

interpreting performance (e.g., Christoffels et al., 2006). Finally, consecutive interpreting is a form of public speaking where interpreters convey a message to an audience, often in a formal setting; therefore, the capacity to control anxiety has traditionally been considered one of the requisites for interpreting (e.g., Moser-Mercer et al., 1998) and a predictor of interpreting competence (e.g., Alexieva, 1997).

To examine how language proficiency, working memory and anxiety may impact interpreting disfluencies, I conducted an experiment where 53 student interpreters consecutively interpreted an English speech into Chinese. I also measured their proficiency in English (the source language), working memory span in English listening, working memory span in Chinese speaking, and their general anxiety about public speaking. Interpreting output of the student interpreters was coded for disfluencies (see **Table 1** for a taxonomy of interpreting disfluencies). I then used regression analyses to examine the relationship between student interpreters' cognitive traits and interpreting disfluencies.

MATERIALS AND METHODS

Participants

Fifty-three fourth-year college students (45 females and 8 males; the imbalance of gender reflects female dominance in interpreting students in China) majoring in interpreting and translation participated in the consecutive interpreting test in a session of their interpreting module. These students all spoke Mandarin Chinese as their first language and had learned English as a second language since primary school. In addition, they all majored in English in college and had used English in both their courses and daily life. Thus, they were all unbalanced Chinese-English bilinguals who were proficient in English. All these participants trained in English language in the first 2 years of their university education and started to train in interpreting from the 3rd year onward (i.e., they had already had 1 year of interpreting training at the time they participated in this study).

The Language Proficiency Test

All 53 participants further took part in a language proficiency test. I developed our test on the basis of the Test for English Majors Band 8, which is a national official test of English language proficiency for English majors in the fourth BA year (such as our participants). As some of the test items were not relevant to language proficiency (e.g., test items on linguistics and English literature), I selected only test items that were related to proficiency in real language usage; these included the reading comprehension part, the listening comprehension part, and the writing composition part (see **Supplementary Appendix 1** for a description of the test items). The total score was 56. In the test, after test papers and answer sheets were distributed to the participants, they began the test with the listening comprehension part, followed by the reading comprehension part, and then by the writing composition part.

The Working Memory Test

The working memory test was adapted from the paradigm developed in Mizera (2006), in which participants memorized a list of Chinese words [e.g., 数学 (math), 现代 (modern), 面积 (area)] and then made a sentence for each word. The materials were 100 two-character Chinese words; all were high-frequency words according to the Modern Chinese Word Frequency Dictionary. There were 5 sets of test items, respectively, with 2, 3, 4, 5, and 6 memory words in a trial. There were 5 trials in each set, with a total of 25 trials. In each trial, participants first read the words one by one on a computer screen, with each word being presented for 1 s. After the presentations, a cue sentence appeared on the screen asking participants to make up a sentence for each of the words presented. Participants pressed the spacebar and made up the sentences. All responses were digitally recorded. Trials were randomly presented. There was a practice session with two trials, one with 2 memory words and one with 3. The score for the test was the proportion of words (out of 100) with which a grammatical sentence was composed.

The Anxiety Questionnaire

Note that anxiety in interpreting may be a multifaceted factor that consists of a student interpreter's general daily anxiety (e.g., when dealing with people and when doing a job) and his/her anxiety about interpreting (e.g., not being very good at English or having a poor memory). In order to exclude language-related and memory-related factors (which were covered by language proficiency and working memory tests already), I decided to use a scale developed in Zhang and Schwarzer (1995) and translated into Chinese by Dong et al. (2013b; see **Supplementary Appendix 2** for sample questions). The scale consisted of two parts. Part 1 tested self-efficacy anxiety (i.e., the anxiety one feels regarding whether he can do a particular task) and Part 2 tested state-trait anxiety (anxiety level as a personal characteristic). An answer was scored 1, 2, 3, or 4 points depending on the response, and a person's total score for anxiety was the sum of all the points in the 30 test items.

The Interpreting Test

The source language (English) speech was adapted from a real international conference speech on computer technology (see **Supplementary Appendix 3**). The original speech lasted for about 10 min, with a speech rate about 180 words per minute; such a speech rate is deemed to the most natural and pleasing speed for broadcasting (Boyd, 2003). The speech was delivered in a standard American accent. The speech was segmented to make it suitable for consecutive interpreting. In line with the common practice of the China Aptitude Test for Translators and Interpreters (CATTI) for consecutive interpreting Level II, following each segment of speech (2–5 sentences in length) was a pause that lasted for about 1.5 times the duration of the preceding segment, where student interpreters provided their interpreting.

The interpreting test was conducted by a teacher in a multimedia lab where participants had their interpreting classes. Participants sat in front of a computer with their headphones. The teacher gave verbal instructions regarding the interpreting test. In the test, participants heard the speech segment by segment, during which note-taking was allowed. At the end of each segment, participants heard an audio signal "ding" as a cue to start their interpreting. Participants' interpreting was individually recorded. The test lasted about 25 min.

I then invited two experienced professional interpreters to rate student interpreters' performance. Both raters had worked as professional consecutive and simultaneous interpreters for over 8 years and taught interpreting courses on BA and MA level at a university for 6 years by the time of rating. They took part in a rater training session on the rating scale before conducting the rating. They then rated two interpreting recordings (not part of the recordings in the current study) using the scale. For rating discrepancies, they discussed and reached a common ground. After this, they separately rated each student's interpreting according to the rating scale (with a full score of 100; Zhao and Dong, 2013). I computed an average score for each participant. Then the recordings of interpreting were transcribed. On the basis of the transcriptions, disfluencies were coded according to the taxonomy I reviewed above (see also **Table 1** for examples).

RESULTS

I first tested how interpreting score varied as a function of the three cognitive factors (see **Supplementary Material** for the data). Participants' interpreting scores increased as a function of their language proficiency ($\beta = 1.07$, SE = 0.33, $t = 3.4$, $p = 0.002$), increased as a function of their working memory ($\beta = 0.36$, SE = 0.10, $t = 3.40$, $p = 0.002$), and decreased as a function of their anxiety level ($\beta = -0.22$, SE = 0.08, $t = -2.66$, $p = 0.011$). I also found a significant correlation between a participant's interpreting score and their total disfluency rate ($r = -0.40$, $t = -3.11$, $p = 0.004$): participants who had a higher overall disfluency rate tended to do more poorly in their interpreting performance.

I next examined the occurrence of disfluencies and how they might be impacted by cognitive factors. In general, there were about 45 disfluencies out of 1,000 morphemes/characters in the target language output. Among the different disfluency types, the most common one is fillers, followed by repetitions. Pauses and articulatory disfluencies were rare (see **Table 2**). I conducted regression analyses on the rate of total disfluencies (i.e., number of disfluencies out of 1,000 characters in the output speech), using language proficiency, working memory and anxiety as predictors. As shown in **Table 3**; see also **Figure 1**, there is no significant effect. There is a marginally significant effect of working memory, with a trend of disfluencies decreasing as a function of working memory. There is a significant effect of anxiety, with increasing disfluencies as a function of participants' anxiety level.

Finally, I looked at how different types of disfluencies varied as a function of the three cognitive factors. Pauses did not vary as a function of any cognitive factor. Fillers did not vary as a function of language proficiency and working memory, but increased as a function of anxiety. Repetitions did not vary as a function of language proficiency and working memory, but increased as a function of anxiety. Finally, articulatory disfluencies did not vary as a function of any cognitive factor. It should, however, be noted that the occurrences of pauses and articulatory disfluencies were rare and the lack of cognitive influences on these disfluencies could be due to a floor effect.

TABLE 2 | Descriptive statistics of different disfluency rates (out of 1,000 characters in target output).

Type	Range	Mean	SD
Total disfluencies	8.6–144.4	45.3	26.6
Pauses	0–16.4	2.0	3.1
Fillers	0.4–128.4	34.5	24.2
Repetitions	0–28	7.4	6.5
Articulatory disfluencies	0–4.4	1.1	1.1

DISCUSSION

In this study, I explored how a student interpreter's cognitive traits, namely language proficiency, working memory and anxiety level, impacted speech disfluencies in target language delivery. Student interpreters were judged as worse in interpreting performance if they produced more disfluencies. Importantly, I showed that the occurrence of disfluencies is influenced by a student interpreter's anxiety level but not their language

TABLE 3 | Different types of disfluencies as a function of the cognitive factors (significant *p*-values in bold).

	Estimate	*SE*	*t*	*p*
Total disfluencies				
(Intercept)	50.62	60.40	0.84	0.406
Language proficiency	–0.53	1.12	–0.47	0.639
Working memory	–0.62	0.35	–1.75	0.086
Anxiety	0.92	0.28	3.25	**0.002**
Pauses				
(Intercept)	–2.48	8.13	–0.31	0.762
Language proficiency	0.03	0.15	0.17	0.866
Working memory	0.02	0.05	0.42	0.674
Anxiety	0.03	0.04	0.81	0.422
Fillers				
(Intercept)	37.66	57.49	0.66	0.516
Language proficiency	–0.23	1.06	–0.22	0.829
Working memory	–0.55	0.34	–1.63	0.109
Anxiety	0.70	0.27	2.60	**0.012**
Repetitions				
(Intercept)	12.74	15.88	0.80	0.426
Language proficiency	–0.28	0.29	–0.97	0.339
Working memory	–0.08	0.09	–0.83	0.413
Anxiety	0.17	0.07	2.27	**0.028**
Articulatory disfluencies				
(Intercept)	1.27	2.76	0.46	0.649
Language proficiency	–0.01	0.05	–0.14	0.887
Working memory	–0.01	0.02	–0.60	0.551
Anxiety	0.01	0.01	0.92	0.361

significant p-values in bold.

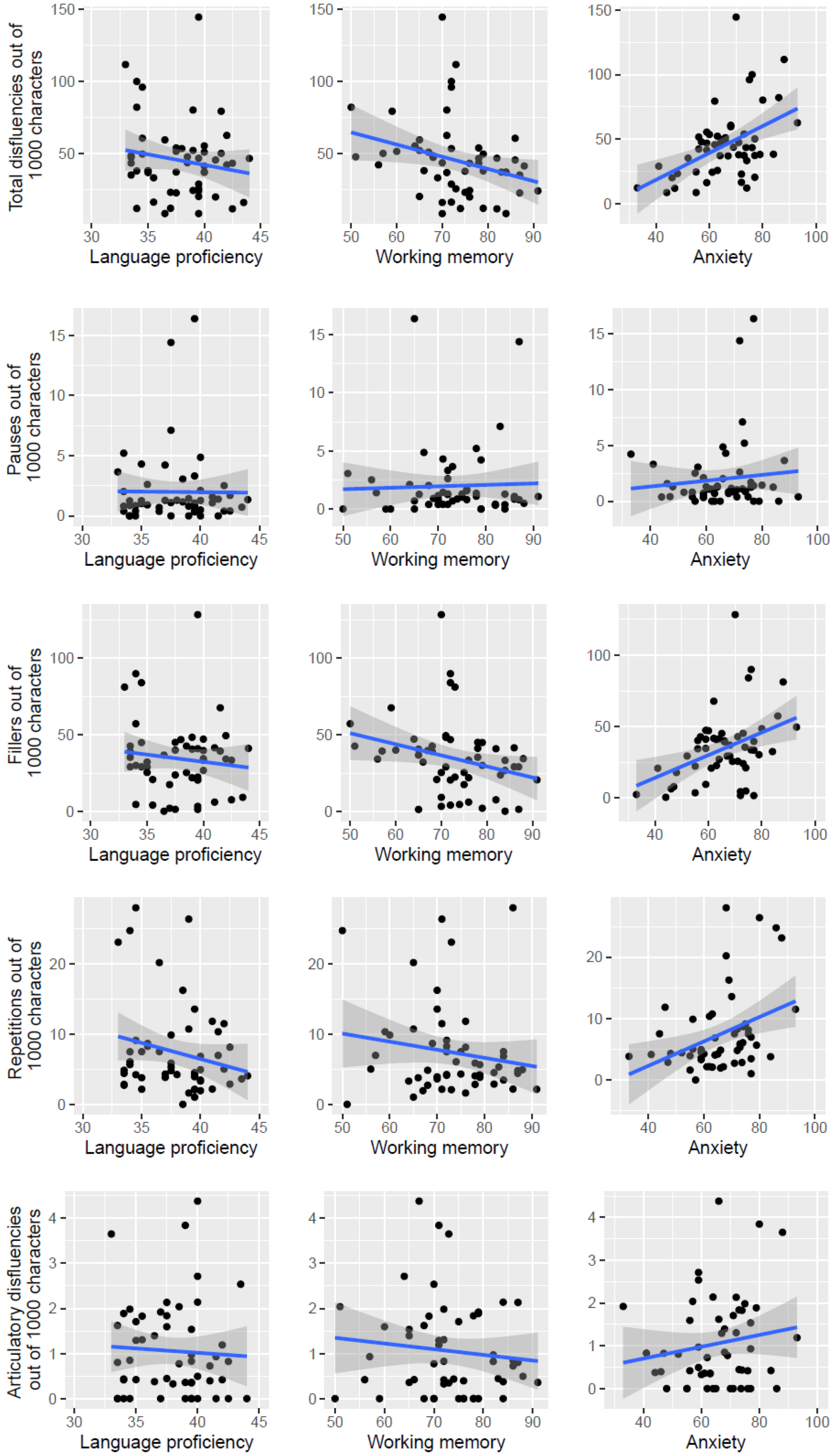


FIGURE 1 | Total disfluencies, pauses, fillers, repetitions, and articulatory disfluencies as a function of language proficiency, working memory, and anxiety.

proficiency or working memory. In particular, student interpreters with higher anxiety tended to have more fillers and more repetitions in their interpreting output. My findings of the anxiety effects are thus in line with previous theorizing that anxiety control is an important part of interpreting ability (Cooper et al., 1982; Moser-Mercer, 1985; Klonowicz, 1994; Gile, 1995; Alexieva, 1997; Moser-Mercer et al., 1998).

But how does anxiety affect the fluency of target language delivery in interpreting? According to the attentional control theory of anxiety (Eysenck et al., 2007), an influential theory that specifically addresses how anxiety affects cognitive performance (which includes interpreting), anxiety increases stimulus-driven attention (i.e., automatic attention to salient things, e.g., a loud sound) but decreases goal-driven attention (i.e., attention needed to complete a goal, e.g., interpreting a speech). More specifically, when an individual feels anxious, he/she attends more to salient properties in the surrounding environment; when the properties are not goal-related (e.g., a cough from the audience when an interpreter is working), the individual is easily distracted, thus leading to processing difficulties (e.g., at finding an appropriate translation word) and in turn to disfluencies (e.g., fillers).

According to the attentional control theory, an anxious individual is impaired in his/her cognitive functions that are necessary for completing a goal. These cognitive functions include inhibition, shifting, and updating (see also Miyake et al., 2000). Inhibition is a cognitive process whereby an individual is less likely to respond to things (e.g., responding to a goal-irrelevant in the audience during interpreting). Shifting is needed to allocate cognitive resources among different sub-tasks (e.g., listening to the source speech while retrieving target language expressions) when a cognitive performance requires multitasking. Finally, updating is a process that helps to update and monitor working memory representations (e.g., semantic representations retrieved from the comprehension of the source speech). Both attention and executive functions (e.g., inhibition and updating of information) are necessary in language processing and thus interpreting. For example, in language production such as interpreting, attentional resources are necessary to monitor whether a produced speech contains errors, and executive functions such as updating are necessary to integrate the message from a new sentence into the context to build a coherent model of the topic being comprehended. Thus, it is expected that an interpreter's anxiety has an all-round impact on interpreting (e.g., comprehension of source language, content delivery in the target language), not just disfluencies.

Given the crucial role of anxiety in interpreting, helping student interpreters to become less anxious (especially in public) should be an important component of the interpreting training curriculum. To do this, it is crucial that we understand the anxiety level of each student interpreter. We can then build an anxiety profile for each student by regularly testing their anxiety level (e.g., Dong et al., 2013b). For anxious students, more opportunities should be offered for them to speak in public.

In summary, we showed that speech disfluencies, especially fillers and repetitions, tended to increase as a function of a student interpreter's anxiety level. Given that interpretation is expected to be as fluent as possible, the finding suggests that interpreting teaching and learning should place more emphasis on reducing student interpreters' anxiety, especially in public speaking.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Guangdong University of Foreign Studies. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

The author would like to thank Yanping Dong for her supervision on the project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.881778/full#supplementary-material

REFERENCES

Alexieva, B. (1997). A typology of interpreter-mediated events. *Translator* 3, 153–174. doi: 10.1080/13556509.1997.10798996

Andrade, M., and Williams, K. (2009). Foreign language learning anxiety in Japanese EFL university classes: physical, emotional, expressive, and verbal reactions. *Sophia Jr. Coll. Fac. J.* 29, 1–24.

Arnaiz, P., and Pérez-Luzardo, J. (2014). Anxiety in Spanish EFL university lessons: causes, responsibility attribution and coping. *Stud. Angl. Posnaniensia* 49, 57–76. doi: 10.2478/stap-2014-0003

Belke, E. (2008). Effects of working memory load on lexical-semantic encoding in language production. *Psychon. Bull. Rev.* 15, 357–363. doi: 10.3758/pbr.15.2.357

Blasco Mayor, M. J. (2015). L2 proficiency as predictor of aptitude for interpreting: an empirical study. *Transl. Interpret. Stud.* 10, 108–132. doi: 10.1075/tis.10.1.06bla

Boyd, A. (2003). *Broadcast journalism: Techniques of radio and television news.* Oxford: Focal Press.

Cho, J., and Roger, P. (2010). Improving interpreting performance through theatrical training. *Interpret. Transl. Train.* 4, 151–171. doi: 10.1186/s13063-016-1291-7

Christoffels, I. K., de Groot, A., and Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: the role of expertise and language proficiency. *J. Mem. Lang.* 54, 324–345. doi: 10.1016/j.jml.2005.12.004

Christoffels, I. K., de Groot, A., and Waldorp, L. J. (2003). Basic skills in a complex task: a graphical model relating memory and lexical retrieval to simultaneous interpreting. *Biling. Lang. Cogn.* 6, 201–211. doi: 10.1017/s1366728903001135

Clark, H. H., and Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition* 84, 73–111. doi: 10.1016/s0010-0277(02)00017-3

Clark, H. H., and Wasow, T. (1998). Repeating words in spontaneous speech. *Cogn. Psychol.* 37, 201–242. doi: 10.1006/cogp.1998.0693

Cooper, C. L., Davies, R., and Tung, R. L. (1982). Interpreting stress: sources of job stress among conference interpreters. *Multilingua* 1, 97–108.

Craig, A. (1990). An investigation into the relationship between anxiety and stuttering. *J. Speech Hear. Disord.* 55, 290–294. doi: 10.1044/jshd.5502.290

Daneman, M., and Carpenter, P. A. (1980). Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19, 450–466.

Daneman, M., and Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *J. Exp. Psychol.* 9, 561–584. doi: 10.1037/0278-7393.9.4.561

Darò, V., and Fabbro, F. (1994). Verbal memory during simultaneous interpretation: effects of phonological interference. *Appl. Linguist.* 15, 365–381. doi: 10.1016/j.cortex.2013.07.012

Dong, Y., Cai, R., Zhao, N., and Lin, J. (2013a). An empirical study on interpreting competence structures in student interpreters. *Foreign Lang.* 36, 76–86.

Dong, Y., Chen, H., and Yu, Z. (2013b). 口译焦虑量表的研制. *Foreign Lang. World* 6, 57–64.

Eysenck, M. W., Derakshan, N., Santos, R., and Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion* 7:336. doi: 10.1037/1528-3542.7.2.336

Fox, B. A., and Jasperson, R. (1995). "A syntactic exploration of repair in English conversation," in *Alternative linguistics: Descriptive and theoretical modes*, ed. P. Davis (Amsterdam: John Benjamins Publishing), 77–134. doi: 10.1111/1460-6984.12178

Fox Tree, J. E. (1993). *Comprehension after speech disfluencies*. Ph.D. thesis. California: Stanford University

Gerver, D. (1975). A psychological approach to simultaneous interpretation. *Meta* 20, 119–128. doi: 10.7202/002885ar

Gile, D. (1995). Fidelity assessment in consecutive interpretation: an experiment. *Target* 7, 151–164. doi: 10.1075/target.7.1.12gil

Gósy, M. (2007). Disfluencies and self-monitoring. *Govor* 24, 91–110.

Hodáková, S. (2009). *Pamäť v simultánnom a konzekutívnom tlmočení. [Memory in simultaneous and consecutive interpreting.]*. Ph.D. thesis. Nitra: Konštantin Filozof University

Jiménez Ivars, A., Pinazo Catalayud, D., and Ruiz i Fores, M. (2014). Self-efficacy and language proficiency in interpreter trainees. *Interpret. Transl. Train.* 8, 167–182. doi: 10.1016/j.acap.2013.12.002

Klonowicz, T. (1994). "Putting one's heart into simultaneous interpretation," in *Bridging the gap: Empirical research in simultaneous interpretation*, eds S. Lambert and B. Moser-Mercer (Amsterdam: Benjamins), 213–224. doi: 10.1075/btl.3.16klo

Köpke, B., and Nespoulous, J.-L. (2006). Working memory performance in expert and novice interpreters. *Interpreting* 8, 1–23. doi: 10.1075/intp.8.1.02kop

Kurz, I. (1997). "Interpreters: stress and Situation-Dependent Control of Anxiety," in *Transferre Necesse Est. Proceedings of the 2nd International Conference on Current Trends in Studies of Translation and Interpreting*, eds K. Kinga and J. Kohn (Budapest: Scholastica), 201–206.

Kurz, I. (2003). Physiological Stress during Simultaneous Interpreting: a Comparison of Experts and Novices. *Interpret. Newsletter* 12, 51–67.

Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition* 14, 41–104.

Liu, M., Schallert, D. L., and Carroll, P. J. (2004). Working memory and expertise in simultaneous interpreting. *Interpreting* 6, 19–42. doi: 10.1075/intp.6.1.04liu

Longley, P. (1989). "The use of aptitude testing in the selection of students for conference interpretation training," in *The Theoretical and Practical Aspects of Teaching Conference Interpretation*, eds L. Gran and J. Dodds (Udine: Campanotto), 105–108.

MacIntyre, P. D., and Noels, K. A. (1996). Using social—psychological variables to predict the use of language learning strategies. *Foreign Lang. Ann.* 29, 373–386. doi: 10.1111/j.1944-9720.1996.tb01249.x

Martin, R. C., and Slevc, L. R. (2014). "Language production and working memory," in *The Oxford Handbook of Language Production*, eds M. Goldrick, V. Ferreira, and M. Miozzo (Oxford: Oxford University Press), 437–450.

Mead, P. (2000). Control of pauses by trainee interpreters in their A and B languages. *Interpret. Newsletter* 10, 89–102.

Menzies, R. G., Onslow, M., and Packman, A. (1999). Anxiety and StutteringExploring a Complex Relationship. *Am. J. Speech-Lang. Pathol.* 8, 3–10.

Messenger, M., Onslow, M., Packman, A., and Menzies, R. (2004). Social anxiety in stuttering: measuring negative social expectancies. *J. Fluen. Disord.* 29, 201–212. doi: 10.1016/j.jfludis.2004.06.002

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Mizera, G. J. (2006). *Working memory and L2 oral fluency*. Pennsylvania: University of Pittsburgh.

Moser, B. (1978). "Simultaneous interpretation: a hypothetical model and its practical application," in *Language interpretation and communication*, eds D. Gerver and H. W. Sinaiko (Boston, MA: Springer), 353–368. doi: 10.1097/EDE.0000000000000236

Moser-Mercer, B. (1985). Screening potential interpreters. *Meta* 30, 97–100. doi: 10.7202/003631ar

Moser-Mercer, B., Künzli, A., and Korac, M. (1998). Prolonged turns in interpreting: effects on quality, physiological and psychological stress (Pilot study). *Interpreting* 3, 47–64. doi: 10.1075/intp.3.1.03mos

Pöchhacker, F. (1995). "Those Who Do.": a Profile of Research (ers) in Interpreting. *Target* 7, 47–64. doi: 10.1007/s00330-019-06183-y

Postma, A., Kolk, H., and Povel, D. J. (1990). On the relation among speech errors, disfluencies, and self-repairs. *Lang. Speech* 33, 19–29.

Timarová, Š (2008). "Working memory and simultaneous interpreting," in *Translation and its Others: Selected Papers of the CETRA Research Seminar in Translation Studies 2007*, ed. P. Boulogne. Available online at: [https://lirias.kuleuven.be/retrieve/214722

Tissi, B. (2000). Silent pauses and disfluencies in simultaneous interpretation: a descriptive analysis. *Interpret. Newsletter* 10, 103–127.

Tzou, Y.-Z., Eslami, Z. R., Chen, H.-C., and Vaid, J. (2011). Effect of language proficiency and degree of formal training in simultaneous interpreting on working memory and interpreting performance: evidence from Mandarin–English speakers. *Int. J. Biling.* 16, 213–227. doi: 10.1177/1367006911403197

Van Dijk, R., Christoffels, I., Postma, A., and Hermans, D. (2012). The relation between the working memory skills of sign language interpreters and the quality of their interpretations. *Biling. Lang. Cogn.* 15, 340–350. doi: 10.1017/s1366728911000198

Waters, G., and Caplan, D. (2005). The relationship between age, processing speed, working memory capacity, and language comprehension. *Memory* 13, 403–413. doi: 10.1080/09658210344000459

Zhang, J. X., and Schwarzer, R. (1995). Measuring optimistic self-beliefs: a Chinese adaptation of the General Self-Efficacy Scale. *Psychologia* 38, 174–181.

Zhao, N. (2015). *Speech monitoring in consecutive interpreting: The effects of language proficiency, working memory and anxiety*. Ph.D. thesis. China: Guangdong University of Foreign Studies

Zhao, N., and Dong, Y.-P. (2013). A validating study using many-facet Rasch analysis on a consecutive interpreting. *J. PLA Univ. Foreign Lang.* 1, 86–90.

Zhao, N., Dong, Y.-P., and Cai, Z. G. (2021). *Speech errors in consecutive interpreting: Effects of language proficiency, working memory, and anxiety*. England: University of Essex.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINAL RESEARCH
published: 12 July 2022
doi: 10.3389/fpsyg.2022.872497

An Empirical Study on Imagery and Emotional Response in Chinese Poetry Translation—The Visual Grammar Perspective

Yuan Yuan[1] *and Tu Guoyuan*[2]*

[1] School of Foreign Languages, Central South University, Changsha, China, [2] Faculty of Foreign Languages, Ningbo University, Ningbo, China

The study investigated the evocation of mental imagery and emotional responses when English–Chinese bilinguals read classic Chinese poems and their English translations to examine (1) the target readers' formation of non-verbal text representations of Chinese poetry and (2) whether different translations affect the target readers' imagery cognition. A total of 20 English–Chinese speaker students enrolled in a Chinese university read a classic Chinese poem in Chinese and its four versions of translation in English. Through questionnaires and interviews, participants rated the visualized words used in the poems for the degree of mental imagery and emotional response evoked based on three indicators of narrative process, salience value, and emotive validity in the theoretical framework of visual grammar. Results showed considerable individual variances in the cognitive differences in forming mental imagery in all versions of the poems and there were also effects of translation strategy. Moreover, visual language information in poetry reading and its translations evoked different emotional responses depending on the use of visual words with cultural features. Our study demonstrates the applicability and accessibility of visual language in describing different readers' mental imagery and the interrelation and interaction between the poetry language system and the emotional, social, and cultural contexts involved in poetry translation.

Keywords: mental imagery, poetry translation, reader's cognition, visual grammar, visual image

Edited by:
Haoyun Zhang,
The Pennsylvania State University (PSU), United States

Reviewed by:
Zhenguang Cai,
The Chinese University of Hong Kong, Hong Kong SAR, China
Han Jingquan,
Guangdong University of Foreign Studies, China

Correspondence:
Tu Guoyuan
tuguoyuan@nbu.edu.cn

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *09 February 2022*
Accepted: *13 June 2022*
Published: *12 July 2022*

Citation:
Yuan Y and Guoyuan T (2022) An Empirical Study on Imagery and Emotional Response in Chinese Poetry Translation—The Visual Grammar Perspective. Front. Psychol. 13:872497. doi: 10.3389/fpsyg.2022.872497

INTRODUCTION

Mental imagery is a cognitive process that involves forming mental representations and images (Richardson, 1969; Kosslyn, 1994; Guarnera et al., 2019). The quasi-sensory and quasi-perceptual experiences in mental imagery can be formed without external stimuli that physically produce those sensory and perceptual experiences (Nanay, 2010). As the core of traditional Chinese aesthetics, mental imagery is a literary device widely applied in classic poetry in which vivid words are used to evoke an image or concept in the reader's mind. Through these words, poets seek to elicit emotional responses, rather than just painting a picture in the reader's mind, so that the imagery becomes a source of pleasure to readers (De Koning and van der Schoot, 2013). Empirical studies have examined the psychological and aesthetic responses to mental imagery during literary reading in both monolingual and bilingual contexts (Gunston-Parks, 1985; Goetz et al., 1993; Sadoski et al., 2000; Tierney and Readence, 2000; Sadoski, 2001). For instance, studies have shown the effects of imageability levels in mental imagery on single word processing in the form of both processing speed and brain activity (Steffensen et al., 1999; Ehlers-Zavala, 2005; Krasny and Sadoski, 2008; Connell and Lynott, 2012). Some evidence has also revealed contributions of different narrative styles to forming mental imagery elicited by a natural string of words (Magyari et al., 2020). However, translating imagery words from one language to another involves more complex cognitive operations than reading imagery words in either the native language or the second language. The cognitive operations underlying translating highly imageable and emotional poetry remain a topic unexplored.

In literary works, imagery is often considered a static visual description, i.e., a rendition of the story world wherein objects are ascribed visual properties but are separated from the characters' interactions with them (Wolf, 2004). However, Halliday (1967) argued that language is a social sign, "an atent system of meaning" and "the grammar of language" is not a set of systems but a resource for making meaning. Based on Halliday (1978)'s study on the social semiotic perspective, Kress and Van Leeuwen (1996, 2021) incorporated and extended the meta-functional theory (i.e., ideational function, interpersonal function, and textual function) by applying it to the level of visual modality to create a visual grammar, which interprets how images express various types of meanings (e.g., symbolic meaning, interactive meaning, and meaning of the composition). As Kress and Van Leeuwen (2021) put it, "grammar of the visual describes how depicted elements—people, places, and things—combine in the visual statements of greater or lesser complexity and extension".

In recent years, research in visual grammar has primarily developed in three perspectives: first, audience studies such as corpus studies or eye-movement experiments to confirm (or falsify) an argument (Holsanova, 2012; Bateman, 2014); second, theoretical development and innovation that integrate different disciplines, revise and supplement the visual grammar theory based on corpus analysis and propose new theoretical frameworks (e.g., Bateman, 2008; Painter et al., 2013); third, multi-modal discourse analysis through authentic images or pictures (Serafini, 2011; Feng and O'Halloran, 2012; Foncubierta-Rodríguez et al., 2017; Teng and Miao, 2018). Most multi-modal translation studies focus on non-verbal texts, but little attention has been paid to pure verbal literary texts. The current study attempts to employ visual grammar as a theoretical framework. It focuses on three critical indicators—narrative process (i.e., identification of narrative structures), salience value, and emotive validity—to examine readers' perceptions of imagery in poetry translation. Narrative structures refer to the structure which represents aspects of reality in terms of unfolding actions and events, processes of change, transitory spatial arrangements, and so on (Kress and Van Leeuwen, 2021). Narrative structures can generally be categorized into action process, reactional process, and speech and mental process. In the "action process," the participant who sends the vector is the actor, and the participant to whom the vector is directed is the goal. When the vector is formed by an eyeliner, by the direction of the glance of one or more of the participants, the structure is reactional. A special kind of vector in "speech and mental process" is formed by the thought bubbles and dialogue balloons in the image that connect drawings of speakers or thinkers to their speech or thought. Salience is one of the principles used to represent the compositional meaning of the image (Kress and Van Leeuwen, 1996). It refers to the elements (participants as well as representational and interactive system) that are made to attract the viewer's attention to different degrees, as realized by such factors as placement in the foreground or background, relative size, contrasts in tonal value (or color), the difference in sharpness, and so on (Kress and Van Leeuwen, 2021). Poetry writing has more opportunities to convey salience through the use of visual words and rhythmic features. The term "validity" encompasses what seem to be different types of truth that are realized in different semiotic modes, and expresses the social semiotic core idea that modality is based on the values, beliefs, and social needs of social groups. Abstraction and amplification of validity are described by abstraction and amplification of validity markers (Kress and Van Leeuwen, 2021). These markers can be various and abundant visual words that enhance the emotive attractiveness of the poems.

We also investigate how bilingual readers respond to and accept imagery by comparing their emotional reactions when reading the original and translated poems. Imagery in poetry is a real art sublimated from everyday life that begins with sensory contact and is subsequently processed and formed by a poet (Qu, 2002). A sinologist, Waley (2000), argues that imagery is the soul of poetry. Pound (2004), an American poet, believes that imagery can convey intellectual and emotional depth quickly in a poem or literary work. In literature, imagism emphasizes the union of emotion and form. According to Langer (1957), literature is a symbolic manifestation of human feelings, and the artistic symbol is the ultimate imagery full of passion, life, and personality. Imagery serves as the vehicle by which the poet expresses his or her emotional state, and the poet's rich emotional experience pervades the reproduction of imagery in literary translation. Therefore, the imagery reproduced must also be the reproduction of emotion, for only imagery with nearly the same emotional effect can faithfully convey the poet's thoughts and feelings while also reflecting the aesthetic value of the original poem in order to achieve its second life (Rojo et al., 2014). The purpose of the current study is to comment on poetry translation beyond the verbal system and highlight the importance of imagery creation, aesthetic experience, and cultural connotation in poetry translation.

Although the imagery word in poetry is difficult for readers to comprehend, particularly those with various cultural backgrounds, it has attracted universal attention and concern from poets and researchers. Cognitive and emotional investigation of mental imagery can help better understand how the text and the non-verbal system (mental imagery) work together in poetry translation, as well as how the source poem and the target poem are linked and reformed at the mental imagery level.

The poet, the translator, and the target reader are all involved in poetry translation. The perception and interpretation of the imagery vary as the participants' identities alter during the cognitive process of poetry translation. The poet creates imagery in the poem based on real-world objects and the poet's own experiences and emotions. The translator accepts and transforms

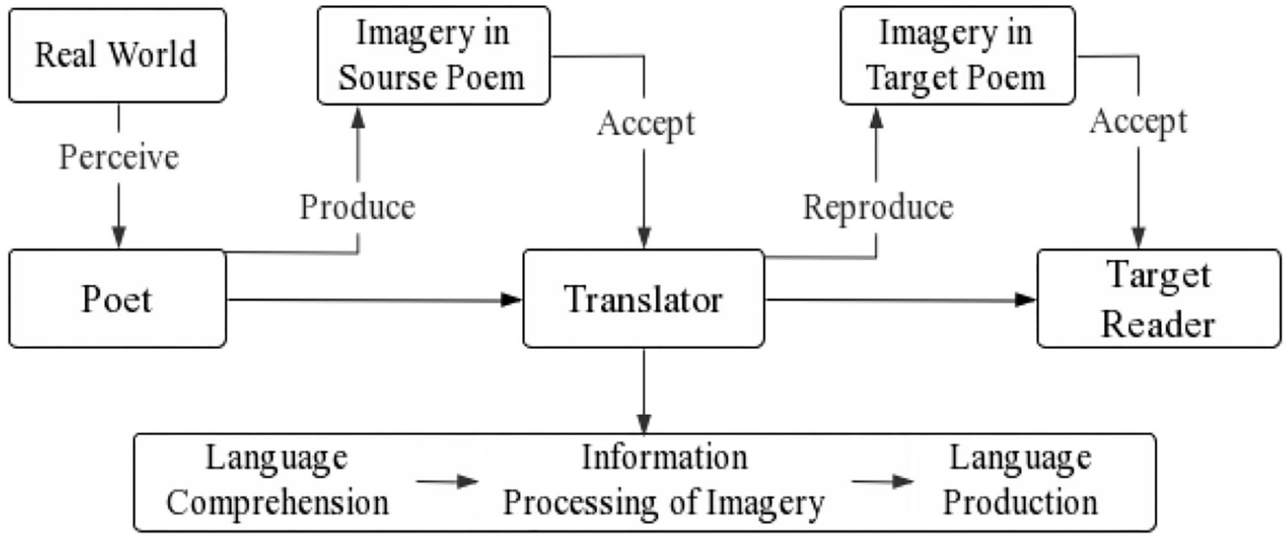


FIGURE 1 | The cognitive process of poetry imagery translation.

the imagery by reproducing it in the target language for the target reader (see **Figure 1**). The translator's cognitive process includes language comprehension, information processing of mental imagery, and language production, in which the translator has to determine the literal, emotional, and cultural meaning of the visual words in the source poem while perceiving the same aesthetic experience as the poet. Then the translator must search for equivalent words with the imagery effect in the target language for the target reader. The translator and the target reader might be thought of as recipients of the source and target poems. Therefore, the whole process of translating imagery contents in a poem involves the transition from visual symbols (words) to imagery, then from imagery to symbols (words in the target language), and finally from symbols to imagery on the part of the target reader.

METHODS

Subjects

Twenty participants (7 females, 13 males), including senior and post-graduate international students currently studying at a Chinese university, participated in the study. Twelve participants speak English as their native language, and eight are fluent in English (L1). All of them speak Chinese as their second language (L2), and 15 of them have passed Level V of the HSK (Chinese Proficiency Test).[1] All participants met or exceeded college curricular expectations for reading and writing in the Chinese language. Participation in the study was completely voluntary.

[1]The HSK (Level V) assesses test takers' abilities in the application of everyday Chinese. It is the counterpart of the Level V of the Chinese Language Proficiency Scales for Speakers of Other Languages and the C1 Level of the Common European Framework of Reference (CEF). Test takers who are able to pass the HSK (Level V) can read Chinese newspapers and magazines, enjoy Chinese films and plays and give a full-length speech in Chinese (see http://www.chinesetest.cn).

Materials

A classic Chinese poem and its four versions of translations were used as materials. The title of the poem is "天净沙·秋思" in Chinese. The poem, which contains 28 Chinese characters, was written by Ma Zhiyuan in the Yuan Dynasty (1271–1368). It sketches a picture of a journey at dusk in an autumn field, enumerating the spatial imagery of nine scenes in the landscape of ancient Chinese villages. There are eight noun phrases with the "modifier-head" structure, divided into three groups, each composed of three visual words. On the surface, there appears to be no link between the nine scenes, but the context of the poem reveals that all of the scenes are utilized to portray a profound and emotional love for one's birthplace. The four versions of English translations are titled "Autumn", "Tune: Tian Jin Sha", "Tune to "Sand and Sky"-Autumn Thoughts", and "Autumn Thoughts" by translators Weng Xianliang, a Chinese translator (Weng, 1978), Wayne Schlepp, a sinologist (Wen, 1989), Zhao Zhentao, a Chinese translator (Gu, 1993), and Ding Zuxin, a Chinese translator, and Burton Raffel, an American writer, translator (Ding and Raffel, 1992), respectively.

Procedure

All participants were first familiarized with the concept of visual grammar and its three critical indicators, "narrative structure", "salience value", and "emotive validity" in the theoretical framework. Each participant was then given five copies of the poem, one original Chinese version and its four versions of English translations. Participants were asked to read the materials and filled a questionnaire regarding the observed visual images presented by the words used in the materials. An experimenter then conducted an interview with each participant to collect explanations and reasons for the participant's answers to the questionnaire and the cognitive process behind it. According to the previous empirical study on mental imagery in literary translation and visual grammar (Chen and Li, 2021), the questionnaire was divided into four parts. In the first part, participants needed to divide the poem (and its four versions of translations) into pauses (i.e., sections) by finding the visual words and forming imagery. In the second part, participants needed to describe the type of defined mental imagery using the narrative structure categorization. In the third part, the subjects needed to determine the salient elements they find at first sight and describe the ways to achieve the imagery's meaning based on the content of the salient values. Finally, participants had to report any mental imagery that evoked emotional reactions. At the same time, they read the poem and explain the connections between the imagery and emotion. During the survey questionnaire, participants could look up information about the visual words to ensure their choices.

RESULTS

The Identification of Visual Images

After reading the original (source) poem (i.e., ST) and its four versions of English translations (i.e., TT1, TT2, TT3, and TT4), 15 participants identified 12 separate images in the source poem, TT2, and TT3. Sixteen reported 13 images in TT1, involving one visual word, "the far bank" which does not exist in the source poem. For TT4, sixteen participants identified the same 12 images as ST, but four participants identified more than 12 images. These four participants explained that "returning crows" and "at dusk" formed two separate images in their mental imagery and the same as "a narrow bridge" and "below the bridge."

However, single visual pictures must be connected to disclose the poem's idea and produce an emotional and aesthetic experience on the reader's part. For example, the same linguistic construct (i.e., a modified and a noun) was used to communicate three images in the first three lines of the source verse, which might be connected to make more significant pictures. Participants were asked to combine those images in each line and report whether or not they could "see" the pictures or determine the representational meaning of imagery superimposition. As can be seen in **Figure 1**, the connection of the three images in Line 1 is the weakest among the three lines. Only six participants suggested that they could think of pictures about the season or the desert, which are the target meaning of the verse. The remaining 14 participants could not even find any connections between the images. Nevertheless, both Line 2 and 3 produced imaginative pictures, especially in TT1 and TT4. Much like linguistic structures, visual structures point to particular interpretations of experience and particular forms of social interaction (Kress and Van Leeuwen, 1996). Without cultural accumulation, the effect of the imagery superimposition of parallel nouns (such as withered vine, old tree, and crow at dusk in Line 1 of ST and TT3) could not be easily achieved. However, in target poems, it is possible to find how these images can be linked to each other. Some are linked in spatial, locative terms, such as Line 2, "Yonder is a tiny bridge over a sparkling stream, and on the far bank, a pretty little village" (TT1), and "A few houses hidden past a narrow bridge, and below the bridge a quiet creek running" (TT4). All participants suggested that these words evoked imagery of pastoral life in their minds. Verbs can relate to others, such as Line 3, "west wind moaning, horse groaning" (TT1), and "A lean horse comes plodding" (TT4), which formed imagery of a very burdened, tired horse.

Recognition of imagery superimposition in Line 1, Line 2, and Line 3 in the source poem and its four versions of English translations, respectively, can be seen in **Figure 2**. Columns represent the number of participants who formed connections between mental imageries depicted in each line of each version of the poem.

Narrative Structure of Imagery

The current study selected the Chinese verb "下", which means "come" or "go down" in English, as the critical visual word

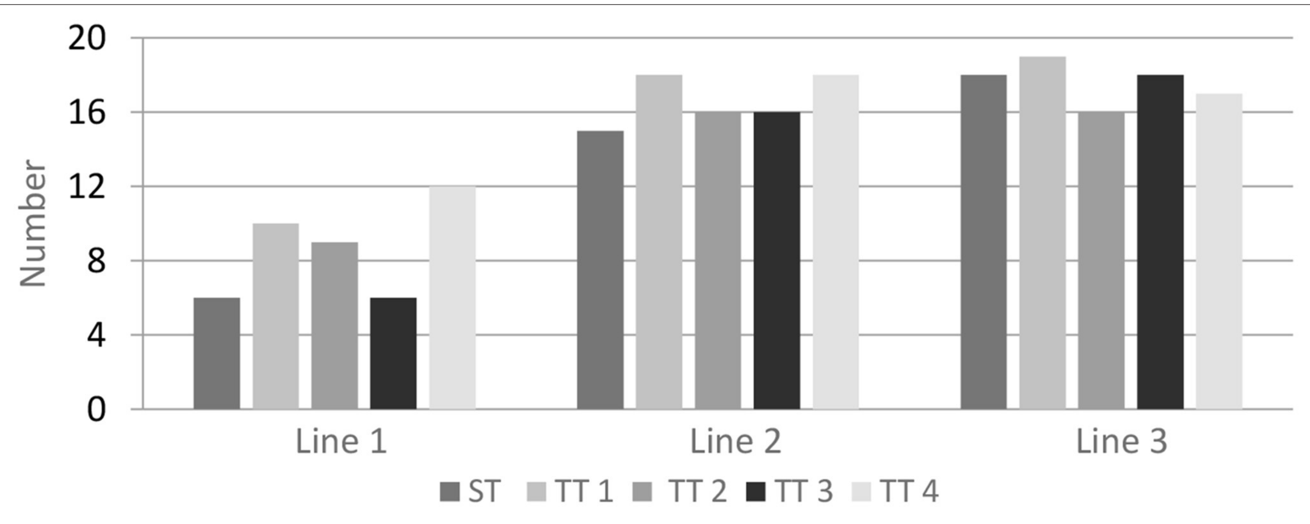


FIGURE 2 | Recognition of imagery superimposition.

TABLE 1 | Imagery comparison in verbal and non-verbal systems.

Poem	Text (verbal system) transitivity process type	Imagery (non-verbal system) narrative structure
ST: 夕阳西下，断肠人在天涯	Material process (20)	Action process (16)/reactional process (4)
TT1: Trudging toward the ***sinking*** sun, farther and farther away from home.	Material process (20)	Action process (10)/reactional process (10)
TT2: The sun ***westering***, And one with breaking heart at the sky's edge.	Material process (20)	Action process (11)/reactional process (9)
TT3: The sun is ***setting***, Broken man far from home roams and roams	Material process (20)	Action process (15)/reactional process (5)
TT4: The sun ***dips down*** in the west, and the lovesick traveler is still at the end of the world.	Material process (20)	Action process (9)/reactional process (11)

to examine participants' narrative structure of imagery. Both a verbal (the transitivity system; Halliday, 1994) and a non-verbal system (i.e., the visual narrative system; Chen and Li, 2021) are used to categorize participants' responses to the source poem and its four translations. Participants also judged if any forms of imagery were represented in the narrative structures.

As can be seen in **Table 1**, in the verbal system, all participants judged that the verb "下" is used as the material process (e.g., a process of doing or happening) in all clauses in the poem, as the actor is the "sun", and in the English translations, verbs including "sinking", "westering", "setting", and "dips down", as different translations of the Chinese verb "下", are processes of abstract doing pertinent to the "sun". However, to represent meaning through imagery, sixteen participants suggested that they could see the dynamic movement of the sunset in the source poem and thought that the day was about to finish (i.e., Action process). Four participants felt more assertive about the representation of the sun instead of the sunset (Reactional process). These variations between individual participants existed in all four versions of the English translations of the source poem. Furthermore, ten participants mentioned they could find the related process in the verbal system of TT1 due to the adding verb "trudging toward". They argued that the choice of this verb phrase was ineffective in translation, for it reduced the dynamic effect of the sun and could not enhance the logic-semantic relation in the poem.

Salience of Imagery

Participants were asked to determine the salience of the visual words from both the ST and TTs and indicate how they were achieved. Multiple explanations were allowed in their answers.

As can be seen in **Table 2**, the salience values observed from the visual words representing imagery in the poems varied considerably between participants. In the ST, the ways to achieve the representational imagery of "夕阳" (sunset in English) are the largest. Participants reported that they could imagine the position in the west, the enlarging size, and even the golden color of the sun, enjoying its warmth and beauty, or feeling the passing of the day. But in the TTs, most participants could only feel the setting sun in position, for all the verbs showed a dynamic movement of the sun. Then, the majority of the participants suggested that the most significant visual element in ST was "断肠人" (a heart-broken man), and most of them chose "cultural factor" as its way to achieve its function and considered this element to be the embodiment of traditional Chinese culture. Because "断肠人" is a frequently used literary expression in Chinese poems, often to describe the overwhelming emotional stimulation with the feeling of extreme sadness. Lack of a culture-based understanding of this expression, readers have to spend more cognitive effort when forming imagery of it. Nine participants formed imagery with an unfortunate person from its literal meaning "break one's intestine" (getting people hurt).

TABLE 2 | Distribution of imagery salience of subjects.

Text	Salience	Ways to achieve
ST	夕阳 (8)	Cultural factor (8) position (6) size (1) color (2)
	断肠人 (8)	Cultural factor (19)
	天涯 (4)	Cultural factor (14) position (4)
TT1	Trudging toward (5)	Position (2)
	Sinking sun (9)	Cultural factor (8) position (7) size (1)
	Farther and farther (6)	Position (19)
TT2	Sun westering (7)	Position (18)
	Breaking heart (9)	Cultural factor (19)
	The sky's edge (4)	Position (11) cultural factor (5) size (1)
TT3	Sun is setting (5)	Position (13) size (4)
	Broken man (13)	Cultural factor (18)
	Roams and roams (2)	Position (17)
TT4	Sun dips down (6)	Position (18)
	Lovesick traveler (9)	Cultural factor (2) position (3)
	End of the world (5)	Cultural factor (14) position (5)

As compared to ST, TTs correspond to a more significant variation in readers' perceptions of their mental imagery in terms of salience values. People or objects often generate the most considerable imagery elements, but some readers tend to "see" the action details referring to the narrative process. For example, 18 participants perceived the imagery salience of TT2 as "westering" which is a more dynamic expression of a pleasant or sad state of mind than the sunset action itself. Interestingly, in TT 4, nine participants gave salience to "lovesick traveler," but only two chose "cultural factor," and three chose "position." As they could not find any relevance of this expression to the poem's theme, it represented the wrong imagery compared to other visual words in the poem. The same case happened in TT1's "trudging toward" which misled participants more to focus on the action of people instead of the sunset imagery. Regardless of the degree to which different visual words are connected to each other, salience value can create a hierarchy of importance among elements, selecting some as more important, more worthy of attention as compared to others (Kress and Van Leeuwen, 2021). As a principal function of sound in the poem, the rhyme will also give different salience to strengthen the effect of imagery and the sense of beauty. After reading all the poems aloud, fourteen participants reported that they had successive sensations of salience about the stressed syllables and the regular rhyme in ST and TT3, which are more poetic. Because the rhythmic features of poems are not the focus of this paper, these data were not analyzed further.

Emotive Validity From Imagery

Participants were asked to choose the theme emotion of the poem after reading ST and four translations (multiple choices were available). Most participants could feel the nostalgia and the sadness in the poem, but it was still hard for the five of them, who knew less about Chinese poetry, to construct a complete picture in their minds, such as the sunset scenery in autumn. Nine participants found calmness, positioning themselves for the poem's second line, which describes the peaceful life through three images: bridge, stream, and homes. Participants also expressed their emotional responses to each salient element from **Table 2**, and the types of emotions all came from the genuine feelings that may arise from representing the meanings of single imagery (see **Table 3**).

As seen in **Table 3**, most participants chose "sorrow" for the miserable traveler in ST, TT2, and TT3, which corresponds to the poem's theme. Nevertheless, 2 of them chose "sorrow" for "lovesick traveler"; they suggested that the traveler was not lovesick but homesick, indicating that mistranslated imagery words cannot evoke the same emotional response. Moreover, words with Chinese cultural characteristics, such as "天涯" in ST, are also confused by readers who are not so familiar with Chinese literary works. Thus, the imagery from "at the sky's edge" and "at the end of the world" brings readers different emotional experiences. The verbs as salient elements in ST and TTs failed to deepen readers' emotional experience. As shown in **Figure 2**, the emotive validity is based on the affective appeal of the salient imagery words, and the imagery words in translated poems do not always convey the same emotive validity as the original poem. The predominant emotional tone of the original is sorrow. By adding up all the numbers of sorrow responses in **Figure 3**, ST almost realizes the emotive validity according to its background. However, the validity of such a feeling has decreased a lot in TT1, TT3, and TT4. The second emotional factor of a poem is loneliness, which is represented similarly in the four English versions of the poem, and TT2 almost achieves its emotive validity. Finally, calmness, which is not apparent in the original poem, is reflected in participants' reports to varying degrees in the four translated versions, which may violate the validity of the original affection.

TABLE 3 | Emotional response of subjects to the salient imagery.

Text	Salience	Emotional response
ST	夕阳 (8)	Joy (6) interest (5) sorrow (6)
	断肠人 (8)	Sorrow (14) horror (3)
	天涯 (4)	Loneliness (12) sorrow (8)
TT1	Trudging toward (5)	Tired (12) sorrow (6)
	Sinking sun (9)	Calmness (12) sympathy (7)
	Farther and farther (6)	Anxious (5) loneliness (6) sorrow (7)
TT2	Sun westering (7)	Calmness (16) sympathy (4)
	Breaking heart (9)	Sorrow (18)
	The sky's edge (4)	Loneliness (11) sorrow (8)
TT3	Sun is setting (5)	Calmness (16) sympathy (2)
	Broken man (13)	Sorrow (15) confusion (3)
	Roams and roams (2)	Loneliness (5) boredom (8) anxious (6)
TT4	Sun dips down (6)	Calmness (8) sympathy (10)
	Lovesick traveler (9)	Confusion (16) sympathy (2) sorrow (2)
	End of the world (5)	Loneliness (8) hopelessness (9) sorrow (3)

DISCUSSION

The evocation of mental imagery and emotions can assist in understanding and appreciating poetry in different cultural backgrounds. The recognition that target readers access both verbal and non-verbal imagery when reading poems is essential in selecting translation strategies. These strategies can exploit non-verbal aspects (e.g., imagery) of comprehension with target readers whose vocabulary, linguistic, and cultural knowledge of the poem's native language is rather limited. From a psychological point of view, imagery transformation is the process of selecting and matching the imagery to languages. In addition to the multiple influences of the translator's individual variables, such as knowledge, experience, and cognitive style, the conversion is also subject to external factors, such as social, cultural, and thinking traditions. Compared with the aesthetic imagery presented in the original poem, what is restored and reproduced by the translator may be the same, similar, or even wholly different "imagery."

The findings of the current study showed that the cognitive transformation of imagery contains considerable individual variances. Therefore, the ultimate goal of poetry translation is to achieve imagery-to-imagery equivalence, which refers to the fact that the imagery created has the same semantic and aesthetic implications in the translated versions as in the original language, based on the commonality of thinking and psychological identity. Although different cultures have distinct languages and ways of thinking, there are similarities in understanding the same thing and the translator's style of thinking, processing, and aesthetic experience. As a result, the "meaning" and "imagery" can be connected in the original as in translated languages. At this point, the translator will utilize more image-based reasoning to extract pictures of particular items in the original language, as well as the visual implications that go along with them. Translators will also search the translation system for objective items that fit the meaning and emotions so that they can link the two.

For the imagery in translation works to have the same or similar emotional effect like the one in the original, "image for image" strategy can also be used. This strategy refers to searching for different imagery in the target culture to replace

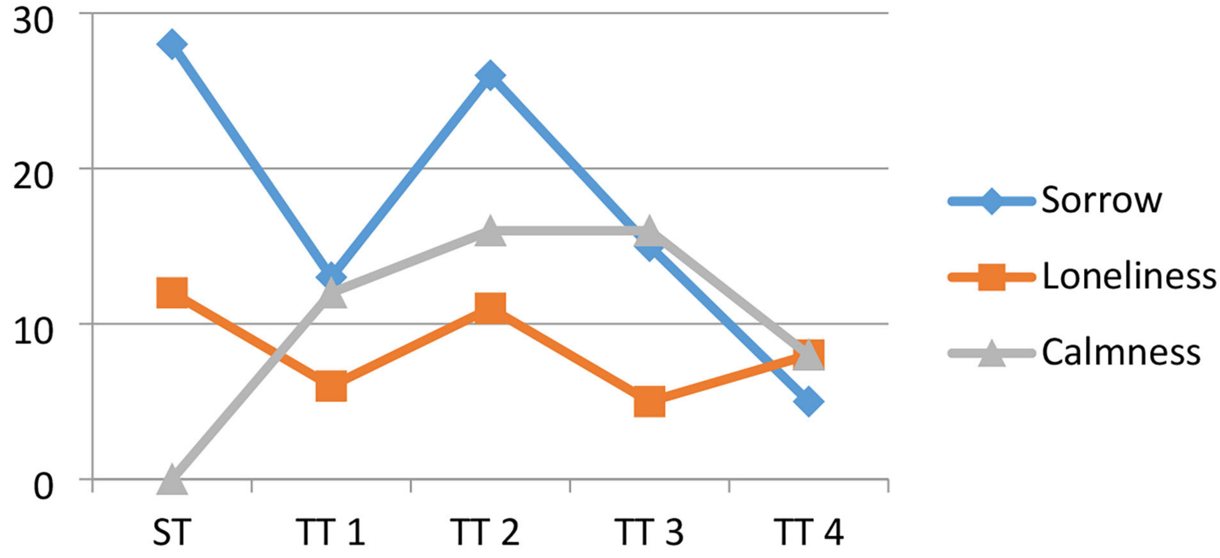


FIGURE 3 | The number of the main emotions in different versions of the poem.

the original one with the translator's imagination. Due to external environmental factors, the materials or information available for people's imagination might be different. That is to say, the equivalent image in the translated context cannot be found to represent the "image" in the original poem to a perfect level. However, since human beings have a high degree of consistency in their understanding of objective things in nature, the content of thinking (the "meaning" in imagery) will stay the same or comparable even if the output of thinking (language) changes in imagery reduction (Zoltan, 2000).

Furthermore, our results suggest that extracting the translated picture with the same visual impact from memory is not always possible. To be "faithful" to the original poem or let the target readers understand the exotic flavor, the translator needs to retain or highlight the imagery of the original work. The semantic or formal equivalence between the "original image" and the "translated image" may be obtained through this process, and the translator must make the meaning of the image explicit by adding notes or supplementary explanations to convey the original imagery's meaning and emotion. Finally, sometimes it is necessary to discard the image to reach the meaning, which means that the translator can ignore images developed in the original poem entirely and only look for the visual word with a similar meaning or effect. This strategy prevents the tautologies of adding long explanations or notes for imagery words in the original poem. This strategy also releases the translator from the burden of looking for expressions in the target language to restore and reproduce the original imagery. The translator may choose to abandon the imagery words in the transformation process based on his aesthetic judgment.

Our findings suggest that the theoretical framework of visual narrative provides a new perspective for analyzing poetry translation. First, we can comprehensively and systematically compare different visual narratives between the original and translated poems. Second, we can explain in more depth how visual imagery regulates readers' emotional engagement and attracts their interests by selecting visual words with almost the same imagery effect and providing theoretical guidance and practical strategies for poetry translation. The concepts of "narrative process," "salience value," and "emotive validity" help examine the readers' representational imagery from reading both the source poem and the translated poems. It can be found that there are interpersonal differences in recognizing imagery through visual words and between parallel poems involved in translation because meanings arise in social contexts and personal interaction, which are variables of cultural backgrounds (Kress and Van Leeuwen, 2021).

Imagery in poetry is a vivid and vibrant form of description that produces visual effects and appeals to the readers' senses and imagination (Lewis, 1984). Many translation strategies are studied from culture, linguistics, and translation studies in order to enhance the reproduction of imagery that is presented in target poems. However, some cognitive differences can still be described concretely through the principles of visual grammar. Thus, it is also confirmed that visual grammar can effectively serve as a reference mechanism for a particular reader to construct imagery meaning and evaluate different translations or present different readers' perceptions of the same poem. The poem selected for the current, "天净沙·秋思", is a classic work of Chinese poetry, and research on its English translations has been conducted extensively. This study attempts to describe the readers' imagery in poetry translation using visual language, peeking into the involvement of verbal and non-verbal systems in translation. Accordingly, it is argued that more intrinsic connections and interactions between verbal and non-verbal systems can be explored in translation studies in conjunction with examining mental imagery.

Imagery in poetry is uncertain in different cultures; it is not as solid and constant as real images, and they vary from one individual to another. The purpose of observing visual words and imagery is not to solve the problem of uniformity but rather to investigate them in-depth as the subjectivity of the cognitive process of translation triggers further studies. Reading poems with high imagery content can be seen as a visualization of words in the reader's mind. Visual grammar offers a new method to explore the interaction between the verbal and the visual, the semiotic and the non-semiotic, and between individual expression and social semiosis.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

YY and TG contributed to the conception of the study. YY conducted the experiments and drafted the manuscript. TG contributed to the revision of the manuscript. All authors have approved the final version of the manuscript.

FUNDING

This research was supported by National Social Science Foundation of China (16BYY008).

REFERENCES

Bateman, J. A. (2008). *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. New York: Springer.

Bateman, J. A. (2014). *Text and Image: A Critical Introduction to the Visual/Verbal Divide*. London: Routledge.

Chen, W. H., and Li, B. (2021). Readers' visual mental perceptions of literary translations: an empirical study based on visual grammar. *New Perspect. Trans. Stud.* 1, 22–33.

Connell, L., and Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition* 125, 452–465. doi: 10.1016/j.cognition.2012.07.010

De Koning, B. B., and van der Schoot, M. (2013). Becoming part of the story! Refueling the interest in visualization strategies for reading comprehension. *Educ. Psychol. Rev.* 25, 261–287. doi: 10.1007/s10648-013-9222-6

Ding, Z. X., and Raffel, B. (1992). *Gems of Chinese Poetry: From The Book of Songs to the Present*. Shenyang: Liaoning University Press.

Ehlers-Zavala, F. (2005). "Bilingual reading from a dual coding perspective," in *ISB4: Proceedings of the 4th International Symposium on Bilingualism*, 656–662.

Feng, D., and O'Halloran, K. (2012). Representing emotion in visual images: a social semiotic approach. *J. Pragmat.* 44, 2067–2084. doi: 10.1016/j.pragma.2012.10.003

Foncubierta-Rodríguez, A., Müller, H., and Depeursinge, A. (2017). From visual words to a visual grammar: using language modelling for image classification. *arXiv:1703.05571 [cs.CV]*, 1–17. doi: 10.48550/arXiv.1703.05571

Goetz, E. T., Sadoski, M., Stowe, M. L., Fetsco, T. G., and Kemp, S. G. (1993). Imagery and emotional response in reading literary text: quantitative and qualitative analyses. *Poetics* 22, 35–49. doi: 10.1016/0304-422X(93)90 019-D

Gu, Y. L. (1993). Appreciation of Ma Zhiyuan's autumn thoughts. *J. Foreign Lang.* 2, 12–14.

Guarnera, M., Pellerone, M., Commodari, E., Valenti, G. D., and Buccheri, S. L. (2019). Mental images and school learning: a longitudinal study on children. *Front. Psychol.* 10, 2034. doi: 10.3389/fpsyg.2019.02034

Gunston-Parks, C. A. (1985). *Effects of Imagery Ability and Text Bound Guided Imagery on Comprehension by Low Readers*. (Unpublished Doctoral Dissertation), Arizona State University, USA.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English. *J. Ling.* 1, 37–81. doi: 10.1017/S0022226700012949

Halliday, M. A. K. (1978). *Language as Social Semiotic: The Social Interpretation of Language and Meaning.* London: Hodder Education.

Halliday, M. A. K. (1994). "Spoken and written modes of meaning," in *Media Texts, Authors and Readers: A Reader*, eds D. Graddol and O. Boyd-Barrett (Bristol: Multilingual Matters).

Holsanova, J. (2012). New methods for studying visual communication and multimodal integration. *Vis. Commun.* 11, 251–257. doi: 10.1177/1470412912446558

Kosslyn, S. M. (1994). *Image and Brain: The Resolution of the Imagery Debate*. Cambridge: M.I.T. Press.

Krasny, K. A., and Sadoski, M. (2008). Mental imagery and affect in English/French bilingual readers: a cross-linguistic perspective. *Can. Modern Lang. Rev.* 64, 399–428. doi: 10.3138/cmlr.64.3.399

Kress, G., and Van Leeuwen, T. (1996). *Reading Images: The Grammar of Visual Design.* London: Routledge.

Kress, G., and Van Leeuwen, T. (2021). *Reading Images: The Grammar of Visual Design.* London: Routledge.

Langer, S.K. (1957). *Problems of Art.* New York, NY: Charles Scribner's Sons.

Lewis, C. D. (1984). *The Poetic Image: The Creative Power and the Visual Word.* Los Angeles: Jeremy P. Tarcher.

Magyari, L., Mangen, A., Kuzmičov,á, A., Jacobs, A. M., and Lüdtke, J. (2020). Eye movements and mental imagery during reading of literary texts with different narrative styles. *J. Eye Mov. Res.* 13, 3. doi: 10.16910/jemr.13.3.3

Nanay, B. (2010). Perception and imagination: A modal perception as mental imagery. *Philos. Stud.* 150, 239–254. doi: 10.1007/s11098-009-9407-5

Painter, C., Martin, J. R., and Unsworth, L. (2013). *Reading Visual Narratives: Image Analysis of Children's Picture Books*. London: Equinox.

Pound, E. (2004). "A retrospect," in *Poetry in Theory: An Anthology 1900-2000*, ed J. Cook (Oxford: Blackwell Publishing).

Qu, G. (2002). *On the Imagery in Classical Chinese Poetry*. Beijing: Chinese Social Science Today, 162–171.

Richardson, A. (1969). *Mental Imagery*. New York: Springer.

Rojo, A., Ramos, M., and Valenzuela, J. (2014). The emotional impact of translation: a heart rate study. *J. Pragmat.* 71, 31–44. doi: 10.1016/j.pragma.2014.07.006

Sadoski, M. (2001). Resolving the effects of concreteness on interest, comprehension, and learning important ideas from text. *Educ. Psychol. Rev*. 3, 263–281. doi: 10.1023/A:1016675822931

Sadoski, M., Goetz, E. T., and Rodriguez, M. (2000). Engaging texts: effects of concreteness on comprehensibility, interest, and recall in four text types. *J. Educ. Psychol.* 1, 85–95. doi: 10.1037/0022-0663.92.1.85

Serafini, F. (2011), Expanding perspectives for comprehending visual images in multi-modal texts. *J. Adolesc. Adult Liter.* 54, 342–350. doi: 10.1598/JAAL. 54.5.4

Steffensen, M. S., Goetz, E. T., and Cheng, X. (1999). The images and emotions of bilingual Chinese readers: a dual coding analysis. *Read. Psychol.* 20, 301–324. doi: 10.1080/027027199278376

Teng, D., and Miao, X.W. (2018). Meaning construction of multi-modal metaphors in the picture book discourse from the grammar of visual design. *Foreign Lang. Res.* 5, 53–59. doi: 10.16263/j.cnki.23-1071/h.2018. 05.010

Tierney, R. J., and Readence, J. E. (2000). Expository reading-writing think-sheets. *RJ Tierney JE Readence Read. Strategies Pract.* 5, 278–284.

Waley, A. (2000). *Chinese Poem*. New York: Dover Publications Inc.

Wen, S. (1989). *The Selection of English Translation of Chinese Poetry*. Beijing: Foreign Language Teaching and Research Press.

Weng, X. L. (1978). The anthology of ancient poetry translation. *Modern Foreign Lang.* 4, 88.

Wolf, W. (2004). Aesthetic illusion as an effect of fiction. *Style* 38, 325–351.

Zoltan, K. (2000). *Metaphor and Emotion: Language, Culture, and Body in Human Feeling*. Cambridge, New York: Cambridge University Press.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

APPENDIX

Source Text
天净沙·秋思
(元)马致远
枯藤老树昏鸦，小桥流水人家，古道西风瘦马。夕阳西下，断肠人在天涯。

Target Text 1

Autumn
by Ma Chi-yuan

Crows hovering over rugged old trees wreathed with rotten vine—the day is about done. Yonder is a tiny bridge over a sparkling stream, and on the far bank, a pretty little village. But the traveller has to go on down this ancient road, the west wind moaning, his bony horse groaning, trudging towards the sinking sun, farther and farther away from home.

–tr. by Weng Xianliang

Target Text 2

Tune to"Sand and Sky"
—Autumn Thoughts
by Ma Zhiyuan

Dry vine,old tree,crows at dusk,
Low bridge,stream running,cottages,
Ancient road,west wind,lean nag,
The sun westering
And one with breaking heart at the sky's edge.

-tr.by Wayne Schlepp

Target Text 3

Autumn Thoughts
by Ma Zhiyuan

Withered vines, olden tree, evening crows;
Tiny bridge, flowing brook, hamlet homes;
Ancient road, wind from west, bony horse;
The sun is setting,
Broken man, far from home, roams and roams.

-tr.by Zhao Zhentao

Target Text 4

Tune:Tian Jin Sha
by Ma Zhiyuan

Withered vines hanging on old branches,
Returning crows croaking at dusk.
A few houses hidden past a narrow bridge,
And below the bridge a quiet creek running.
Down a worn path,in the west wind,
A lean horse comes plodding.
The sun dips down in the west,
And the lovesick traveller is still at the end of the world.

-tr.by Ding Zuxin and Burton Raffel.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 08 September 2022
DOI 10.3389/fpsyg.2022.986450

Check for updates

OPEN ACCESS

EDITED BY
Yan Jing Wu,
Ningbo University, China

REVIEWED BY
Mingjun Wu,
Jiangsu University of Science and Technology, China
Zhenguang Cai,
The Chinese University of Hong Kong, Hong Kong SAR, China

*CORRESPONDENCE
Tao Zeng
taozengclarry@hnu.edu.cn

SPECIALTY SECTION
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

RECEIVED 05 July 2022
ACCEPTED 01 August 2022
PUBLISHED 08 September 2022

CITATION
Zeng T, Chen C and Guo J (2022) First language translation involvement in second language word processing. *Front. Psychol.* 13:986450.
doi: 10.3389/fpsyg.2022.986450

First language translation involvement in second language word processing

Tao Zeng*, Chen Chen and Jiashu Guo

School of Foreign Languages, Hunan University, Changsha, China

Studies on bilingual word processing have demonstrated that the two languages in a mental lexicon can be parallelly activated. However, it is under discussion whether the activated, non-target language gets involved in the target language. The present study aimed to investigate the role of the first language (L1, the non-target one) translation in the second language (L2, the target one) word processing. The tasks of semantic relatedness judgment and lexical decision were both adopted, to explore the relation of the possible L1 involvement and the task demand. Besides, bilinguals with relatively higher and lower L2 proficiency were recruited, to clarify the potential influence of L2 proficiency. Results showed that the manipulation of L1 translation exerted an influence on bilinguals' task performances, indicating that L1 translation was involved, but did not just serve as a by-product when bilinguals were processing L2 words. And about the influence of L2 proficiency, the higher proficiency bilinguals performed better than the lower proficiency ones when the L1 translations could be taken advantage of, indicating a better access to L1 translation in L2 word processing, as bilinguals' L2 proficiency increased. As for the task demands, the L1 translation was partially involved in Experiment 1 while a full involvement was observed in Experiment 2, suggesting a differed depth of L1 translation involvement, if the task demands allowed. The present study supplemented the previous ones due to its participants (the intermediate bilinguals) and tasks (the tasks of semantic relatedness judgment and lexical decision); besides, it provided an interesting view into interpreting the "task schema" of the BIA+ model.

KEYWORDS

second language, first language, translation, language proficiency, word processing

Introduction

A core issue in research on bilingualism is how bilinguals access words in their two languages. Numerous studies have been conducted accordingly, proposing two competing views. Early research suggested the language-selective view, an idea that bilinguals selectively activate words from only the target language (see, e.g., Watkins and Peynrcoğlu, 1983; Scarborough et al., 1984; Gerard and Scarborough, 1989). The non-selective view, on the other hand, holds that words from both the target and the non-target languages are activated (see, e.g., Beauvillain and Grainger, 1987; Green, 1998; Meuter and Allport, 1999). In recent decades, there has been a bulk of studies supporting the non-selective standpoint: it is claimed that words in the non-target language are also activated, even when bilinguals are only engaged in the target one. Corresponding evidence comes from experiments on homographs (Schwartz and Kroll, 2006), cognates (Van Hell and de Groot, 2008; Van Assche et al., 2009), semantically related words (Kroll et al., 2008), etc. There are also some theoretical models of bilingual lexical processing which defend the non-selective view, such as the IC model (Inhibitory Control Model, Green, 1998), BIA+ model (Bilingual Interactive Activation Model+, Dijkstra and Van Heuven, 2002; Dijkstra, 2005), and The Three-Stage Model of L2 Lexical Development (Jiang, 2000).

Despite the compelling evidence for non-selective activation, the role of the non-target language is still being discussed: does the activated non-target language act as a by-product (it is only activated), or does it get involved in the target language processing (it is activated, meanwhile helps to process the target language)? Several experiments were therefore conducted, with a special interest in the role of the first language (L1, the non-target one) during the second language (L2, the target one) processing. The participation of L1 in L2 processing was therefore revealed: the psychological experience of L1 was automatically activated in L2 tasks (e.g., Vukovic and Williams, 2014; Ahlberg et al., 2017); the neural networks of L1 orthographic, semantic and phonetic processing were active in L2 tasks (e.g., Tan et al., 2003; Xue et al., 2004; Nelson et al., 2009; Gao et al., 2015); the L1 knowledge on lexical meaning and collocation played a role when one dealt with L2 words (e.g., Wolter and Gyllstad, 2011; Zhang et al., 2017; Cao, 2018). To be noticed, the above studies focused on the involvement of L1 language, e.g., the processing strategy and knowledge that take a part in L2 use. Such L1 language involvement is at a much more general scale, possibly with no access to a certain L1 word. We thus turn to the specific aspect, that is, whether the L1 translation equivalents get involved in L2 use.

Some prior studies have been conducted to resolve the issue, adopting the paradigms of cross-language priming (Beauvillain and Grainger, 1987; Gerard and Scarborough, 1989; Dijkstra et al., 2000) or word translation (De Groot and Hoeks, 1995; Dufour and Kroll, 1995). The above paradigms took L2 word as the prime; the L1 translation was also presented as the target or spoken out as the response, which requires the obligate processing of both L2 and L1. Therefore, it is difficult to tell whether the possible L1 translation involvement is due to the L1 mediation in L2 word processing, or just due to the task requirement (Grosjean, 1998). For that reason, recent studies switched to monolingual tasks, within which the experimental stimuli and the required responses are only associated with L2 words. It is in such a manner that the obligatory processing of L1 can be avoided. The recently employed tasks are semantic relatedness judgment (e.g., Thierry and Wu, 2004, 2007) or lexical decision (e.g., Jiang et al., 2019). In both tasks, participants are asked to accomplish L2 tasks. Unknown to them, the words' L1 translations are manipulated: in the semantic relatedness judgment task, the L1 translations of an L2 word-pair could share a logo-graphic character or not (translation repetition/non-repetition); and in the lexical decision task, the L1 translation of a word has a high or low lexical frequency (high/low translation frequency). Supposing that, in either task, the manipulated L1 translation leads to participants' differing performances, the L1 translation involvement in L2 word processing can thus be verified.

The above two tasks, however, might be different in nature. In the lexical decision task, there exists a full activation of the L1 translation word, due to the manipulation of its word frequency. But as for the task of semantic relatedness judgment, it has been proposed that the manipulation of a shared character in L1 translation, essentially, creates a kind of L1 form-repetition. It was the repetition of a character that helped to promote L2 word processing, without the need to activate a whole L1 translation word (Costa et al., 2017; Jiang et al., 2019). And in that case, the L1 translation word may get fully involved in the lexical decision task, but partially involved in the task of semantic relatedness judgment. It is therefore worth exploring, whether the L1 translation involvement, if it exists, varies in depth due to the different task demands? To tackle this issue, both the task of semantic relatedness judgment and lexical decision will be adopted in the present study to test the same participants. If the L1 translation is involved in both tasks, the varied depth of L1 translation involvement on different task demands can be demonstrated.

Additionally, L2 proficiency is another factor that can affect the potential L1 translation involvement. Some previous studies agreed on the involvement of L1 translation, but with the disagreement on how the bilingual's L2 proficiency plays a part. Li et al. (2018) analyzed high- and low-proficient Chinese-English bilinguals' responses in the task of semantic relatedness judgment. They found that while the target words were semantically related, the effect of L1 activation (measured in reaction times) was greater in high proficiency group, than in its low proficiency counterpart. It indicated an increasing involvement of L1 translation with the L2 promotion. Hu and Qi (2014) reached the opposite conclusion, however. In their experiment, high- and low-proficient Chinese-English bilinguals performed a lexical decision task. The high proficiency group showed no different reaction times for the target English words (whose Chinese translations were acquired in advance) and the non-target ones (with no such acquisition), yet the significantly shorter reaction times for target English words were observed in the low-proficient group. The contrast suggested a decrease in L1 translation involvement as L2 proficiency was promoted. As shown above, results of the previous studies are not consistent. The present study will thus recruit the higher- and lower-proficient bilinguals; once the two proficiency groups are differed in task performances, the influence of L2 proficiency can be revealed.

In considering all those facts, the present study attempts to investigate the role of L1 translation in L2 word processing. Both the tasks of semantic relatedness judgment and lexical decision are adopted, with the consideration of a possible influence of task demand. Besides, bilinguals of higher- and lower- L2 proficiency are recruited. The two groups' task performances are to be compared, through which we may investigate how bilingual's L2 proficiency affects the possibly involved L1 translation in L2 word processing.

Accordingly, the research questions are stated as: (1) What is the role of the L1 translation in L2 word processing (get involved, or serve as a by-product)? (2) If the L1 translation gets involved in L2 word processing, what is the influence of L2 proficiency and task demands?

Experiment 1: Semantic relatedness judgment

Method

Participants

Fifty-eight bilinguals aging averagely around 21.3 years of age from a major university in Hunan participated in the experiment. All participants are right-handed with normal or corrected to normal eyesight. They are all native Chinese speakers who learn English as their second language, without experience of living or studying abroad. The participant's L2 proficiency level is measured by their scores on the Test for English Majors-Band 8 (TEM-8, an English proficiency test for English majors in China; the passing of the test indicates a relatively high level of proficiency) and College English Test-Band 4 (CET-4, an English proficiency test for Chinese college students). Participants were asked to complete the Oxford Quick Placement Test (OQPT), grouped according to their test score, as well as L2 proficiency level and learning duration of English: the higher proficiency group, consisting of 29 participants (mean age = 23.1, OQPT = 48.5/60, learning duration = 13 years and 7 months, with a proficiency level not lower than TEM-8 or a CET-4 score higher than 600/710); and the lower proficiency group, consisting of 29 participants (mean age = 19.5, OQPT = 31.1/60, learning duration = 10 years and 7 months, with a proficiency level not higher than CET-4 or a CET-4 score lower than 400/710).

Materials

Forty English word pairs (eighty words) were selected from the British National Corpus (BNC) as the target stimuli. The words in a pair were matched in length and frequency. In each pair, the words were either semantically related (e.g., college-student) or unrelated (e.g., sport-fork). And unknown to participants, the Chinese translations of the word pair may share a Chinese character (e.g., college-student, translated as 学院 – 学生) or not (e.g., sport-fork, translated as 运动 – 叉子) (the repeated character, if any, was of the same position in their Chinese translations). Four conditions were therefore created: related & repeated, unrelated & repeated, related & unrepeated, and unrelated & unrepeated (see Table 1).

Procedure

After giving informed consent, participants were tested individually. They were instructed to perform the semantic relatedness judgment as fast and accurately as possible. Stimuli of the judgment were all presented visually, using the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

The participants were first familiarized with the task in a practice session with 4 trials; the formal experiment would not start until participants' accuracy reached 90%.

Each trial started with a 500 ms fixation sign. The first word in a pair was then presented for 500 ms, followed by the second word. Participants were asked to judge whether words in the pair were semantically related by pressing the button D (related) or J (unrelated), as soon as they saw the second word. Once the responses of participants were registered, the second word disappeared from the screen. An average experimental session lasted about 15 min.

Data preparation

Participants' reaction time (RT) and accuracy (ACC) were collected and analyzed. Trials with an incorrect response (4.7% of all trials) were excluded from the RT analysis, and so were trials with excessively fast or slow reactions (RT below 300 ms or above 3,000 ms, 5.0% of all trials). Only trials with excessively fast or slow reactions (RT below 300 ms or above 3,000 ms, 5.0% of all trials) were trimmed in ACC analysis. Table 2 displays the RT and ACC means for all conditions.

Data analyses were run using R (R Core Team, 2019). There were two types of models conducted: (a) to investigate the possible effect of semantic-relatedness and translation-repetition, fixed effects in the model were established as: semantic relatedness (related, unrelated), translation repetition (repeated, unrepeated), and L2 proficiency (the higher proficiency group, the lower proficiency group); (b) to investigate the influence of L2 on the possible L1 translation involvement, we conducted models under two different translation repetition conditions, in which the fixed effect was

TABLE 1 Conditions (semantic relatedness and repetition in translation of the word pair) in the task of semantic relatedness judgment.

Repetition in translation (Implicit factor)	Semantic relatedness (Explicit factor)	
	Semantic related (S+)	Semantic unrelated (S−)
Translation repeated (T+)	**College—Student** 学院 – 学生 SRE 4.82 (0.13) SRC 4.81 (0.17)	**Angel—Genius** 天使 – 天才 SRE 1.49 (0.46) SRC 1.54 (0.49)
Translation unrepeated (T−)	**Milk—Bread** 牛奶 – 面包 SRE 4.35 (0.41) SRC 4.03 (0.52)	**Sport—Fork** 运动 – 叉子 SRE 1.34 (0.37) SRC 1.34 (0.46)

(1) A group of 10 high-level bilinguals (of Chinese and English), with no participation in the present study, was required to measure the mean semantic relatedness of the English stimuli (SRE, measured on a scale of 1–5), and the mean semantic relatedness of the stimuli translated into Chinese (SRC, measured on a scale of 1–5). (2) Standard deviations of SRE and SRC scores are given in parentheses. (3) The factor of Semantic Relatedness and Repetition in Translation were not correlated (SRE and Repetition in Translation: r = 0.156, p = 0.337; SRC and Repetition in Translation: r = 0.095, p = 0.560).

set as L2 proficiency (the higher proficiency group, the lower proficiency group). For both model types, the random effects structure included by-participant and by-item effects.

The above two model types were both associated with an RT, as well as an ACC analysis. In RT analysis, raw data was log-transformed to compensate for the lack of a normal distribution. The models were then conducted as linear mixed-effect ones with the *LmerTest* package (Kuznetsova et al., 2017). Analysis-of-variance was calculated using the function of *anova*. *Post-hoc* group comparisons were made using the *emmeans* package (Lenth et al., 2020). And in ACC analysis, models were conducted as logic mixed-effect ones with the *LmerTest* package. Analysis-of-variance was calculated using the package *CAR* (Fox and Weisberg, 2019). Similarly, *post-hoc* group comparisons were made using the package *emmeans*.

Results

The semantic relatedness effect

There existed the main effect of semantic relatedness in RT [$F(1) = 22.28$, $p < 0.001$], as well as in ACC [$\chi^2(1) = 6.39$, $p = 0.011$]. *Post-hoc* group comparison demonstrated that, participants exhibited a faster reaction ($\beta = -0.25$, SE = 0.05, $t = -4.68$, $p < 0.001$) and higher accuracy ($\beta = 0.70$, SE = 0.40, $z = 1.76$, $p = 0.078$) to semantically related word pairs than to unrelated ones. Such a better performance suggested the semantic relatedness effect.

The translation repetition effect

Importantly, the main effect for translation repetition was reported in both RT [$F(1) = 4.39$, $p = 0.043$] and ACC analysis [$\chi^2(1) = 23.97$, $p < 0.001$]. *Post-hoc* group comparison revealed that word pairs with a repeated character in translation, compared to those without, induced a slower reaction ($\beta = 0.10$, SE = 0.05, $t = 2.09$, $p = 0.043$) and lower accuracy ($\beta = -0.85$, SE = 0.41, $z = -2.08$, $p = 0.037$). Such translation repetition effect showed that the manipulation of L1 translation exerted an influence on participants' performances, indicating the L1 involvement when processing L2 words.

To further investigate the effect of translation repetition, we focused on the two-way interaction of semantic relatedness and translation repetition, which reached marginal significance in RT [$F(1) = 3.65$, $p = 0.062$] and significance in ACC [$\chi^2(1) = 22.17$, $p < 0.001$]. Further *post-hoc* analysis showed that on the semantic-unrelated condition, exposure to word pairs with translation overlaps impeded participants' judgment (shown as slower reaction: $\beta = 0.19$, SE = 0.07, $t = 2.77$, $p = 0.008$; and lower accuracy: $\beta = -2.31$, SE = 0.57, $z = -4.05$, $p < 0.001$), in comparison to word pairs without the translation overlaps. No such effect was found on the semantic-related condition. This may be due to the over-generalization of participants: without the explicit information given on semantics, an association may be assigned between the semantically unrelated words (i.e., they were wrongly assumed to be semantically related) because their L1 translations are somehow associated. But once words are related in meaning, which explicitly reveals the semantic relatedness, participants can judge easily without any help from such generalized "association." That is why the translation repetition effect was found only under the condition of semantic-unrelatedness.

To sum up, in the task of semantic relatedness judgment (Experiment 1), only a manipulated character of L1 translation was enough to affect the L2 word processing. It is thus necessary to take into account results of the lexical decision task (Experiment 2), in which the L1 translation word is manipulated as a whole. In such a manner we can examine whether a whole L1 translation word can get involved in L2 word processing.

Performances between proficiency groups

The main effect of L2 proficiency was reported in RT [$F(1) = 8.13$, $p = 0.006$] and ACC analysis [$\chi^2(1) = 28.00$, $p < 0.001$]. Unsurprisingly, the *post-hoc* group comparison provided evidence for a faster response ($\beta = -0.13$, SE = 0.04, $t = -2.85$, $p = 0.006$) and higher accuracy ($\beta = 1.02$, SE = 0.18, $z = 5.75$, $p < 0.001$) in the more proficient group.

To clarify how the two proficiency groups were affected by the manipulated L1 translation, we conducted models on different conditions of translation repetition, focusing on the effect of L2 proficiency. Under the condition of no translation repetition, the main effect of L2 proficiency was significant [RT: $F(1) = 11.57$, $p < 0.001$; ACC: $\chi^2(1) = 24.99$, $p < 0.001$). The condition of translation repetition also led to a significant main effect of L2 proficiency [RT: $F(1) = 12.70$, $p < 0.001$; ACC: $\chi^2(1) = 10.87$, $p < 0.001$]. Further *post-hoc* comparisons revealed that the more proficient

 10.3389/fpsyg.2022.986450

TABLE 2 Mean reaction time (RT, in ms) and accuracy (ACC, in %) of the two proficiency groups in semantic relatedness judgment.

	More proficient group (n = 29)		Less proficient group (n = 29)	
Condition	**ACC**	**RT**	**ACC**	**RT**
S+T+(n = 10)	93.57 (24.57)	1162 (535)	83.67 (37.04)	1272(560)
S+T−(n = 10)	86.35 (34.40)	1150 (515)	82.44 (38.11)	1217(574)
S−T+(n = 10)	61.54 (48.75)	1469 (551)	54.27 (49.92)	1567 (591)
S−T−(n = 10)	96.98 (17.14)	1275(528)	81.33 (39.04)	1545 (628)

Four conditions are presented: semantic related (S+), semantic unrelated (S−), translation repeated (T+), and translation unrepeated (T−). Standard deviations are given in parentheses.

TABLE 3 Conditions (the mean frequency, length, and translation frequency of target words) in the lexical decision task.

Condition	Target words (Examples)	Frequency (Per million)	Length (Number of letters)	Chinese translations	Translation frequency (Per million)
HTF	Research	22.1	7.5	研究	810.3
LTF	Evidence	23.9	7.5	证据	27.8
HLF	Room	307.8	6.3	房间	76.3
LLF	Carpet	21.6	6.6	地毯	77.8

(1) Target words were all nouns or verbs, each had a unique disyllabic Chinese translation. According to Jiang et al. (2019), the high/low frequency of the English word was based on Brysbaert and New (2009); that of the Chinese word was based on Beijing Language Institute [BLI], 1986. (2) The factors of Frequency and Translation Frequency were not correlated: ($r = -0.181, p = 0.152$).

TABLE 4 Mean reaction time (RT, in ms) and accuracy (ACC, in %) of the two proficiency groups in lexical decision task.

	More proficient group (n = 29)		Less proficient group (n = 29)	
Condition	**ACC**	**RT**	**ACC**	**RT**
HLF (n = 16)	99.12 (9.33)	895 (401)	97.16(16.64)	972 (458)
LLF (n = 16)	97.98 (14.09)	1,015 (471)	91.80 (27.46)	1,091 (518)
HTF (n = 16)	96.20 (19.15)	1,057 (443)	93.02 (25.51)	1,167 (540)
LTF (n = 16)	98.86 (10.64)	1,151 (520)	89.86 (30.22)	1,211 (561)

Four word conditions are presented as words with high frequency (HLF), low frequency (LLF), high translation frequency (HTF), and low translation frequency (LTF). Standard deviations are given in parentheses.

group exhibited a quicker reaction and higher accuracy, on the condition of unrepeated translation (RT: $\beta = -0.09$, SE = 0.03, $t = -3.40$, $p = 0.0007$, ACC: $\beta = 0.95$, SE = 0.20, $z = 4.84$, $p < 0.001$) and repeated translation (RT: $\beta = -0.11$, SE = 0.03, $t = -3.56$, $p = 0.0004$; ACC: $\beta = 0.48$, SE = 0.15, $z = 3.28$, $p = 0.001$). In sum, no contrast was found between the two conditions.

According to the task of semantic relatedness judgment (Experiment 1), it remained unclear which proficiency group gained greater influence of the involved L1 translation word. That brings the necessity to check the results of the lexical decision task (Experiment 2).

Experiment 2: Lexical decision

Method

Participants

Same as Experiment 1.

Materials

The experiment follows the materials by Jiang et al. (2019). Our test materials consisted of 64 English words (the target stimuli), 48 non-words and 16 English filler words. All of the materials were randomly presented.

Among the target stimuli, there were 32 words matched for length and lexical frequency, but differed in translation frequency (the frequency of its Chinese translation): half of them were with relatively high-frequency Chinese translations (HTF, high translation frequency), and the other half were with low-frequency ones (LTF, low translation frequency). Another 32 words, matched for length and translation frequency, were differed in lexical frequency: half of them were relatively high-frequency words (HLF, high lexical frequency) while the other half were low-frequency ones (LLF, low lexical frequency) (see **Table 3**).

Procedure

Participants were tested individually with their informed consent. Instructions of the lexical decision task were presented to participants orally and visually, with an emphasis on both speed and accuracy. Stimuli of the decision task were all presented visually, using the E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

The participants were first familiarized with the task in a practice session with 10 trials; the formal experiment would not start until participants' accuracy reached 90%.

Each trial began with a fixation sign, lasting for 500 ms. Thereafter the letter string was presented at the center of the screen, remaining until participants made the decision. Participants decided on whether the string made up a word or not, by pressing the button D (yes) or J (no). An average experimental session lasted about 15 min.

Data preparation

Participants' reaction time (RT) and accuracy (ACC) were collected and analyzed. In RT analysis, trials with an incorrect response (4.7% of all trials), as well as those with excessively fast or slow reactions (RT below 300 ms or above 3,000 ms, 5.0% of all trials) were excluded. In ACC analysis, only trials with excessively fast or slow reactions (RT below 300 ms or above 3,000 ms, 5.0% of all trials) were trimmed. Table 4 displays the RT and ACC means for all conditions.

In the present experiment, the same methods and packages for data analysis in Experiment 1 were adopted. Differed from Experiment 1, though, three types of models were conducted: (a) to verify the effect of lexical frequency, we focused on target words with LLF and HLF, setting the fixed effects as lexical frequency (high, low) and L2 proficiency (the higher proficiency group, the lower proficiency group); (b) to explore the possible effect of translation frequency, we chose the target words with LTF and HTF, setting the fixed effects as translation frequency (high, low) and L2 proficiency (the higher proficiency group, the lower proficiency group); (c) to investigate the influence of L2 on the possible L1 translation involvement, we conducted models on two different translation frequency conditions, in which the fixed effect was set as L2 proficiency (the higher proficiency group, the lower proficiency group). For each of the three types of models, the random effects structure included by-participant and by-item effects.

Results

The lexical frequency effect

A main effect of lexical frequency was revealed in RT [$F(1) = 33.11$, $p < 0.001$]. It can be seen from the *post-hoc* comparison that, the words of higher frequency led to participants' quicker reactions compared to the words of lower frequency ($\beta = -0.11$, SE = 0.02, $t = -5.82$, $p < 0.001$), known as the lexical frequency effect. However, no significant main effect of lexical frequency was reported in ACC.

The translation frequency effect

To be noted, RT analysis yielded a main effect for translation frequency [$F(1) = 7.42$, $p = 0.007$]. *Post-hoc* group comparison revealed that participants responded to higher translation frequency words faster than to lower translation frequency words ($\beta = -0.06$, SE = 0.02, $t = -2.75$, $p = 0.006$). It was an influence brought about by the manipulation of L1 translation, i.e., the translation frequency effect, which indicated the L1 involvement in L2 word processing.

The translation-frequency effect was not reported in ACC analysis. Yet a visual inspection of the data in the less proficient group (see Table 4) might indicate an ACC difference in conditions of higher- and lower-translation frequency (the more proficient group, however, who exhibited a rather high ACC, may exhibit a ceiling effect): words with a higher translation frequency induced higher accuracy, compared to those with a lower translation frequency. It indicated an easier processing of words with higher translation frequency, possibly serving as a verification of translation frequency effect in RT analysis.

Further, the translation frequency effect was independent of the lexical frequency effect: in the present experiment, the translation frequency effect was obtained from words matched in their lexical frequency. In this way, a more reliable conclusion can be reached.

Performances between proficiency groups

A main effect of L2 proficiency was shown in target words with manipulated lexical frequency [RT: $F(1) = 12.55$, $p < 0.001$; ACC: $\chi^2(1) = 5.09$, $p = 0.024$], as well as in those with manipulated translation frequency [RT: $F(1) = 9.92$, $p = 0.002$; ACC: $\chi^2(1) = 4.48$, $p = 0.034$]. *Post-hoc* comparison demonstrated that, the more proficient group exhibited a quicker response and higher accuracy, no matter when given target words varying in lexical frequency (RT: $\beta = -0.07$, SE = 0.02, $t = -3.53$, $p < 0.001$; ACC: $\beta = 1.33$, SE = 0.35, $z = 3.86$, $p < 0.001$) or in translation frequency (RT: $\beta = -0.06$, SE = 0.02, $t = -3.13$, $p = 0.002$; ACC: $\beta = 1.46$, SE = 0.29, $z = 5.13$, $p < 0.001$).

Additionally, the effect of L2 proficiency was analyzed separately in two translation frequency conditions, through which we can distinguish how different proficiency groups were affected by the manipulated L1 translation. On condition of lower translation frequency, the main effect of L2 proficiency was reported in only ACC analysis [$\chi^2(1) = 37.56$, $p < 0.001$]; in *post-hoc* analysis, the higher accuracy in the more proficient group was revealed ($\beta = 2.28$, SE = 0.48, $z = 4.78$, $p < 0.001$). However, on condition of the higher translation frequency, the main effect of L2 proficiency was observed in both RT [$F(1) = 8.70$, $p = 0.003$] and ACC [$\chi^2(1) = 4.48$, $p = 0.034$]; and in the *post-hoc* analysis, the faster response ($\beta = -0.08$, SE = 0.03, $t = -2.95$, $p = 0.003$), along with higher accuracy ($\beta = 0.64$, SE = 0.31, $z = 2.07$, $p = 0.038$) was obtained in the more proficient group. Overall, a contrast has been revealed: words with lower translation frequency led to nearly the same reaction speed across the two groups; when processing words with higher translation frequency, however, the more proficient group was significantly quicker than its less proficient counterpart. We can tell from the contrast that it was the more proficient participants, not the less proficient ones, who "gained more benefits" from the high-frequent L1 translation. It suggested a greater influence of L1 translation once the bilingual reaches a higher proficiency level.

General discussion

The present study adopted the tasks of semantic relatedness judgment and lexical decision, testing both the higher- and lower-proficiency bilinguals. Results revealed that participants performed better on semantically related word pairs and higher frequency words. More importantly, participants' worse performances were observed when the word pair shared a repeated character in L1 translations, or when the target words were with lower-frequency L1 translations, demonstrating that the manipulated L1 translation exerts an influence on bilinguals' task performances. We also found that the more proficient bilinguals "gained more benefits" from the high-frequency L1 translation, suggesting that the manipulated L1 translation had an even greater influence on the more proficient bilinguals, than their less proficient counterparts. Further discussions on the effects by the manipulated L1 translation, the L2 proficiency and the task demands, are presented as follows.

The involvement of L1 translation in L2 word processing

The present study aimed to explore the role of L1 translation in L2 word processing. Results showed that the words that share a character in their Chinese translations led to worse performances in the judgment of semantic relatedness (the translation repetition effect); words with a high frequency Chinese translations promoted the lexical decision (the translation frequency effect). The manipulation of not only the form, but the frequency of L1 translation had an effect when bilinguals processed L2 words, which revealed the involvement of L1 translation in L2 word processing.

That is, we suggested an L1 translation mediation, but not a strong and direct association between L2 lexicon and the word meaning. In fact, for the unbalanced bilinguals (as recruited in the present study), the L2 lexicon is learned after the complete construction of the concept (i.e., the word meaning). The connection between the concept and L2 lexicon can thus be weak. Such weak connection is supported by the asymmetric cross-language priming (which is weaker from L2 to L1 than that from L1 to L2, see, e.g., Keatley et al., 1994; Jiang, 1999), as well as the reduced emotional responses in L2 processing (see, e.g., Costa et al., 2014).

The L1 translation involvement has been well documented in a variety of participants and conditions. In the task of semantic relatedness judgment, for example, Thierry and Wu (2004, 2007) revealed an unconscious translation into L1 in proficient bilinguals' L2 comprehension. Zhang et al. (2012) replicated this pattern of results in late, non-proficient bilinguals. Xiao and Ni (2016) subdivided the condition "character repetition (in the target word's Chinese translation)" into the "first/final-character repetition," suggesting a translation into L1 even at the sub-lexical level. And in a lexical decision task, Jiang et al. (2019) recruited bilinguals with immersion experience in L2 and found also the influence of L1 translation, establishing that L1 translation served as an integral part of L2 word processing.

The phenomenon can be discussed in Jiang's Three-Stage Model of L2 Lexical Development (Jiang, 2000), according to which the developing L2 lexicon consists of three gradual stages: (a) the formal stage when a link between L2 words and L1 translations is established; (b) the L1 lemma mediation stage with a stronger L1–L2 link, when L2 word processing is mediated by lemmas of its L1 translation; (c) the L2 integration stage when L2 word information is represented independently, causing the L1–L2 link to be unnecessary in L2 word processing. The L2 mental lexicon of our participants may have reached the second stage (the intermediate level of L2 proficiency), during which L2 words are processed *via* their L1 translation equivalents. It may explain why any manipulation of the L1 translation (its repetition of form, or its frequency) will exert a significant effect on L2 word processing. And notably, the mental lexicon of our higher-proficiency bilinguals stopped at the same stage (the second stage) as their lower-proficiency counterparts. In other words, their lexical development, to some extent, is fossilized (for similar discussion, see Ma, 2015). And it is the L1 translation involvement that can be a major cause. In L2 word processing, the word meaning is mediated by the L1 translation; bilinguals may get used to the "walking stick" and pay less attention to the semantic context, within which the L2 word meaning can be analyzed and acquired (Mestres-Missé et al., 2007, 2014; Wu and Feng, 2014).

The influence of L2 proficiency

Additionally, we attempted to clarify how L2 proficiency affects the involvement of L1 translation. In RT analysis of the lexical decision task, two proficiency groups performed no differently on words with lower translation frequency; however, the higher-proficiency group was more sensitive than the lower-proficiency one while responding to words with higher translation frequency. The contrast indicated that participants with higher L2 proficiency appeared to benefit more from the L1 translation in L2 word processing. It suggested a better access to L1 translation as one's L2 proficiency increases.

The phenomenon can also be discussed within Jiang's Three-Stage Model of L2 Lexical Development. The stages in the model are not that clear-cut; the representations of bilingual's mental lexicon are in transition from one stage to another (Jiang, 2000). The model's first stage is with a weak L1–L2 link, which develops to be strong at the second stage, it is thus possible that the L1–L2 link at the second stage is gaining strength progressively, as the bilingual gets more proficient in L2. On this basis, we believe that although our two proficiency groups are undergoing the same stage of L1 lemma mediation, their L1–L2 link can differ in strength. Compared to the lower-proficiency ones, the higher-proficiency group exhibited a stronger link, making possible their easier access to both the L2 word and its L1 translation.

In our lexical decision task, the higher-proficiency bilinguals showed better access to L1 translation, which facilitated the task behavior. It seemed to be a kind of "advantage." In some tasks of semantic relatedness judgment, however, the higher-proficiency bilinguals' better access to L1 translation was found and reported as a "disadvantage": the higher-proficiency bilinguals exhibited worse performance on the condition of L1 translation repetition, compared with the condition of no repetition (Li et al., 2018; Qu, 2019). The distinction may result from the design of the two tasks. In the task of semantic relatedness judgment, the semantic association can be wrongly established due to the form-repeated L1 translation; the L1 translation involvement behaved as a hinderance. And in the lexical decision task, the word recognition can be improved due to the high-frequency L1 translation; the L1 translation involvement acted as an assistance. We thus propose that it is the task itself, rather than the involved L1 translation, that creates such a "disadvantage" or "advantage."

But notably, we suggest that the more proficient bilinguals exhibit a better access to L1 translation, which does not mean they rely more on the L1 translation involvement. In fact, the involvement of L1 translation seems to decrease, as one's L2 proficiency increases. In some studies, for example, no better performances were reported even if the L1 translation could be taken as a clue to promote the lexical decision (Hu and Qi, 2014; Jiang et al., 2019). It is because that differed from the present study, participants of those studies have reached quite a high level of L2 (they were university teachers with overseas education experience, see Hu and Qi, 2014; graduate students and visiting scholars studying at an American university, see Jiang et al., 2019). For those highly proficient bilinguals, their L2 mental lexicon can thus be approaching the stage of L2 integration, which makes the L1–L2 link and the L1 word information unnecessary in the L2 word processing. Therefore, in lexical decision tasks, those highly proficient bilinguals were almost independent of the L1 translation.

The task demands and the depth of L1 translation involvement

Additionally, it was found in the task of semantic relatedness judgment that, a form-repetition in L1 translation was enough to mediate L2 word processing. As for the lexical decision task, however, it was a whole L1 translation word that mediated the L2 processing. The contrast revealed that the depth of L1 translation involvement would change with the task demand; if a character of L1 translation word is sufficient for a certain L2 task, the involvement of a whole L1 translation word becomes unessential.

Such an effect of task demands has been discussed in the BIA+ model. The model is concerned with the processing of bilingual words. It consists of an identification system, which provides the word representations at the semantic, orthographic (lexical) and phonological levels; and a task schema, which takes a decision on how the response will be made. The task schema tends to optimize one's performances, based on an internal criterion (e.g., the shortest reaction time and the highest accuracy enough for accomplishing the task, which depends mainly on the task demand) (see Dijkstra and Van Heuven, 2002; Dijkstra, 2005). It seems to be in accord with the cognitive economy principle. And in the present study, the participant was able to "take a shortcut": they used only a form-repetition of, but not a whole L1 translation to mediate L2 word, through which a quicker reaction became possible. It is in essence an optimization of one's performances, or a kind of task-modulation, conducted by the task schema.

The optimization may be associated with the independence of task schema. The task schema to some extent functions on its own, separate from the identification system (Dijkstra, 2005). In this view, it is possible that in the semantic relatedness judgment, the representation of a word's L1 translation is activated, waiting for the possible reaction optimization. The task schema thereafter conducts the reaction optimization, according to which only a form-repetition of the L1 translation will be selected and taken advantage of, so that the L2 word processing can be mediated. That is to say: the whole L1 translation word is activated (which serves as a by-product), while it is only a form-repetition in L1 translation that gets involved in L2 word processing, due to the effect of task demand.

The effect of task demand, according to BIA+ model, is usually depicted on a broader scale. However, we found in the present study that it can work at a narrower range. It is obvious that in the task of semantic relatedness judgment, due to the L1 translation involvement, both the lexical and the semantic level of L1 words are useful to arrive at a response; but as for the lexical decision task, the processing at the semantic level becomes less critical. Thus in differed tasks, the language (L1) involvement may reach different levels (for similar discussion, see Wang et al., 2011). It can be regard as a task-modulation across the language levels, conducted by the task schema. That is the most common interpretation based on the BIA+ model. The present study, however, turned to a more specific scale. In both tasks of the present study, the L1 translation involvement has reached the lexical level. Experiment 2 reported a fully involved L1 translation word at the lexical level, while Experiment 1 suggested an involvement of L1 form-repetition; it is a partially involved L1 translation at the lexical level. The contrast of the two experiments showed that even at a certain language level, the depth of word processing could differ, if the task demands allowed. In other words, the task-modulation (derived from the operation of task schema, according to BIA+ model) may play a role not only across the language levels, but just at a certain level.

Conclusion

This article investigated the role of L1 translation in L2 word processing, while taking into consideration the influence of L2 proficiency and task demands. Results showed that the performances of participants were affected by the manipulated L1 translation, indicating an involvement of L1 translation in L2 word processing. It was also found that compared with the lower-proficiency ones, the higher-proficiency bilinguals could be more sensitive to the manipulated L1 translation, demonstrating their better access to both the L2 word and its L1 translation. Additionally, the depth of L1 translation involvement was found to vary with the task demands. It suggested a kind of task-modulation at a certain language level, which may provide an uncommon viewpoint in interpreting the BIA+ model.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

TZ: conceptualization, project administration, writing—review and editing, and supervision. CC: validation, data curation, formal analysis, and writing—original draft. JG: methodology and investigation. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the National Social Science Foundation of China (Grant No. 19BYY079).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.986450/full#supplementary-material

References

Ahlberg, D. K., Bischoff, H., Kaup, B., Bryant, D., and Strozyk, J. V. (2017). Grounded cognition: comparing language × space interactions in first language and second language. *Appl. Psycholinguist.* 39, 437–459. doi: 10.1017/s014271641700042x

Beauvillain, C., and Grainger, J. (1987). Accessing interlexical homographs: some limitations of a language-selective access. *J. Mem. Lang.* 26, 658–672.

Beijing Language Institute [BLI] (1986). *A Frequency Dictionary of Modern Chinese*. Beijing: Beijing Language Institute Press.

Brysbaert, M., and New, B. (2009). Moving beyond Kučera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behav. Res. Methods* 41, 977–990. doi: 10.3758/BRM.41.4.977

Cao, Y. (2018). The influence of L1 semantic information on Chinese EFL learners' processing of L2 formulaic sequences. *For. Lang. Their. Teach.* 5, 109–117. doi: 10.13458/j.cnki.flatt.004531

Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., et al. (2014). Your morals depend on language. *PLoS One* 9:e94842.

Costa, A., Pannunzi, M., Deco, G., and Pickering, M. J. (2017). Do bilinguals automatically activate their native language when they are not using it? *Cogn. Sci.* 41, 1–15. doi: 10.1111/cogs.12434

De Groot, A. M. B., and Hoeks, J. C. J. (1995). The development of bilingual memory: evidence from word translation by trilinguals. *Lang. Learn.* 45, 683–724. doi: 10.1111/j.1467-1770.1995.tb00458.x

Dijkstra, T. (2005). *Bilingual Visual Word Recognition and Lexical Access. Handbook of Bilingualism: Psycholinguistics Approaches*. New York, NY: Oxford University Press.

Dijkstra, T., Timmermans, M., and Schriefers, H. (2000). On being blinded by your other language: effects of task demands on interlingual homograph recognition. *J. Mem. Lang.* 42, 445–464. doi: 10.1006/jmla.1999.2697

Dijkstra, T., and Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Bilingualism* 5, 175–197. doi: 10.1017/S1366728902003012

Dufour, R., and Kroll, J. F. (1995). Matching words to concepts in two languages: a test of the concept mediation model of bilingual representation. *Mem. Cogn.* 23, 166–180. doi: 10.3758/BF03197219

Fox, J., and Weisberg, S. (2019). *An R Companion to Applied Regression*, 3rd Edn. Thousand Oaks, CA: Sage.

Gao, Y., Wei, N., Wang, Z. K., Jian, J., Ding, G. S., Meng, X. Z., et al. (2015). Interaction between native and second language processing: evidence from a functional magnetic resonance imaging study of Chinese-English bilingual children. *Acta Psychol. Sin.* 47, 1419–1432. doi: 10.3724/SP.J.1041.2015.01419

Gerard, L. D., and Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 305–315. doi: 10.1037/0278-7393.15.2.305

Green, D. W. (1998). Mental control of the bilingual lexico-sematic system. *Bilingualism* 1, 67–81. doi: 10.1017/S1366728998000133

Grosjean, F. (1998). Studying bilinguals: methodological and conceptual issues. *Bilingualism* 1, 131–149. doi: 10.1017/S136672899800025X

Hu, J. H., and Qi, S. (2014). The effect of second language proficiency on lexical access mechanism of Chinese English learners. *For. Lang. Educ.* 35, 59–63. doi: 10.16362/j.cnki.cn61-1023/h.2014.06.022

Jiang, N. (1999). Testing processing explanations for the asymmetry in masked cross-language priming. *Bilingualism* 2, 59–75. doi: 10.1017/S1366728999000152

Jiang, N. (2000). Lexical representation and development in a second language. *Appl. Lingusit.* 21, 47–77. doi: 10.1093/applin/21.1.47

Jiang, N., Li, M., and Guo, T. M. (2019). A tale of two frequency effects: toward a verification model of L2 word recognition. *Appl. Psycholinguist.* 41, 1–22. doi: 10.1017/S0142716419000481

Keatley, C. W., Spinks, J. A., and de Gelder, B. (1994). Asymmetrical cross-language priming effects. *Mem. Cognit.* 22, 70–84. doi: 10.3758/BF03202763

Kroll, J. F., Bobb, S. C., Misra, M., and Guo, T. M. (2008). Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychol.* 128, 416–430. doi: 10.1016/j.actpsy.2008.02.001

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). LmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26.

Lenth, R., Singmann, H., Love, J., Buerkner, P., and Herve, M. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. Available online at: https://CRAN.R-project.org/package=emmeans (accessed June 22, 2022).

Li, J. H., Zheng, Y., and Yang, J. (2018). The influence of second language proficiency on the automatic activation of Chinese during English lexical access in Chinese-English bilinguals. *For. Lang. Teach. Res.* 50, 101–113.

Ma, Z. (2015). The motivation and path of concept reconstruction in second language vocabulary acquisition. *J. Pla. Univ. For. Lang.* 38, 74–80.

Mestres-Missé, A., Münte, T. F., and Rodriguez-Fornells, A. (2014). Mapping concrete and abstract meanings to new words using verbal contexts. *Second Lang. Res.* 30, 191–223. doi: 10.1177/0267658313512668

Mestres-Missé, A., Rodriguez-Fornells, A., and Münte, T. F. (2007). Watching the brain during meaning acquisition. *Cereb. Cortex* 17, 1858–1866. doi: 10.1093/cercor/bhl094

Meuter, R. F. I., and Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *J. Mem. Lang.* 40, 25–40. doi: 10.1006/jmla.1998.2602

Nelson, J. R., Liu, Y., Fiez, J., and Perfetti, C. A. (2009). Assimilation and accommodation patterns in ventral occipitotemporal cortex in learning a second writing system. *Hum. Brain. Mapp.* 30, 810–820. doi: 10.1002/hbm.20551

Qu, C. H. (2019). The influence of L2 lexical proficiency on the L1 automatic activation during L2 lexical processing of Chinese-English bilinguals. *For. Lang. Teach. Res.* 51, 560–569.

R Core Team (2019). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. Available online at: https://www.R-project.org/ (accessed June 23, 2022).

Scarborough, D. L., Gerard, L., and Cortese, C. (1984). Independence of lexical access in bilingual word recognition. *J. Verb. Learn. Verb. Behav.* 23, 84–99. doi: 10.1016/S0022-5371(84)90519-X

Schwartz, A. I., and Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *J. Mem. Lang.* 55, 197–212. doi: 10.1016/j.jml.2006.03.004

Tan, L. H., Spinks, J. A., Feng, C. M., Siok, W. T., Perfetti, C. A., Xiong, J. H., et al. (2003). Neural systems of second language reading are shaped by native language. *Hum. Brain. Mapp.* 18, 158–166. doi: 10.1016/S1053-8119(01)91955-X

Thierry, G., and Wu, Y. J. (2004). Electrophysiological evidence for language interference in late bilinguals. *Neuroreport* 15, 1555–1558. doi: 10.1097/01.wnr.0000134214.57469.c2

Thierry, G., and Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proc. Natl. Acad. Sci. U. S. A.* 104, 12530–12535. doi: 10.1073/pnas.0609927104

Van Assche, E., Duyck, W., Hartsuiker, R. J., and Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychol. Sci.* 20, 923–927. doi: 10.1111/j.1467-9280.2009.02389.x

Van Hell, J. G., and de Groot, A. M. B. (2008). Sentence context modulates visual word processing and translation in bilinguals. *Acta Psychol.* 128, 431–451. doi: 10.1016/j.actpsy.2008.03.010

Vukovic, N., and Williams, J. N. (2014). Automatic perceptual simulation of first language meanings during second language sentence processing in bilinguals. *Acta Psychol.* 145, 98–103. doi: 10.1016/j.actpsy.2013.11.002

Wang, R. M., Deng, H. S., Li, J. J., Li, L., and Fan, M. (2011). The activation of non-attended language in language comprehension of Chinese-English bilinguals. *Acta Psychol. Sin.* 43, 771–783. doi: 10.1631/jzus.B1000278

Watkins, M. J., and Peynrcoğlu, Z. F. (1983). On the nature of word recall: evidence for linguistic specificity. *J. Verb. Learn. Verb. Behav.* 22, 385–394. doi: 10.1016/S0022-5371(83)90246-3

Wolter, B., and Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Appl. Lingusit.* 32, 430–449. doi: 10.1093/applin/amr011

Wu, S. Y., and Feng, X. H. (2014). The cognitive mechanisms in the processing of metaphors by L2 learners: evidence from a cross-modal priming experiment. *For. Lang. Teach. Res.* 46, 889–900.

Xiao, W., and Ni, C. B. (2016). The L1 automatic activation of foreign language learners: evidence from ERPs. *For. Lang. Teach. Res.* 48, 236–248.

Xue, G., Dong, Q., Jin, Z., Zhang, L., and Wang, Y. (2004). An fMRI study with semantic access in low proficiency second language learners. *Neuroreport.* 15, 791–796. doi: 10.1097/00001756-200404090-00010

Zhang, H. Y., Ma, F. Y., Chen, B. G., and Guo, T. M. (2012). Automatic activation of Chinese during English word reading in non-proficient Chinese-English bilinguals. *For. Lang. Teach. Res.* 44, 719–727.

Zhang, P., Wei, J. S., Hu, B., Yu, J. Y., and Chen, Y. Y. (2017). Accelerated priming effect of Chinese L1 intralexical knowledge in English collocational processing. *For. Lang. Teach. Res.* 1, 60–72. doi: 10.13458/j.cnki.flatt.004329

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 01 December 2022
DOI 10.3389/fpsyg.2022.980967

Check for updates

OPEN ACCESS

EDITED BY
Haoyun Zhang,
The Pennsylvania State University (PSU),
United States

REVIEWED BY
Tao Zeng,
Hunan University,
China
Zhongyi Lu,
Hebei Normal University,
China

*CORRESPONDENCE
Wenguang He
hewg@qfnu.edu.cn

SPECIALTY SECTION
This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Psychology

RECEIVED 29 June 2022
ACCEPTED 20 September 2022
PUBLISHED 01 December 2022

CITATION
Bai Y and He W (2022) Involvement of the sensorimotor system in less advanced L2 processing: Evidence from a semantic category decision task.
Front. Psychol. 13:980967.
doi: 10.3389/fpsyg.2022.980967

Involvement of the sensorimotor system in less advanced L2 processing: Evidence from a semantic category decision task

Yating Bai and Wenguang He*

Department of Psychology, Qufu Normal University, Qufu, China

There is increasing evidence indicating that the sensorimotor system is involved in advanced L2 processing, which raises the question of what role sensorimotor information plays in the course of less advanced L2 comprehension. In the current study, two experiments were conducted using a lexical decision task (LDT) and semantic category task (SCT). The results showed that, in the LDT, a task more likely to result in participants making judgments based on the physical properties of words (e.g., familiarity, orthography), "up" words (e.g., *sun, plane*) did not result in faster upward than downward responses, and "down" words (e.g., *tunnel, cave*) also did not result in faster downward than upward responses. In the SCT, compatibility effects were found; specifically, searching for the up target after "up" words was faster than after "down" words and searching for the bottom target after "down" words was faster than after "upward-pointing" words. Hence, we concluded that L2 sensorimotor association, at least for L2 with low proficiency, not automatic in nature and is dependent upon deeper semantic task demands.

KEYWORDS

embodied cognition, L2 processing, lexical decision task, semantic category task, bilingualism

Introduction

Recent research evidence has indicated that native language (L1) processing is more than a purely symbolic process (Kintsch, 1988; Glenverg and Kaschak, 2002; Felisatti et al., 2022), in which the sensorimotor system is also activated (Barsalou, 1999). For example, in studies investigating the embodied effect in sentence comprehension (Zwaan et al., 2002), researchers found that participants responded significantly faster to pictures consistent with the implied sentential content. Hence, they concluded that perceptual symbols are routinely activated in language comprehension. Similarly, a considerable number of behavioral studies found that people simulate a range of other perceptual features, such as orientation (Stanfield and Zwaan, 2001), location (Bergen et al., 2007), visibility conditions (Horton and Rapp, 2003), and motion (Kaschak et al., 2005). Neuroimaging studies also proved that language comprehension crucially involves the simulation of sensory, motor, and emotional content (De Grauwe et al., 2014; Birba et al., 2020). Recently, *f*MRI studies showed that reading action verbs reliably activated the motor cortex (Kiefer and Pulvermüller, 2011); these areas of the brain were even recruited while reading nouns expressing graspable objects (Desai et al., 2016).

Although evidence increasingly supports embodied cognition, the exact mechanism of these activations remains subject to debate. Some scholars have argued that embodied mechanisms are indeed an inseparable and functionally crucial part of language processing (Vukovic et al., 2017), but others have contended that these mechanisms might just be a by-product of language processing, functionally "redundant," and irrelevant to efficient semantic comprehension (Kühne and Gianelli, 2019). To disentangle the above-mentioned issues, studies focusing on the embodied effects in L2 processing have been conducted. The results indicated that, as in L1 processing, the sensorimotor system is also involved in L2 processing, although some differences were found in terms of degree or time course (De Grauwe et al., 2014; Dudschig et al., 2014; Vukovic and Williams, 2014; Foroni, 2015; Buccino et al., 2017; Ratcliffe and Tokarchuk, 2020).

Based on the above findings, some researchers have argued that embodied effects are universal across L1 and L2. However, such conclusions were drawn with caution because the bilingual speakers recruited in these studies were all highly proficient. Some bilingual models, such as the Revised Hierarchical Model (RHM; Kroll and Stewart, 1994) and the Bilingual Interactive-Activation Model (BIA-d; Grainger et al., 2010), were used to show that the semantic link between the L2 and the conceptual store begins to strengthen with the development in L2 proficiency, such that eventually L1 mediation may not be necessary if a high enough level of proficiency is reached. Thus, highly proficient L2 learners have a similar representation mechanism as L1 speakers. However, for bilinguals with low proficiency, the link between L2 words and the conceptual store is weak, and L2 semantic access is accomplished *via* the activation of L1 counterparts. People with low L2 proficiency tend to be late bilinguals who acquired the L2 explicitly in a school context, in which L2 learning often takes place in a specific and limited setting, without the direct contact with events or entities described in language which takes place in L1 learning. Therefore, L2 words are claimed to have "less rich" semantic representations, i.e., they may be associated with fewer senses than L1 words or advanced L2 words. Such differences motivated us to consider whether, as with L1 processing, the sensorimotor system is also engaged in L2 processing for bilinguals with low proficiency.

Another issue still under debate is the embodied effect mechanism in L2 processing. Some studies have argued that the sensorimotor system is involved in the initial stages of L2 processing, as is the case in L1 processing. For example, using the Stroop task, Dudschig et al. (2014) found that action–sentence compatibility effects were triggered automatically for both implicit location words and emotion words in L2 processing. In a go/no-go paradigm, Buccino et al. (2017) discovered that the motor system was involved in processing nouns regarding graspable objects as compared with non-graspable ones. These studies suggested that L2 sensorimotor associations are automatic in nature and do not depend on deeper semantic task demands. However, in contrast to these accounts, studies on the simulation of language comprehension (Fischer and Zwaan, 2008; Kiefer and Pulvermüller, 2011) suggested that semantic information is kept in a distributed fashion in modality-specific sensory and motor areas. Hence, language processing is accomplished by the use of motor, perceptual, and emotional systems to simulate the situations described by the words or sentences. Thus, embodied effects in language comprehension might not reflect a process that is basic to language processing, but is rather the result of participants' conscious decision to imagine a described scene after they have already understood the meaning. As Meteyard et al. (2012) noted, perceptual and sensorimotor information is activated when semantic representations are accessed.

In sum, there are two main perspectives on the mechanism of embodied effects in language processing. One holds that the embodied effect found in language processing is the result of the processing of the physical attributes of words (Dudschig et al., 2014), while the other contends that semantic representation plays a crucial role in the embodied effect (Fischer and Zwaan, 2008).

Experiments

To disentangle the above-mentioned issues, a lexical decision task (LDT) and semantic category task (SCT) were implemented in the current study. In the LDT, participants were asked to decide whether the stimuli that appeared was a real word or not as quickly and accurately as possible. Due to the stimulus being presented very quickly (about 100 ms in duration), subjects did not have enough time to fully access the semantic information, and their judgments were mainly influenced by superficial attributes of words such as orthography, acoustic aspects, or word frequency. If embodied effects were found in the task, this would indicate that the effect occurs in the early stage of word processing. In the SCT, participants were asked to classify the stimuli presented. Owning to the stimuli being presented for about 800 ms, participants had enough time to access the semantics, and so according to the simulation of language comprehension, embodied effects would be found in the SCT.

In the LDT, participants are rapidly presented with some words and asked to decide whether the lexeme they saw was a true word or not; therefore, they may rely more on familiarity-based information (e.g., word frequency, orthography) to discriminate a word from a pseudo-word (Zunini and Renoult, 2017). In the SCT, participants are required to point out which category the word they saw belongs to; hence, they may have to determine the specific meaning of a word or at least require more access to semantic information than in the LDT to make a decision. Some studies have indicated that semantic information is not fully accessed in the LDT, meaning that participants' performance is mainly affected by shallow lexical factors such as word frequency, word length, and familiarity of words (Balota and Chumbley,

TABLE 1 The biographical data of the participants in low L2 proficiency group (SD).

Testing item	Rating scale (E1)	
	Experiment 1	Experiment 2
College English Test-Band 6 (CET6)	456.2 (35.8)	461.7 (36.4)
Oxford Placement Test (OPT)	40.45 (3.21)	40.65 (3.02)
Age of acquisition L2	9.38 (1.43)	9.35 (1.52)
Duration of L2 learning	11.34 (1.54)	11.36 (1.48)
Listening	2.85 (1.28)	2.98 (1.41)
Speaking	2.06 (1.32)	2.07 (1.37)
Reading	3.46 (1.35)	3.49 (1.41)
Writing	3.29 (1.57)	3.33 (1.55)

1985). If we found embodied effects in the LDT, this would mean that the sensorimotor system is also involved in the processing of superficial language components such as word frequency and orthography. If the embodied effect is derived from semantic access, then a strong embodied effect would be found in the SCT but not in the LDT, because participants need to deeply access the semantic information of words in order to succeed in the SCT.

Experiment 1

Methods

Participants

Forty-two native Chinese speakers (L1) took part in the experiment (16 male, $M_{age} = 21.98$, $SD_{age} = 1.24$); they were all late bilinguals with low proficiency in English (L2) and none of them had ever lived in an English-speaking country. The participants started learning the L2 at age 9–10. All the participants were right-handed, had normal or corrected-to-normal vision, and had no history of hearing or language difficulties or neurological/psychiatric impairment based on self-report. They signed a consent form before the experiment began and were paid for their participation. Ethical approval was given by the Committee of Protection of Subjects at Qufu Normal University. Participants were recruited based on four participant-selection criteria: duration of English language learning, College English Test-Band 6 (CET 6),[1] Oxford placement test,[2] and self-rating of L2 skills. The self-rating of L2 skills was based on a six-point scale assessment (ranging from 1 = "quite poor" to 6 = "highly proficient"). These tests were demonstrated to be valid measures of overall language proficiency (Hulstijn, 2012). Detailed biographical data of participants are presented in Table 1.

Materials

With reference to the existing research (Estes et al., 2007) and the purpose of the study, 60 English nouns (see Appendix) denoting different locations were used. Of them, 20 were "up" words (e.g., *roof*), 20 were "down" words (e.g., *root*), and 20 did not denote a location (e.g., *book*). Sixty pseudo-words were also selected, and were constructed by substituting one or two consonants or vowels in each noun (e.g., *griss* instead of *grass*). Through this procedure, the pseudo-words contained orthographically and phonologically permissible syllables in the English language. Words were controlled for frequency,[3] length, and typical location (on the vertical axis). For this purpose, 20 volunteers who had passed CET 6 rated 60 true nouns on a seven-point Likert scale (1 = "very down," 4 = "not sure about the location," 7 = "very up"). Words selected as "down" words had rating values smaller or equal than 2.2 ($M = 1.88$, $SD = 0.27$), words selected as "up" words had rating values equal or larger than 5.6 ($M = 6.02$, $SD = 0.4$), and words that did not denote a location had rating values around 3.6–4.5 ($M = 4.05$, $SD = 0.24$). The three categories of nouns did not differ significantly with regard to frequency, $F(2, 57) = 0.116$, $p = 0.891$, or length, $F(2, 57) = 0.42$, $p = 0.658$, but did differ significantly for the rated position, $F(2, 57) = 64.14$, $p < 0.001$.

Procedure and design

The experiment was conducted in a sound-attenuated and dimly illuminated experiment room. Participants sat comfortably in front of a PC screen (HP 21.5′ LCD, 1,920 × 1,080-pixel resolution, and 60 Hz refresh rate) at a distance of around 60 cm. Their task was to decide whether the target word was true or not as quickly as possible.

Figure 1A displays the experimental procedure. Each trial began with a centrally presented fixation cross for 500 ms. Then a target word appeared at the same location for 120 ms. After the target word presentation and 50 ms for masking stimuli, three "?" appeared in the same location as the target word, cuing participants to make decisions as quickly and as accurately as possible. If the participants perceived the target word to be true, they would press the up arrow in the top part of keypad with their right index finger, or if the target word was a non-word, then they would press the down arrow in the bottom part of keypad. If the participants did not respond within 3,000 ms, the cued signal would disappear. If they were inaccurate, the word "INCORRECT" appeared in red font for 500 ms. After an inter-trial interval (800 ms) the next trial started. Eight practice trials were conducted before four blocks consisting of 120 trials each (60 trials for true

1 The CET 6, designed by the Ministry of Education of China, is used in all universities in China to evaluate the English proficiency of non-English majors. It consists of tasks on listening comprehension, reading comprehension, vocabulary knowledge, grammar knowledge, and writing. The total score is 710, and the cutoff point (set by the Ministry of Education) for success and failure in the test is 427.

2 The Oxford Placement Test (OPT) includes 25 multiple-choice questions and a cloze test, and the total score is 50.

3 http://subtlexus.lexique.org

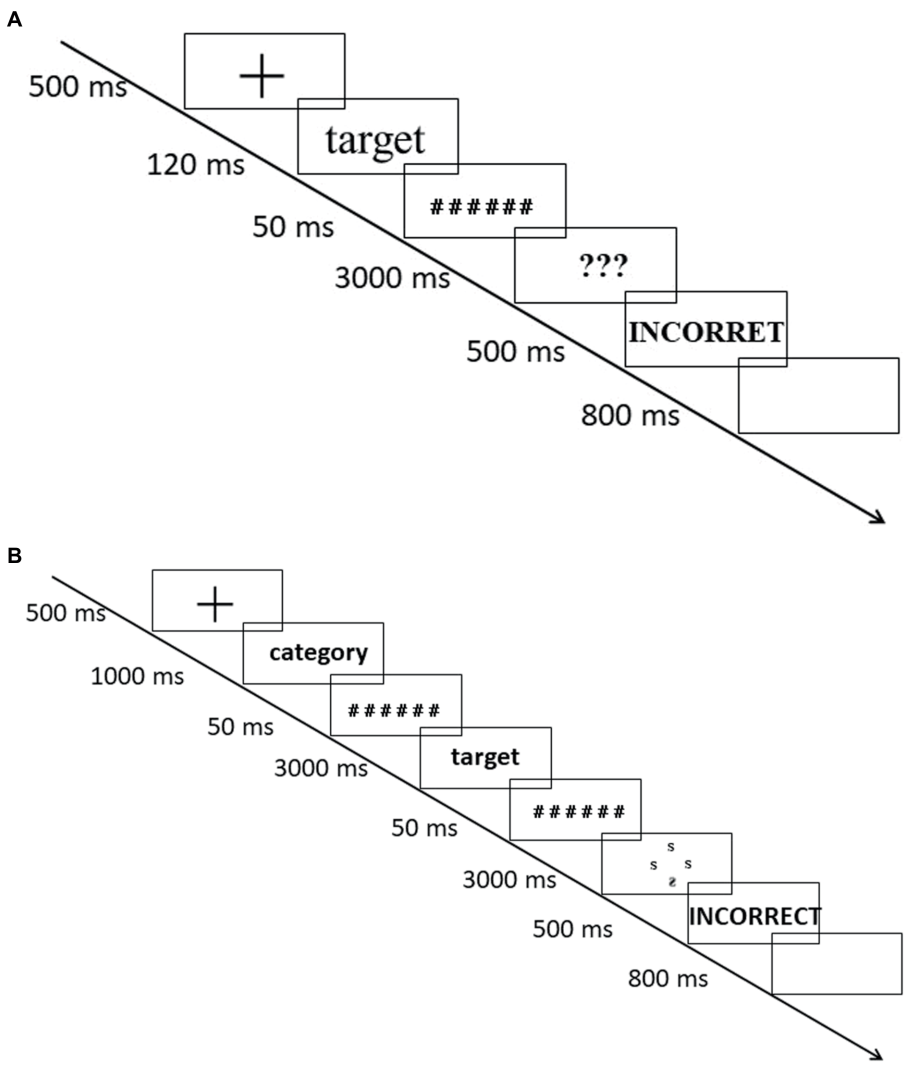


FIGURE 1
(A) The procedure of Experiment 1. **(B)** The procedure of Experiment 2.

words and 60 trials for non-words). Of the four blocks, two blocks required participants to press the up arrow on the keypad for true words, and the other two blocks required them to press the up arrow for non-words. Within each block, the order of presentation was randomized for each participant.

If L2-sensorimotor associations are automatic in nature and do not depend on deeper semantic task demands or L2 proficiency, cognitive advantages would be found in the congruent condition (e.g., seeing an "up" word and pressing the up arrow on the keypad) compared with the incongruent condition (e.g., seeing an "up" word and pressing the down arrow) or baseline condition (e.g., seeing a word that did not denote a location).

Results and discussion

The data from two participants were excluded due to the low accuracy of these participants (<80%). Erroneous trials and reaction times (*RTs*) out of 2.5 *SD* were not included in the analysis, reducing the dataset by 3.14%. The results are shown in Table 2.

RTs were analyzed with an ANOVA with word-direction and response-direction. There was no main effect of word-direction, $F(1, 39) = 1.64$, $p = 0.208$, $\eta_p^2 = 0.04$, but the main effect of response-direction was significant, $F(1, 39) = 10.83$, $p = 0.002$, $\eta_p^2 = 0.217$, where the response to press the up arrow was faster than pressing the down arrow (492.67 ms vs. 570.71 ms). The interaction between word-direction and response-direction was not significant, $F(1, 39) = 0.96$, $p = 0.33$, $\eta_p^2 = 0.024$. A paired t-test between the congruent condition (e.g., seeing an "up" word and pressing the up arrow) and incongruent condition (e.g., seeing an "up" word and pressing the down arrow) also indicated no significance, $t(81) = 1.08$, $p = 0.285$, Cohen' s $d = 0.055$. Analysis of accuracy showed that no significant differences were found, $FS < 1$.

TABLE 2 Reaction time (ms) and accuracy (%) in lexcial decision task (LDT).

Response	Reaction time			Accuracy		
	Up-pointing word	Down-pointing word	Irrespective location word	Up-pointing word	Down-pointing word	Irrespective location word
Up-response	493.76 (49.93)	491.59 (53.82)	514.94 (79.46)	90.19 (6.59)	91.31 (7.45)	90.50 (6.08)
Down-response	580.56 (84.13)	560.25 (54.53)	565.31 (73.73)	89.63 (7.31)	90.63 (7.27)	90.50 (7.25)

In sum, no differences were noted in Experiment 1, suggesting that L2 word decisions could not activate spatially directed motor response automatically for late L2 learners with low proficiency. The results are inconsistent with findings from Dudschig et al. (2014). Using the vertical Stroop paradigm and similar materials, they found that L2 processing automatically activated motor responses similar to L1 processing for late L2 learners; participants were not required to actively read or evaluate word meaning. They concluded that L2 sensorimotor associations are automatic in nature and do not depend on deeper semantic task demands. Although participants in Dudschig et al.'s (2014) study were late L2 learners, they may have been highly proficient in the L2 (especially at the level of written word identification), because English shares more similarities with German, compared with Chinese and English. Current cumulative evidence arguing in favor of the involvement of the sensorimotor system in L2 processing have mainly involved semantic tasks. Thus, a further investigation was conducted in Experiment 2 using a semantic category decision task. If the L2 sensorimotor system is the result of accessing semantic representations, then cognitive advantages would be found in congruent conditions compared with incongruent conditions or baseline conditions.

Experiment 2

Methods

Participants

Forty native Chinese speakers (L1) took part in the experiment (12 male, $M_{age} = 21.80$, $SD_{age} = 1.32$); they were all late bilinguals with low proficiency in English (L2) and none of them had ever lived in an English-speaking country or taken part in Experiment 1. Participants had started learning the L2 at age 9–10. All the participants were right-handed, had normal or corrected-to-normal vision, and had no history of hearing or language difficulties or neurological/psychiatric impairment based on self-report. The other procedure was the same as in Experiment 1. Detailed biographical data of the participants are presented in Table 1.

Materials

Sixty true nouns, as used in Experiment 1, were mainly from six categories: nature entities (e.g., *sun*/*cloud*/*river*), living entities or organisms (e.g., *bird*/*leg*/*seed*/*grass*/*root*), household items (e.g., *hat*/*glass*/*cake*/*wallet*), buildings (e.g., *roof*/*ceiling*/*tomb*/*tunnel*/*well*/*cave*), food (e.g., *bread*/*soap*/*cookie*), and aircraft or vehicle (e.g., *plane*/*kite*/*wheel*).

Procedure and design

Figure 1B displays the experimental procedure. Each trial began with a centrally presented fixation cross for 500 ms, and then a cue of the category (e.g., natural entities) appeared at the same location for 1,000 ms. After cue presentation and 50 ms for the pre-masked stimuli, the signal target (e.g., *sun*) appeared at the same location for 800 ms. After 50 ms of post-masked stimuli, three normal "S" and an inverted "S" arranged into a cross (S S S S) appeared in the center of the frame. If participants considered the target word to be from the category presented, they would search for the inversed "S" by pressing the corresponding arrow on the keypad. If the inversed "S" was at the bottom of the screen, participants would press the down arrow, and they would press the up arrow if the inversed "S" was at the top of the screen. If the target word was not from the category presented, participants did not need to respond and the stimulus would disappear in 3,000 ms. Eight practice trials were performed before four blocks consisting of 120 trials each. The location of the inverted "S," participants' response, and the matching of the target word and its category were well counterbalanced. Within each block, the order of presentation was randomized for each participant.

Results and discussion

Due to exploring only the embodied effects in upward or downward location word processing in our study, the data of left or right trials were not collected. Erroneous trials and reaction times (*RT*s) for "up" or "down" trials out of 2.5 *SD* were excluded from the analysis, reducing the data set by 2.42%. The results are shown in Table 3.

*RT*s were analyzed with an ANOVA with word-direction and response-direction. The main effect of word-direction was significant, $F(2, 78) = 17.86$, $p < 0.001$, $\eta_p^2 = 0.31$. Planned comparisons showed that there was a significant difference between the congruent (upward-location words and pressing the up arrow or downward-location words and pressing the down arrow) and the incongruent condition, $t(117) = -4.83$, $p < 0.001$, and the difference between the incongruent condition and baseline condition (words not denoting a location and pressing the up or down arrow) was also significant, $t(117) = 2.98$, $p < 0.01$,

10.3389/fpsyg.2022.980967

TABLE 3 Reaction time (ms) and accuracy (%) in semantic category decision (SCT).

Response	Reaction time			Accuracy		
	Up-pointing word	Down-pointing word	Irrespective location word	Up-pointing word	Down-pointing word	Irrespective location word
Up-response	891.39 (94.05)	1109.88 (78.18)	987.56 (42.89)	97.46 (2.12)	97.75 (2.49)	96.31 (4.47)
Down-response	955.37 (51.47)	1027.64 (84.13)	970.20 (75.44)	96.75 (3.17)	96.68 (3.53)	97.00 (2.33)

but the difference between the congruent and baseline condition was marginal, $t(117) = 1.85$, $p = 0.07$. The main effect of response-direction was not significant, $F < 1$. The interaction between word-direction and response-direction was significant, $F(2, 78) = 3.56$, $p < 0.05$, $\eta_p^2 = 0.08$. Simple effects analyses indicated that searching for the inversed "S" at the top of the screen following "up" words was faster than searching for the inversed "S" at the top of the screen following "down" words, $F(1, 39) = 3.31$, $p < 0.05$, $\eta_p^2 = 0.07$, and searching for the inversed "S" at the bottom of the screen following "down" words was faster than that following "up" words, $F(1, 39) = 11.18$, $p < 0.01$, $\eta_p^2 = 0.27$. Analyses of the accuracy suggested that there were no significant differences for the main and interaction effects, $FS < 1$.

The results of Experiment 2 strongly suggested that location information was indeed activated in the SCT, since responses were faster when the word's referent location in the world was compatible with the participant's response movement. This replicates previous findings regarding the effects of implicit location words and indicates that spatial experiential traces are activated in various tasks involving semantic retrieving (Parker Jones et al., 2012; Vukovic and Shtyrov, 2014; Vukovic and Williams, 2014; Buccino et al., 2017). Using the semantic judgment task, Qian (2016) investigated the embodied effects in processing nouns with high or low power, and the results showed that responses were faster for power words presented in the upper (vs. lower) part of the screen even for L2 speakers with low proficiency.

General discussion

Studies have increasingly suggested that L2 processing relies on embodied representations of meaning and is connected to motor and perceptual processing, as is found in L1 processing (De Grauwe et al., 2014; Dudschig et al., 2014; Buccino et al., 2017; Gianelli et al., 2018). However, there remained some unresolved issues, for example, whether the sensorimotor system is automatically involved in L2 processing or not. The question remained as to the role of language proficiency in embodiment effects. In the current study, two experiments were conducted using the LDT and SCT. The results showed that, in the LDT, a task where participants are more likely to make judgments based on physical properties of words (e.g., familiarity, orthography), "up" words did not result in faster upward than downward responses, and "down" words also did not result in faster downward than upward responses. In the SCT, compatibility effects were found; specifically, searching for the target located at the top of the screen after "up" words was faster than after "down" words and searching for the target at the bottom of the screen after "down" words was faster than after "up" words. Hence, we concluded that L2-sensorimotor association, at least for L2 speakers with low proficiency, was not automatic in nature and did depend on deeper semantic task demands (Qian, 2016).

Increasing evidence has suggested that L2 processing is also based upon modal experiences, and is not separate from the sensory system (Dudschig et al., 2014; Buccino et al., 2017); however, it remained open whether the sensorimotor system is involved in L2 processing or not (Monaco et al., 2019). Some studies argued that L2 processing is "disembodied," and considerable differences between L1 and L2 were noted in the literature such as age of acquisition (AOA), style of learning, and proficiency. Using the LDT, motor and non-motor cognate or non-cognate verbs in Dutch were presented to participants with highly proficient L1-German L2-Dutch and Dutch native speakers. The results indicated a significantly stronger activation in motor and somatosensory areas for motor verbs, regardless of the cognate status of the verbs. This was true of both language groups. De Grauwe et al. (2014) consequently suggested L2 representations to be rich enough to activate similar motor-related areas to L1. However, in contrast to their findings, we did not find an embodiment effect in our experiment using the LDT. One of the reasons for the difference between the two studies may relate to the participants. In our experiment, all the participants were of low proficiency and were late L2 learners. Unlike advanced L2 learners who have similar semantic representations to L1 speakers, L2 learners with low proficiency have a coarse semantic representation, in which the concept representation is detached from the sensorimotor system and environments; thus, some scholars have argued that L2 processing, especially for L2 learners with low proficiency, is "disembodied" (Pavlenko and Aneta, 2017). Another reason may be the research method. Compared with the *f*MRI method used in De Grauwe et al.'s study, the behavior method was not sensitive enough to detect the involvement of the sensorimotor system in L2 processing. A third reason may come from the similarity between the L1 and L2. In the *f*MRI study, German and Dutch are highly related languages with a large number of cognates, i.e., words with similar form and meaning in the two languages; thus, there was a high overlap between L1 and L2 representation. In our study, Chinese and

English are from different language families, meaning that few common words were shared between them.

In the SCT, a stronger embodiment effect was discovered, which was consistent with many other findings. Using the picture-word mapping paradigm, Bergen et al. (2010) found both groups (L1 and L2) responded more rapidly when the picture and the word were matched. In another study (Xue et al., 2015) using EEG techniques, written high and low BOI (body-object interaction) words embedded in segmented sentences characterized by rich and poor sensorimotor context were presented, and participants were asked to make judgments about the acceptability of the sentences. The results showed that action-and perception-related brain areas for L2 words were activated, which indicated that the semantic representations for L2 are plentiful enough for sensorimotor-related activation. Combining the results of the two experiments and other studies, we concluded that the sensorimotor system is also involved in L2 processing, even for L2 learners with low proficiency.

Another key issue subject to serious debate in L2 processing is whether the sensorimotor system is involved automatically in L2 processing or is the result of consciously imagining a described scene after accessing lexical-semantic information. Some studies, such as Dudschig et al. (2014) and Buccino et al. (2017), held that the sensorimotor system is involved automatically in L2 processing and does not depend on deeper semantic task demands. Our results are not consistent with their findings because no embodiment effect was found in the LDT, in which participants made decisions according to the familiarity or form of the word, without needing to access semantic information. The sensorimotor system is automatically involved in L1 processing because, as is widely known, L1 is learned interactively and we often perceive the events and entities or feelings described. Thus, language percepts are typically combined with specific gestures, eye movements, and physical orientation toward the described entity. When seeing a target word, the sensorimotor system, events, and feelings associated with the word would be automatically activated. In contrast, L2 learning in school typically takes place in a very limited setting, whereby interactions with other people and physical experiences are less dominant during the object of inquiry, which is in large part an internal, mental process. In such a view of L2 learning, there is a division between the mind and world, especially for L2 speakers with low proficiency; thus, it is difficult to automatically activate the sensorimotor system or the referents in L2 processing. In contrast, in the SCT, a task which entailed deeper semantic processing, a strong embodiment effect was found. Taken together, we contend that the embodiment effect in L2 processing is not an automatic consequence, but the result of participants consciously imagining a described scene after they have already understood the meaning.

The following question emerges: Why is the sensorimotor system or experiential trace activated automatically in L2 processing for advanced L2 learners, but not for L2 learners with low proficiency? According to RHM (Kroll and Stewart, 1994), when L2 is still emerging, L1 mediates L2 access to the conceptual store. If this is the case, then a large amount of time would be needed to activate the sensorimotor experiential associations, because the equivalent L1 lexeme would be first retrieved, followed by access to the concept *via* the L1, and this would entail later sensorimotor involvement compared with advanced L2 learners. With the development of L2 proficiency, a semantic link begins to strengthen between the L2 and the "conceptual store," such that, eventually, L1 mediation may not be necessary if a high enough level of proficiency is achieved and it is possible to activate the sensorimotor experiential associations instantly by accessing the conceptual store. Emerging research has demonstrated that embodiment processes occur similarly to L1 for highly proficient bilinguals, but may differ in some ways for less proficient L2 learners.

In our opinion, our findings have significant implications for debates both in embodied cognition and bilingual processing. First, although there is increasing evidence indicating that L2 comprehension is achieved by recruiting the very same resources which are used for action, perception, and emotion, most of these studies focused on advanced L2 learners, meaning it was difficult to conclude that the embodiment effect was universal in L2 processing. Our findings gave support to the statements mentioned above. Second, the embodied experience and sensorimotor system should be taken into account while constructing models of L2 representation. Third, if sensorimotor experience, emotion, and settings of lexical learning are essential to L2 processing, teachers should adopt experience-based teaching methods and encourage learners to use bodies, actions, imaginaries, and settings in L2 learning. Of course, there are still some limitations to our study. First, two different tasks were used in Experiment 1 and Experiment 2, which made it difficult to compare the results between the two experiments. Second, the absence of an embodiment effect in Experiment 1 may be due to an insufficiently sensitive approach. In future research, we will use a more sensitive approach to examine the embodied effects in L2 processing, such as by using ERPs and *f*MRI.

In conclusion, the present findings have demonstrated that the sensorimotor system is also involved in less advanced L2 processing, but this outcome was only found in the SCT and not the LDT. This suggests that the sensorimotor system is not automatically involved in L2 processing, but is the result of consciously simulating the referents after accessing semantic information.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Center of Experiment, Qufu Normal University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YB performed the experiment, wrote the manuscript, and analyzed the data. WH proposed the research idea and experimental design. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the Natural Science Fund of Shandong Province to WH (grant number: ZR2019MC048) and Higher Education Youth Innovation Science and Technology Support Program of Shandong Province (2019RWF005).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2022.980967/full#supplementary-material

References

Balota, D. A., and Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: lexical access and/or production? *J. Mem. Lang.* 24, 89–106. doi: 10.1016/0749-596X(85)90017-8

Barsalou, L. W. (1999). Perceptual symbol systems. *Behav. Brain Sci.* 22, 577–660. doi: 10.1017/S0140525X99002149

Bergen, B., Lau, T. T. C., Narayan, S., Stojanovic, D., and Wheeler, K. (2010). Body part representation in verbal semantics. *Mem. Cogn.* 38, 969–981. doi: 10.3758/MC.38.7.969

Bergen, B. K., Lindsay, S., Matlock, T., and Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cogn. Sci.* 31, 733–764. doi: 10.1080/03640210701530748

Birba, A., Beltrán, D., Caro, M. M., Trevisan, P., Kogan, B., Sedeño, L., et al. (2020). Motor-system dynamics during naturalistic reading of action narratives in first and second language. *NeuroImage* 216:116820. doi: 10.1016/j.neuroimage.2020.11682

Buccino, G., Marino, B. F., Bulgarelli, C., and Mezzadri, M. (2017). Fluent speakers of a second language process graspable nouns expressed in L2 like in their native language. *Front. Psychol.* 8:1306. doi: 10.3389/fpsyg.2017.01306

De Grauwe, S., Willems, R. M., Rueschemeyer, S. A., Lemhöfer, K., and Schriefers, H. (2014). Embodied language in first- and second-language speakers: neural correlates of processing motor verbs. *Neuropsychologia* 56, 334–349. doi: 10.1016/j.neuropsychologia.2014.02.003

Desai, R. H., Choi, W., Lai, V. T., and Henderson, J. M. (2016). Toward semantics in the wild: activation to manipulable nouns in naturalistic reading. *J. Neurosci.* 36, 4050–4055. doi: 10.1523/JNEUROSCI.1480-15.2016

Dudschig, C., de la Vega, I., and Kaup, B. (2014). Embodiment and second-language: automatic activation of motor responses during processing spatially associated L2 words and emotion L2 words in a vertical Stroop paradigm. *Brain Language* 132, 14–21. doi: 10.1016/j.bandl.2014.02.002

Estes, Z., Verges, M., and Barsalou, L. W. (2007). Head up, foot down: object words orient attention to the objects' typical location. *Psychol. Sci.* 19, 93–97.

Felisatti, A., Ranzini, M., Blini, E., Lisi, M., and Zorzi, M. (2022). Effects of attentional shifts along the vertical axis on number processing: an eye-tracking study with optokinetic stimulation. *Cognition* 221:104991. doi: 10.1016/j.cognition.2021.104991

Fischer, M. H., and Zwaan, R. A. (2008). Embodied language: a review of the role of the motor system in language comprehension. *Q. J. Exp. Psychol.* 61, 825–850. doi: 10.1080/17470210701623605

Foroni, F. (2015). Do we embody second language? Evidence for 'partial' simulation during processing of a second language. *Brain Cognition.* 99, 8–16. doi: 10.1016/j.bandc.2015.06.006

Gianelli, C., Kühne, K., Mencaraglia, S., and Volta, R. D. (2018). Action processing in the motor system: TMS evidence of shared mechanisms in the visual and linguistic modalities. *PsyArXiv* [Preprint]. doi: 10.31234/osf.io/gw4v9

Glenverg, A. M., and Kaschak, M. P. (2002). Grounding language in action. *Psychomnomic Bull. Rev.* 9, 558–565.

Grainger, J., Midgley, K., and Holcomb, P. J. (2010). "Re-thinking the bilingualinteractive-activation model from a developmental perspective (BIA-d)," in *Language Acquisition Across Linguistic and Cognitive Systems*. eds. M. Kail and M. Hickmann (New York: John Benjamins), 267–284.

Horton, W. S., and Rapp, D. N. (2003). Out of sight, out of mind: occlusion and the accessi-bility of information in narrative comprehension. *Psychon. Bull. Rev.* 10, 104–110. doi: 10.3758/BF03196473

Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Biling. Lang. Congn.* 15, 422–433. doi: 10.1017/S1366728911000678

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M. E., Blanchard, A. A., et al. (2005). Perception of motion affects language processing. *Cognition* 94, B79–B89. doi: 10.1016/j.cognition.2004.06.005

Kiefer, M., and Pulvermüller, F. (2011). Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. *Cortex* 48, 805–825. doi: 10.1016/j.cortex.2011.04.006

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychol. Rev.* 95, 163–182.

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174. doi: 10.1006/jmla.1994.1008

Kühne, K., and Gianelli, C. (2019). Is embodied cognition bilingual? Current evidence and perspectives of the embodied cognition approach to bilingual language processing. *Front. Psychol.* 10:108. doi: 10.3389/fpsyg.2019.00108

Meteyard, L., Cuadrado, S. R., Bahrami, B., and Vigliocco, G. (2012). Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex* 48, 788–804. doi: 10.1016/j.cortex.2010.11.002

Monaco, E., Jost, L. B., Gygax, P. M., and Annoni, J. M. (2019). Embodied semantics in a second language: critical review and clinical implications. *Front. Hum. Neurosci.* 13:110. doi: 10.3389/fnhum.2019.00110

Parker Jones, Ō., Green, D. W., Grogan, A., Pliatsikas, C., Filippopolitis, K., Ali, N., et al. (2012). Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cereb. Cortex* 22, 892–902. doi: 10.1093/cercor/bhr161

Pavlenko, and Aneta. (2017). Do you wish to waive your rights? Affect and decision-making in multilingual speakers. *Curr. Opin. Psychol.* 17, 74–78. doi: 10.1016/j.copsyc.2017.06.005

Qian, W. (2016). Embodied cognition processing and representation of power words by second language learners with different proficiency levels. *Chinese J. Appl. Ling.* 39, 484–494. doi: 10.1515/cjal-2016-0030

Ratcliffe, J., and Tokarchuk, L. (2020). "Evidence for embodied cognition in immersive virtual environments using a second language learning environment," in *2020 IEEE Conference on Games (CoG).* IEEE.

Stanfield, R. A., and Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychol. Sci.* 121, 153–156. doi: 10.1111/1467-9280.00326

Vukovic, N., Feurra, M., Shpektor, A., Myachykov, A., and Shtyrov, Y. (2017). Primary motor cortex functionally contributes to language comprehension: an online rTMS study. *Neuropsychologia* 96, 222–229. doi: 10.1016/j.neuropsychologia.2017.01.025

Vukovic, N., and Shtyrov, Y. (2014). Cortical motor systems are involved second-language comprehension: evidence from rapid mu-rhyth desynchronisation. *NeuroImage* 102, 695–703.

Vukovic, N., and Williams, J. N. (2014). Automatic perceptual simulation of first language meanings during second language sentence processing in bilinguals. *Acta Psychol.* 145, 98–103. doi: 10.1016/j.actpsy.2013.11.002

Xue, J., Marmolejo-Ramos, F., and Xuna, P. (2015). The linguistic context effects on the processing of body-object interaction words: an ERP study on second language learners. *Brain Res.* 1613, 37–48. doi: 10.1016/j.brainres.2015.03.050

Zunini, R. L., and Renoult, L. (2017). N400 effects of semantic richness can be modulated by task demands. *Neuropsychology* 31, 277–291. doi: 10.1037/neu0000327

Zwaan, R. A., Stanfield, R. A., and Yaxley, R. H. (2002). Language comprehenders mentally represent the shape of objects. *Psychol. Sci.* 13, 168–171. doi: 10.1111/1467-9280.00430