Orthographic, Semantic, and Contextual Influences on Initial Processing and Learning of Novel Words During Reading: Evidence From Eye Movements

Wei YI & Shiyi LU
Peking University
Robert DEKEYSER
University of Maryland, USA

Abstract

This study investigates how orthographic, semantic and contextual variables—including word length, concreteness, and contextual support—impact on the processing and learning of new words in a second language (L2) when first encountered during reading. Students learning English as a foreign language (EFL) were recruited to read sentences for comprehension, embedded with unfamiliar L2 words that occurred once. Immediately after this, they received a form recognition test, a meaning recall test, and a meaning recognition test. Eye-movement data showed significant effects of word length on both early and late processing of novel words, along with effects of concreteness only on late-processing eye-tracking measures. Informative contexts were read slower than neutral contexts, yet contextual support did not show any direct influence on the processing of novel words. Interestingly, initial learning of abstract words was better than concrete words in terms of form and meaning recognition. Attentional processing of novel L2 words, operationalized by total reading time, positively predicted L2 learners' recognition of new orthographic forms. Taken together, these results suggest: 1) orthographic, semantic and contextual factors play distinct roles for initial processing and learning of novel words; 2) online processing of novel words contributes to L2 learners' initial knowledge of unfamiliar lexical items acquired from reading.

Keywords: word processing/learning, word length, concreteness, contextual support, eye tracking

1. Introduction

Lexical knowledge plays a critical role in the development of second language skills. L2 learners are able to acquire new words under different learning conditions and through various input modalities, either in or outside the classroom. Despite this, researchers and

ISSN 2192-9505 Chinese J. of Appl. Ling. 45-2 (2022), pp. 194-219 DOI 10. 1515/CJAL-2022-0203

practitioners have widely acknowledged that vocabulary learning is especially challenging for adult L2 speakers (Schmitt, 2008). Vocabulary learning is inextricably related to the way lexical items are processed and internalized by language learners (Pirrelli et al., 2020). Specifically, acquiring new words means establishing form-meaning mappings for novel lexical items and storing such knowledge in memory, thus involving the integration of orthographic, semantic, and contextual information in real time. During the past decades, a large body of research in the field of second language acquisition (SLA) has been carried out to investigate L2 vocabulary learning. Nevertheless, few studies have examined how orthographic, semantic and contextual features influence L2 learners' processing and acquisition of unfamiliar lexical items, as well as how online processing of novel words is linked to the acquisition of vocabulary knowledge. On the one hand, vocabulary learning is an incremental process, with multiple encounters generally required for L2 learners to fully acquire new words. On the other hand, outside lab and classroom settings, it is not always realistic for L2 learners to encounter a novel word repeatedly. Meanwhile, under some circumstances, L2 learners may be able to acquire novel words with only one exposure (Nagy & Anderson, 1984). Recent years have witnessed a growing interest in exposure frequency for L2 vocabulary learning, yielding accumulating evidence supporting its positive influence for acquiring new words, especially under incidental learning conditions (for a meta-analysis, see Uchihara et al., 2019). However, little effort has been made to reveal how novel L2 words are processed and learned when they are first encountered during reading.

The present study designed an eye-tracking experiment in which Chinese learners of English read novel L2 words embedded in sentences. Three types of intralexical and contextual characteristics—namely, word length, concreteness, and contextual support—were manipulated, targeting orthographic, semantic and contextual influences for vocabulary processing and learning respectively. The purpose of this study is threefold. First, we would like to examine how word length, concreteness and contextual support impact L2 learners' processing of novel words in real time. Particularly, by incorporating early- and late-processing eye-tracking measures, we would like to examine whether their influences differ in the time course. Second, we would also like to investigate whether L2 learners' initial knowledge of novel words, measured by receptive and productive offline tests, can be predicted by word length, concreteness, and contextual support. Finally, by combining online (i. e., eye movements) and offline (i. e., test scores) measures, we aim to relate L2 learners' acquisition of different components of vocabulary knowledge to their online processing behavior.

1.1 Orthographic, Semantic, and Contextual Influences on Word Processing and Learning

According to Gibson and Levin (1975, as cited in Laufer, 1990), a word is a composite representation consisting of graphic, phonological, orthographic, semantic, and syntactic information. Over the past decades, research evidence has accumulated, supporting the influences of sublexical, lexical and contextual features on visual word recognition (for a review, see Balota, 1994). When it comes to vocabulary learning, such factors are supposed to moderate the level of difficulty for L2 learners (for a review, see Laufer, 1990). Vocabulary learning is a multi-faceted process. In order to acquire a word from reading, we must not only recognize its orthographic form, but also develop a detailed semantic representation of its meaning, with the aid of contextual support. From this perspective, processing and acquisition of unfamilar words are closely related, both involving interactions in orthographic, semantic and contextual information. In this section, effects of word length, concreteness and contextual support on word processing and learning—targeting orthographic, semantic and contextual influences respectively—are briefly reviewed.

1.1.1 Word Length

Word length is a factor known to affect the perceptual processing of words. Word length can be defined by the number of phonemes/syllables or letters. In the present study, we are primarily interested in word length as an orthographic measure, operationalized by the number of letters. A large number of studies have reported effects of word length for visual word recognition, using techniques such as lexical decision, naming, and eye tracking (for a review, see Barton et al., 2014; New et al., 2006). Overall, word length exerts an inhibitory influence on lexical processing, with more response time needed for longer words than shorter ones. Nonetheless, contrasting patterns of word length effects have also been reported. For instance, based on multiple regression analyses of lexical decision data, New and colleagues (New et al., 2006) found that the effect of word length on lexical decision latency was inhibitory for words of 8-13 letters, but facilitatory for words of 3-5 letters. Word length is one of the main predictors for fixation durations (Kliegl et al., 2006). Effects of word length in reading have been well documented by many eye-tracking studies, with words consisting of more letters receiving longer reading times than shorter words (e.g., Calvo & Meseguer, 2002; Hyönä & Olson, 1995; Juhasz & Rayner, 2003; Kliegl et al., 2004; McDonald, 2006). Compared with shorter words, longer words are also less likely to be skipped (e.g., Calvo & Meseguer, 2002; Kliegl et al., 2004) and receive more fixations (e.g., Calvo & Meseguer, 2002; Hyönä & Olson, 1995; McDonald, 2006). Current evidence (Lowell & Morris, 2014) suggests that the effects of word length for novel and known words may be comparable in terms of single-fixation measures (e.g., the first fixation duration, skipping), yet such effects are significantly larger for novel words than for known words, with respect to multiple-fixation measures, such as the gaze duration.

In the field of short-term memory research, word length effect—that is, serial recall for short words is better than that for long words—has been regarded as one of the benchmark findings (Baddeley et al., 1975). Many studies on L2 vocabulary learning have also revealed similar advantages of short words (for a review, see Laufer, 1990). Hiebert and colleagues (Hiebert et al., 2019) found that number of letters negatively impacted the performance of both L1 and L2 speakers of English in vocabulary knowledge tests. Similarly, Willis and Ohashi (2012) reported that the number of letters correlated with vocabulary learning difficulty for adult Japanese learners of English. In a more recent study, Godfroid et al. (2018) confirmed that shorter words were acquired significantly better than longer words, in terms of form recognition, meaning recognition, and meaning recall.

1.1.2 Concreteness

Lexical processing and vocabulary learning are influenced by various semantic variables, among which concreteness cannot be overestimated. Concreteness represents a fundamental distinction in the way concepts are represented in the mental lexicon (Palmer et al., 2013). By definition, concreteness evaluates the degree to which a concept refers to a perceptible entity (Brysbaert et al., 2014). The concrete/abstract dimension has been one of the heated topics in word recognition, with considerable research demonstrating that concrete words are processed faster than abstract words (for a review, see Schwanenflugel, 1991). In terms of lexical processing during reading, Juhasz and Rayner (2003) found that concrete words received shorter fixation durations (the gaze duration and total reading time) than abstract words. Loomis (2010) had native speakers of English read concrete and abstract words presented in sentences. She reported that concrete words received significantly shorter single fixation duration, first fixation duration, and gaze duration. To account for such processing advantages of concrete words, two groups of theories are currently available. The first group assumes a quantitative difference between concrete and abstract words. For example, according to the dual-coding theory (Paivio, 1991), compared with abstract words, concrete words have more sensory referents and activate perceptual memory codes in addition to verbal codes. An alternative explanation is provided by Schwanenflugel (Schwanenflugel, 1991), which argues that concrete words' processing advantage arises from greater availability of supporting context information. On the contrary, a qualitative-difference account has been developed by Crutch and colleagues (Crutch & Warrington, 2005). According to them, abstract words are represented in an associative neural network, whereas concrete words are organized in a categorical way based on semantic similarity.

Concreteness effects also apply to vocabulary learning. As has been widely reported, concrete words are acquired much easier than abstract words. Mestres-Missé and colleagues (Mestres-Missé et al., 2014) found that native Spanish speakers acquired new meanings of concrete L1 words faster than those of abstract words, when matched for context availability. Similar effects have also been observed by SLA research on vocabulary learning. Using a paired-associate training technique, de Groot and Keijzer (2000) discovered that concrete words were learned more easily and were less susceptible to forgetting than abstract words. Elgort and Warren (2014) asked L2 English participants to read a long text for meaning, in which target pseudowords were embedded. Unsurprisingly, concrete pseudowords were acquired better than abstract pseudowords in terms of meaning production.

1.1.3 Contextual Support

Words encountered by language users are often presented in discourse contexts that vary in terms of the level of semantic support for them. Research has shown that readers may utilize contextual information to predict and activate phonological, morphological, or semantic properties of specific words (Brothers et al., 2015), resulting in the facilitation of lexical retrieval. Many studies have reported that when processing familiar words, those embedded in predictable contexts are more likely to be skipped and receive shorter fixation durations (e. g., Calvo & Meseguer, 2002; Kliegl et al., 2004). Conversely, the influence of contextual support on the processing of novel words is less consistent. Chaffin et al. (2001) explored how native English speakers establish the meaning of monomorphemic pseudowords, which were followed by sentential contexts that varied in informativeness about their meaning. They found that contextual informativeness had no influence on the reading time for pseudowords. Moreover, readers spent significantly more time (i.e., the gaze duration, total reading time) processing informative sentential contexts than neutral ones. Using a design slightly different from Chaffin et al. (2001), Brusnighan and Folk (2012) examined L1 English speakers' processing of novel compound words that were preceded either by informative or neutral sentential contexts. They found contrasting patterns between influences of contextual support on semantically transparent and opaque words. Participants reread novel opaque words significantly more often in informative contexts than in neutral contexts. By contrast, they spent significantly less time rereading novel transparent words in informative contexts than in neutral contexts.

Vocabulary learning is highly contextualized. Graves (1986) estimated that English students acquire on average between 1,000 and 5,000 words from context each school year. Discourse contexts surrounding words provide semantic clues for language learners to infer their meaning. With high-quality contextual support, one exposure may suffice for learners to derive the meaning of novel words (Nagy & Anderson, 1984). Mulder and colleagues (Mulder et al., 2019) investigated the influence of contextual support on L2 vocabulary learning. Dutch learners of English were exposed to L2 words embedded in sentences with different degrees of contextual support and asked to judge the plausibility of each sentence, after being presented with Dutch translations of the target words. As reported by them, vocabulary knowledge measured by a translation task was better for words placed in sentences with stronger contextual support. In a study by Ma and colleagues (Ma et al., 2015), Chinese learners of English were asked to read sentences with either high or low contextual constraint for the interpretation of target L2 pseudowords. Participants' acquisition of vocabulary knowledge was measured by a semantic relatedness judgment task. Results showed that pseudowords embedded in more constraining contexts were acquired better than those in less constraining sentences. Webb (2008) examined the effects of context on incidental L2 vocabulary learning. Japanese EFL students encountered 10 target words in short contexts consisting of one to two sentences while involved in reading comprehension. The sentential contexts varied in the amount of information available to infer the meaning of target words. Consistent with Mulder et al. (2019) and Ma et al. (2015), participants exposed to target words with more contextual clues achieved significantly better performance on meaning recall and recognition tests.

1.1.4 Interactions Among Word Length, Concreteness, and Contextual Support

Both word processing and vocabulary learning are influenced by orthographic, semantic, and contextual variables. In addition to examining the independent contributions of word length, concreteness and contextual support, it would be theoretically interesting to investigate how these factors interact with each other during the processing and learning of words. Calvo and Meseguer (2002) reported that the contribution of word length to fixation durations was independent of contextual support. Juhasz and Rayner (2003) investigated the influences of word length and concreteness on fixation durations during reading, yet they did not find any interaction effect between these variables. Loomis (2010) found no interaction between concreteness and context for fixation durations, concluding that concrete words were processed faster than abstract words regardless of contextual support. When it comes to vocabulary learning, Mulder et al. (2019) is the only study we found that investigated the interaction between word length and contextual support. Their results showed that Dutch learners of English benefited from contextual support regardless of the length of L2 words. As far as we know, not a single study has examined the three-way interaction between word length, concreteness and contextual support, neither for lexical processing nor vocabulary learning.

1.2 Initial Learning of Novel Words Given One Exposure

Vocabulary learning is an incremental process (Barclay & Schmitt, 2019; Fukkink et al., 2001; Schmitt, 1998). Generally speaking, repeated exposures are needed for L2 learners to acquire a word. The positive role of repeated exposures for vocabulary learning has been supported by many SLA studies, with a recent meta-analysis (Uchihara et al., 2019) concluding that frequency of encounters has a medium effect (r = .34) on incidental vocabulary learning. Conversely, little research has been done to examine how novel L2 words are processed and learned when they are first encountered. Outside laboratory and classroom settings, L2 learners do not always have the opportunity to encounter a new word repeatedly. Current studies (e.g., Bisson et al., 2014) have suggested that the effects of repeated exposures for L2 vocabulary acquisition are not constant, with the impact of the first few exposures on vocabulary learning being larger than that of the later exposures. In the field of child language acquisition, researchers generally separate the cumulative process of word learning into two phases, starting with "fast mapping" after the first exposure to novel words, then followed by slow, continuous word learning (Kan & Kohnert, 2012). Nagy and Anderson (1984) estimated that approximately 5% to 12% of the words acquired by L1 English children are learned from a single exposure. Researchers (e.g., Hu, 2012) have also found that L2 children are able to acquire the meaning of new words after one or two exposures. Similarly, a small number of studies (e.g., Borovsky et al., 2012) have claimed that adult L1 speakers can fast-map novel words to their semantically related word meanings after encountering them only once in highly constraining contexts, using sensitive measures such as event-related potentials. Webb (2007) investigated Japanese learners' incidental learning of pseudowords in place of familiar English concepts from reading. He reported that participants' immediate vocabulary knowledge after one exposure to the target items ranged from around 4% to nearly 60%, measured by ten tests targeting various aspect of word knowledge. Using real English words, Chen and Truscott (2010) replicated the study by Webb (2007). The attainment rates of vocabulary knowledge were lower than those in Webb (2007), varying from almost zero to around 40%, either tested immediately or after two weeks. Similarly, Pellicer-Sánchez and Schmitt (2010) reported that L2 learners' retention of word knowledge after one occurrence of target L2 words in authentic texts ranged from 5% (meaning recall) to 30% (spelling recognition).

1.3 Linking Processing to Acquisition of Novel Words Using Eye Tracking

Eye tracking has become increasingly popular among SLA researchers (for a review, see Roberts & Siyanova-Chanturia, 2013). With the aid of eye trackers, eye movements such as fixations and saccades can be captured in real time through a considerable number of measures. Eye-tracking measures are generally classified into early- and late-processing categories, depending on the cognitive processes indexed by them (Reichle et al., 1998). Although there is still ambiguity when interpreting eye-tracking measures in relation to particular language processing stages (Rayner & Liversedge, 2011), it is widely acknowledged that early-processing measures, such as the first fixation duration, the gaze duration and skipping rates, reflect processing steps such as familiarity check and lexical access. Meanwhile, late-processing measures, including but not limited to second-pass reading time, total reading time and fixation count, represent higher-order processes, such as reanalysis/integration of information and recovery from processing difficulties (Reichle et al., 1998). By combining early and late eye-tracking measures, it is possible to tease apart early processes from later ones during language processing (Roberts & Siyanova-Chanturia, 2013).

Eye movements are thought to reflect the allocation of overt attention. Broadly speaking, attention can be overt or covert (Carrasco, 2011); the former is usually accompanied with eye movements, whereas the latter can be employed without directing the eye gaze towards visual targets. According to the E-Z Reader model (Reichle et al., 1998), overt attention executed through eye movements is assumed to be allocated sequentially, processing one word at a time. Consequently, temporal eye-tracking measures, such as total reading time, can reflect the amount of overt attention allocated to each word. By quantifying the amount of overt attention allocated to linguistic stimuli using gaze durations, SLA scholars have examined the relationship between attentional processing and vocabulary acquisition, thus linking lexical processing and vocabulary acquisition. Focusing on incidental vocabulary learning from reading, researchers have reported a positive relationship between total reading time and L2 learners' vocabulary learning outcomes, including form recognition (Mohamed, 2018), meaning recall (Godfroid et al., 2018; Pellicer-Sánchez, 2016), and meaning recognition (Mohamed, 2018). Perez and colleagues (Perez et al., 2015) compared L2 students' learning and processing of novel French words through video-watching with L2 subtitles or captions, under intentional and incidental learning conditions. They found that total reading time positively correlated with form recognition of L2 words in the full captioning, intentional learning condition.

2. The Current Study

This study is motivated by the following gaps in the current literature on the processing and learning of novel L2 words. First, few studies have simultaneously examined how word length, concreteness and contextual support impact the processing and learning of novel L2 lexical items. Second, most SLA research has focused on vocabulary learning with repeated exposures, leaving the processing and learning mechanisms when L2 learners first encounter novel words largely unknown. Third, although scholars have begun to address the link between real-time processing and consolidation/learning of novel lexical items, more studies are still needed to investigate the relationship between online lexical processing and vocabulary acquisition. Last but not least, methodologically, a majority of SLA studies as reviewed in the previous section have examined the processing of learning of novel lexical items by using texts/stories with pseudowords mapped to familiar concepts (e.g., *holter*-house in Pellicer-Sánchez, 2016) inserted. Vocabulary learning often involves the mapping between novel orthographic forms and new concepts. Consequently, the choice of pseudowords, although with good experimental control for prior lexical knowledge, may not fully reflect vocabulary acquisition in real life. In contrast, although texts/stories are ecologically valid for investigating vocabulary learning and processing, they provide little control for linguistic and contextual properties.

To bridge these gaps, the current research investigated L2 learners' initial processing and learning of novel L2 words when first encountered during reading, targeting the effects of word length, concreteness, and contextual support—which functions at the orthographic, semantic and contextual level respectively. A sentence-reading task was adopted, with L2 learners exposed to novel concrete and abstract words embedded in sentences that varied in contextual support (i. e., whether sentential contexts provide semantic clues for inferring the meaning of target words). Sentence-reading tasks have been used by many researchers when examining the effects of word length, concreteness, or contextual support on lexical processing (e.g., Brusnighan & Folk, 2012; Calvo & Meseguer, 2002; Chaffin et al., 2001; Juhasz & Rayner, 2003; Kliegl et al., 2004; McDonald, 2006) and vocabulary learning (e. g., Ma et al., 2015; Mestres-Missé et al., 2014; Mulder et al., 2019; Webb, 2008). Compared with reading authentic texts, it enables the researchers to manipulate variables of interest while maximizing experimental control. Eye tracking can help reveal the time course of effects on language processing (Juhasz & Rayner, 2003). Furthermore, it provides a way to quantify overt attention. To identify and compare the temporal loci of effects of word length, concreteness and contextual support, early-processing measures—including the first fixation duration, the gaze duration and skipping rates—were employed, along with late-processing measures, including second-pass reading time, total reading time, and fixation count. Their definitions are provided in Table 1 (adapted from Juhasz & Pollatsek, 2011). Following the practice of SLA researchers (e.g.,Godfroid et al., 2013, 2018; Mohamed, 2018; Pellicer-Sánchez, 2016), a global eye-tracking measure, namely, total reading time, was used as the metric of attentional processing to predict the amount of initial vocabulary knowledge acquired from reading. In addition, three immediate posttests of vocabulary knowledge, measuring form recognition, meaning recall and meaning recognition respectively, were administered.

Given only one exposure to novel L2 words, the following research questions (RQs) were addressed:

1) How do word length, concreteness and contextual support influence L2 learners' online processing of novel words during reading? Do such influences differ over time?

2) How do word length, concreteness and contextual support influence L2 learners' initial knowledge of novel words?

3) Does attentional processing of novel L2 words, operationalized by total reading time, predict the amount of initial vocabulary knowledge acquired from reading?

Table 1. *Definitions of Eye-Tracking Measures*

	Eye-tracking measures	Definition
Early-processing measures	First fixation duration	The duration of the first fixation on the area of interest during the first pass, irrespective of the number of fixations.
	Gaze duration	The sum of all first-pass fixations on the area of interest.
	Skipping rate	The percentage of cases in which the area of interest is not fixated on the first pass.
Late-processing measures	Second-pass reading time	The amount of time spent re-reading the area of interest after the first-pass reading.
	Total reading time	The total time spent reading the area of interest during the first and second pass.
	Fixation count	The number of fixations made within the area of interest.

3. Methodology

3.1 Participants

Sixty-five Chinese EFL college students (55 females) speaking English at the upper-intermediate level were recruited in Beijing. Their mean age was 24 years (*SD* = 2), and they started learning English at an average age of 9 (*SD* = 2). By the time the experiment began, all participants had passed the College English Test at Level Six (CET-6),[1] with an average score of 502 (*SD* = 47). Thirteen participants reported to have studied other foreign languages, including German, French, Japanese, Korean, and Spanish. The participants all had normal or corrected-to-normal vision.

3.2 Materials

Thirty-two low-frequency English nouns were selected from the MRC Psycholinguistic Database (Coltheart, 1981). To start, 1, 537 words consisting of 5-12 letters were obtained. These words were then split into short (5-7 letters) vs. long (8-12 letters) bins in terms of word length. Based on the first author's intuition, 190 nouns (concrete: 87; abstract: 103) that are likely to be unfamiliar to the participants were selected. Subsequently, ten Chinese EFL students who did not participate in the experiment were asked to rate their degree of familiarity with these candidate words, based on a four-point Likert scale (1: *I have never seen this word before*; 2: *I have seen this word before*; 3: *I know the meaning of this word*; 4: *I am very familiar with this word*). After excluding words with familiarity ratings higher than 3 by more than five raters, 32 monomorphemic lexical items were selected as the target words, with eight words for each combination of word length and concreteness. On average, the familiarity rating was 1.37 (*SD* = 0.38). Statistics for these words, including logged word frequency (*M* = 0.43, *SD* = 0.22), mean bigram frequency (*M* = 2.26, *SD* = 0.38) and orthographic/ phonological neighborhood size were retrieved from N-Watch (Davis, 2005). A MANOVA analysis revealed the mean bigram frequency of short words (*M* = 2.04, *SD* = 0.34) was significantly higher than that of long words (*M* = 2.48, *SD* = 0.31). None of the target words had any homophones. Characteristics of the target words are summarized in Table 2.

[1] CET-6 is a norm-referenced English proficiency test developed in China, with a mean of 500 for the norming group (*SD* = 70, *Min* = 220, *Max* = 710). To pass CET-6, one has to achieve a score no less than 425. Students passing the CET-6 achieve the B2 (upper-intermediate) level of the Common European Framework of Reference for Languages.

Table 2. *Characteristics of the Target Novel Words*

Item	Condition	Frequency	MBF	Length	ONS	PNS	FAM
turpentine	CL	0.34	2.22	10	0	1	1.43
chinchilla	CL	0	1.59	10	0	*	1.43
mahogany	CL	0.73	1.71	8	0	0	1.29
albatross	CL	0.25	1.76	9	0	0	1.00
limousine	CL	0.55	2.46	9	0	1	1.00
moccasin	CL	0.14	1.94	8	0	1	1.00
mackerel	CL	0.86	2.39	8	0	0	1.00
amethyst	CL	0.25	1.63	8	0	0	1.00
soprano	CS	0.5	2.25	7	0	0	1.57
apricot	CS	0.39	1.9	7	0	1	1.40
canary	CS	0.55	2.71	6	1	1	1.86
diadem	CS	0	2.46	6	0	*	1.00
pliers	CS	0.39	2.75	6	1	3	1.00
casket	CS	0.47	2.84	6	1	1	1.20
walrus	CS	0.18	1.89	6	0	0	1.20
satchel	CS	0.31	2.69	7	0	0	1.00
oblivion	AL	0.66	2.35	8	0	1	2.00
blasphemy	AL	0.57	1.77	9	1	0	1.29
allegory	AL	0.35	2.18	8	0	0	1.29
reprisal	AL	0.37	2.59	8	0	1	1.00
delirium	AL	0.48	2.35	8	0	1	2.00
panorama	AL	0.63	2.21	8	0	0	2.00
equanimity	AL	0.46	1.92	10	0	0	1.00
etiquette	AL	0.71	1.68	9	0	0	1.00
enigma	AS	0.53	2.1	6	0	1	2.29
malady	AS	0.24	2.61	6	0	1	1.71
mutiny	AS	0.51	2.67	6	0	0	1.29
menace	AS	0.93	2.64	6	1	2	1.60
bequest	AS	0.3	2.55	7	1	2	1.40
carnage	AS	0.37	2.74	7	0	1	1.40
forfeit	AS	0.33	2.37	7	0	3	2.00
caucus	AS	0.38	2.27	6	0	5	1.20

Note. CL: concrete-long; CS: concrete-short; AL: abstract-long; AS: abstract-short. Word frequencies were on the logarithm scale (base: 10). MBF: mean bigram frequency. Word length was measured as the number of letters. ONS/PNS: orthographic/phonological neighborhood size. *: not provided by N-WATCH. FAM: average familiarity ratings (based on a four-point scale).

Two critical sentences that differed in the degree of contextual support were created for each target word: one informative and one neutral. The neutral context offered relatively little information about the meaning of the target word, whereas the informative context provided a semantic clue for the inference of the meaning of the category that the word belongs to (e.g., liquid for the target word "turpentine"). These sentences were created such that they shared the same syntactic structure, starting with a main clause and accompanied by an object relative clause led by the pronoun "that." The target noun was inserted in the final position of the main clause and acted as the object of the verb phrase. Twenty Chinese EFL college students who did not participate in the experiment were asked to read the critical sentences and guess the meaning for each target noun. Based on their responses, critical sentences were revised, such that the semantic category of each target word can be guessed from the informative contexts by most respondents, while such cannot be achieved for the neutral contexts. The average score of contextual support (computed as the percentage of people who guessed the meaning correctly) for the informative and neutral contexts was 80% (*SD* = 10%) and 20% (*SD* = 10%) respectively. Examples of the critical sentences are illustrated below, with the target word in bold.

1) *My neighbor used the **turpentine** that he mixed with the oil*. (informative)

2) *My neighbor used the **turpentine** that he stored in the studio.* (neutral)

Sixty-four filler sentences were borrowed from Jiang et al. (2011) after minor modifications. Additionally, four sentences were created and used for practice. To avoid reading two sentences with the same target noun , two counterbalanced lists were generated, each containing 100 sentences—32 critical sentences, 64 filler sentences, and 4 practice sentences. Two native speakers of English were asked to examine the readability of all sentences, and revisions were made when necessary. Thirty-two yes/no comprehension questions were created, following half of the filler sentences. The sentences were then randomly assigned to eight blocks, with four critical sentences and eight filler sentences in one block. Randomization was carried out both within and across the blocks when running the experiment.

3.3 Procedure

After signing the consent form, participants were randomly assigned to one of the two counterbalanced lists of sentences. They were instructed to focus on understanding the meaning and reading the sentences normally[2] at their own pace. They were notified that there would be comprehension tests following the reading session. An EyeLink 1000 Plus eye tracker (SR Research, Canada) was used. The sampling rate was 1,000 Hz. Sentences were presented in a normal, spaced manner on a 21-inch CRT monitor (resolution: 1,024 × 768 pixels; refresh rate: 150 Hz) that was connected to a Dell PC. Each sentence was displayed in a single line in Times New Roman 20-point font, with the words shown in black (RGB: 0, 0, 0) on a white background (RGB: 255, 255, 255). Participants were seated at a distance of 70 centimeters from the computer monitor, and their head was stabilized by a chin rest and a forehead rest. Participants read the sentences binocularly, but only their right eye was monitored. A three-point horizontal calibration procedure (e.g., White et al., 2008) was run at the beginning of the reading session, with recalibration performed when necessary.

Immediately after the reading session, three surprise vocabulary tests—starting from a form recognition test, then a meaning recall test, and finally a meaning recognition test—were delivered. The form recognition test (Appendix 1) was used to assess L2 learners' initial orthographic knowledge of the target nouns, in which participants had to choose the words

[2] Strategic behavior such as intentionally skipping unfamiliar words was discouraged.

that appeared in the reading session out of 96 words (32 target nouns and 64 distractors). In the meaning recall test, participants had to explicitly explain the meaning of the target nouns, using either English or their L1 (i.e., Mandarin). The meaning recognition test evaluated the receptive knowledge of form-meaning connections, and participants were required to choose the correct meaning out of four choices for each novel word. Distractors for the meaning recognition test were created based on incorrect responses collected when piloting the meaning generation test. They were matched for part of speech with the target definitions. Additionally, the meaning recognition test was created in a way such that the chances of selecting any of the four choices were equal (Appendix 2). Test items in the above tests were created out of context and randomized. Participants received a score of 0 or 1, depending on whether they responded correctly. After completing the experiment, a brief survey was administered to assess participants' prior knowledge of the target nouns. Participants had to rate their degree of familiarity with each target noun based on their experience with each word before participating in this study, using a four-point Likert scale as mentioned in the previous section.

3.4 Statistical Analysis

In order to examine how word length, concreteness and contextual support influence online processing of novel words during reading, as well as how attentional processing of novel words predicts vocabulary learning, two area of interests (AIs) were defined. Eye-movement data were extracted from the target novel nouns (AI_1) and the relative clauses following the relative pronoun (AI_2). Two participants were excluded from data analysis, due to accuracy rates for the comprehension questions being lower than 75%. Before analyzing the eye-movement data, all trials where track loss occurred were removed (2.3%). In addition, individual fixations shorter than 80 ms (7.4%) were also excluded (Betancort et al., 2009). Meaning recall turned out to be too difficult, with almost no participant responding correctly. Therefore, this test was not incorporated for analysis. Since L2 words, instead of pseudowords, were used in this study, self-reported prior knowledge of the target nouns collected from each participant (see *Procedure*) was checked against their performance on the immediate vocabulary posttests. Specifically, if a participant claimed to have already known a target word prior to taking the experiment, and indeed he or she responded to this word correctly in the form and meaning recognition tests, then his or her eye-tracking and vocabulary-test data for this word were excluded from analysis (2.9%). To identify items that participants answered inconsistently in the vocabulary knowledge tests, Rasch analyses were carried out, using the TAM package (version 2.13-15, Robitzsch et al., 2019) in R. Mean squares of the infit statistic were between 0.6 and 1.5. Consequently, no item was removed from analysis for the form and meaning recognition tests.

Fixation durations were transformed using natural log. Word length (the number of letters) and contextual support (as percentages) were treated as numeric variables, whereas concreteness was dummy-coded, with abstract words as the reference group. Depending on whether participants reported to have encountered a target word before participating in the experiment (see *Procedure*), a binary variable "prior exposure" (1. *I have never seen this word* *before;* 2. *I have seen this word before/I know the meaning of this word/I am very familiar with this word*) was dummy-coded and treated as a covariate. Bigram frequency has been found to influence the processing of low-frequency words (Rice & Robinson, 1975). Given that the target novel words were not matched for mean bigram frequency (logged), this variable was also included as a covariate. All numeric variables were centered at their means.

Mixed-effects models were fit to use the *lme4* package (version 1.1-21, Bates et al., 2015) in R (version 3.6.2, R Core Team, 2019). For the target nouns (AI_1), linear mixed-effects models were fit to model fixation durations, whereas mixed-effects logistic and Poisson models were fit to model skipping rates and fixation count respectively. Linear-mixed effects models were also fit to analyze total reading time for the relative clauses (AI_2) to examine how L2 learners process sentential contexts. In order to investigate whether participants' initial vocabulary knowledge (i.e., form and meaning recognition) can be predicted by word length, concreteness, contextual support, and real-time processing of the novel words, mixed-effects logistic models were fit. Statistical models were implemented using a maximum likelihood technique, following forward model selection procedures. Statistical analysis for online processing of novel words (RQ_1) started from models consisting of random intercepts for subjects and items, with word length, concreteness, contextual support and their interactions entered sequentially, followed by the covariates (i. e., mean bigram frequency and prior exposure). Statistical analysis for vocabulary learning (RQ_2 and RQ_3) adopted the same procedure, except that attentional processing, as operationalized by total reading time for the target words, were also added as a predictor. Random slopes of effects for subjects and items were tested after we selected the best-fitted models based on the above procedure. Model comparisons were performed using the *anova* function in the *lme4* package. The significance level *alpha* was set at .05. For each best-fitted model, effect sizes were measured by marginal R^2 (the variance explained by the fixed effects) and conditional R^2 (the variance explained by the whole model), obtained from the *tab_model* function in the sjPlot package (version 2.7.2, Lüdecke, 2019), along with *p*-values.

4. Results

4.1 Online Processing of Novel Words During Reading

Participants answered 89% (*SD* = 6%) of the comprehension questions correctly, indicating that they did follow the instructions and read for comprehension. The average fixation durations on the target novel words were as follows: 305 ms (*SD* = 172 ms) for the first fixation duration, 685 ms (*SD* = 544 ms) for the gaze duration, 610 ms (*SD* = 474 ms) for second-pass reading time, and 1854 ms (*SD* = 1062 ms) for total reading time. The mean skipping rate for the novel words during first-pass reading was 36% (*SD* = 48%). Additionally, each target noun received an average of 6 fixations (*SD* = 3).

4.1.1 Early Processing of Novel Words

Mixed-effects modeling for early-processing eye-tracking measures (i.e., the first fixation duration, the gaze duration, and skipping rates) revealed that early processing of longer novel words took more time than shorter novel words, as indexed by a significant effect of word length for the first fixation duration (*Estimate* = -0.025, *SE* = 0.009, t = -2.863, p = .004). Effects of concreteness and contextual support were not detected in any of the early-processing measures. A significant effect of prior exposure (*Estimate* = 0.100, *SE* = 0.046, t = 2.206, p = .027) was found for the gaze duration, suggesting that L2 learners spent longer time processing novel words reported to have been seen previously than those first encountered. Mean bigram frequency also turned out to influence the skipping rates (*Estimate* = 0.507, *SE* = 0.212, t = 2.383, p = .017), indicating that novel words with higher mean bigram frequencies were more likely to be skipped. Results for early processing of the novel words are summarized in Table 3.

Table 3. *Mixed-Effects Models for Early-Processing Eye-Tracking Measures*

	First fixation duration				Gaze duration				Skipping rate			
Parameters	*Estimate*	*SE*	*t*	*p*	*Estimate*	*SE*	*t*	*p*	*Estimate*	*SE*	*t*	*p*
Intercept	5.587	0.025	223.26	<0.001	6.167	0.050	122.490	<0.001	-0.730	0.146	-4.986	<0.001
Length	-0.025	0.009	-2.863	0.004								
Prior exposure					0.100	0.046	2.206	0.027				
Mean bigram frequency									0.507	0.212	2.383	0.017
Random effects												
Variance	0.029 subject				0.126 subject				0.935 subject			
	0.001 item				0.001 item				0.106 item			
Marginal R^2 / conditional R^2	0.005 / 0.129				0.003 / 0.183				0.008 / 0.247			

Note. Prior Exposure was dummy-coded, with the reference level meaning that participants had no exposure to the target words previously. Mean bigram frequency was logged.

4.1.2 Late Processing of Novel Words

Mixed-effects modeling for late-processing eye-tracking measures (i. e., second-pass reading time, total reading time, fixation count) revealed significant effects of word length for second-pass reading time (*Estimate* = 0.038, *SE* = 0.014, t = 2.743, p = .006), total reading time (*Estimate* = 0.064, *SE* = 0.007, t = 8.786, p < .001), and fixation count (*Estimate* = 0.074, *SE* = 0.007, z = 11.193, p < .001). This indicates that words consisting of more letters received longer second-pass reading time and total reading time, as well as more fixations, than those composed of fewer letters. Additionally, concrete novel words were also found to receive fewer processing resources than abstract ones, indicated by significant effects of concreteness for total reading time (*Estimate* = -0.064, *SE* = 0.020, t = -3.107, p = .002) and fixation count (*Estimate* = -0.043, *SE* = 0.019, z = -2.282, p = .023). Effects of contextual support on the processing of novel words were not found in any of the late-processing measures. No significant interactions among word length, concreteness and contextual support were detected. Results for late processing of the novel words are summarized in Table 4.

4.1.3 Processing of Sentential Contexts

In order to examine how L2 learners process sentential contexts following the target novel words, a statistical analysis was implemented for the relative clauses that followed the target nouns, focusing on total reading time. Mixed-effects modeling revealed that contextual

Table 4. *Mixed-Effects Models for Late-Processing Attentional Measures*

	Second-pass reading time				Total reading time				Fixation count			
Parameters	*Estimate*	*SE*	*t*	*p*	*Estimate*	*SE*	*t*	*p*	*Estimate*	*SE*	*z*	*p*
Intercept	6.120	0.037	163.822	<0.001	7.391	0.054	137.704	<0.001	1.798	0.049	36.752	<0.001
Length	0.038	0.014	2.743	0.006	0.064	0.007	8.786	<0.001	0.074	0.007	11.193	<0.001
Concreteness					-0.064	0.020	-3.107	0.002	-0.043	0.019	-2.282	0.023
Random effects												
Variance	$0.063_{subject}$				$0.168_{subject}$				$0.139_{subject}$			
	0.001_{item}				0.000_{item}				0.000_{item}			
Marginal R^2 / conditional R^2	0.005 / 0.110				0.026 / 0.515				0.068 / 0.491			

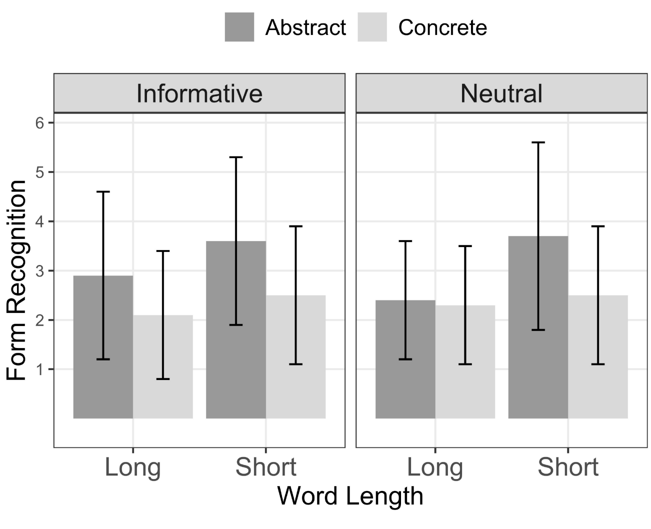
Note. Concreteness was dummy-coded, with abstract words as the reference group.

support had a significant effect for total reading time (*Estimate* = 0.001, *SE* = .000, t = 4.673, p < .001), suggesting that informative contexts were read longer than neutral contexts. Furthermore, a significant interaction between contextual support and word length was found (*Estimate* = 0.0005, *SE* = .0002, t = 2.640, p = .008), indicating that the effect of contextual support for the total reading time in sentential contexts was larger for longer novel words. Effect sizes measured by Marginal R^2 and Conditional R^2 were 0.008 and 0.512 respectively.

4.2 Initial Knowledge of L2 Words Acquired From Reading

Out of the 32 novel words, the average number of orthographic forms and meanings recognized by the participants were 22 (69%) and 18 (56%) respectively. Means and standard deviations of form- and meaning-recognition scores across conditions are presented in Figure 1.

Figure 1. *Form and Meaning Recognition Performance Across Conditions Generated by Word Length (Long vs. Short) and Contextual Support (Informative vs. Neutral)*



Note. Each condition contains a maximum of eight target words. The Y-axes represent the average number of words responded correctly by participants, whereas the error bars illustrate their standard deviations.

Mixed-effects logistic models revealed significant effects of concreteness (*Estimate* = -0.671, *SE* = 0.267, *z* = -2.514, *p* = .012), total reading time (*Estimate* = 0.374, *SE* = 0.137, *z* = 2.718, *p* = .007) and prior exposure (*Estimate* = 1.031, *SE* = 0.154, *z* = 6.677, *p* < .001) on form recognition of the novel words. Such results suggest that the orthographic forms of novel abstract words were easier to recognize than those of novel concrete words. Moreover, total reading time spent on processing the novel words, regardless of concreteness, contributed positively to L2 learners' form recognition performance. Lastly, novel words reported to have been seen by participants prior to the experiment were more likely to be recognized correctly, in terms of orthographic form. Word length and contextual support had no influence on L2 learners' form recognition performance. A significant effect of concreteness was also found for meaning recognition (*Estimate* = -0.496, *SE* = 0.223, *z* = -2.219, *p* = .027), indicating that abstract novel words also outweighed concreteness novel words with respect to the recognition of meaning. Word length and contextual support, as well as attentional processing of novel words operationalized by the total reading time, did not predict L2 learners' meaning recognition performance. Results for the initial vocabulary learning are summarized in Table 5.

Table 5. *Mixed-Effects Models for Vocabulary Learning*

	Form recognition				Meaning recognition			
Parameters	*Estimate*	*SE*	*z*	*p*	*Estimate*	*SE*	*z*	*p*
(Intercept)	-3.530	1.048	-3.370	0.001	-0.784	0.163	-4.817	<0.001
Concreteness	-0.671	0.267	-2.514	0.012	-0.496	0.223	-2.219	0.027
Total reading time	0.374	0.137	2.718	0.007				
Prior exposure	1.031	0.154	6.677	<0.001				
Random effects								
Variance	1.12_{subject}				0.13_{subject}			
	0.45_{item}				0.30_{item}			
Marginal R^2 / conditional R^2	0.082 / 0.378				0.016 / 0.130			

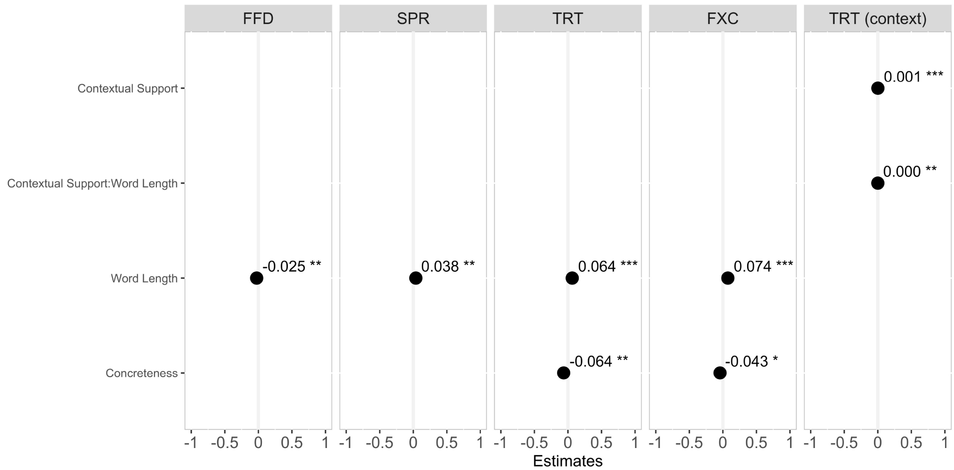
Note. Concreteness was dummy-coded, with abstract words as the reference group. Prior Exposure was also dummy-coded, with the reference level meaning that participants had no exposure to the target words previously.

5. Discussion

5.1 Summary of Results

Using eye tracking, this study investigated second language learners' initial processing and learning of novel L2 words, when first encountered during reading. The following processing patterns were revealed. First, word length showed significant impacts on L2 learners' early and late processing of novel words, as indexed by shorter first fixation durations, longer second-pass reading time and total reading time, as well as more fixations for words consisting of more letters. Second, concreteness exerted influence only on late processing of novel L2 words, with concrete lexical items receiving shorter total reading time and fewer fixations. Third, contextual support did not have any direct influence on the processing of novel L2 words. Meanwhile, informative contexts attracted longer total reading time than neutral contexts, with such an effect being larger for longer words. The effects of word length, concreteness and contextual support for early and late processing of novel L2 words are summarized in Figure 2.

Figure 2. *Summary of the Effects of Word Length, Concreteness, Contextual Support and Their Interaction (Contextual Support: Word Length) on Early (I. E. , FFD) and Late (I. E. , SPR, TRT, FXC) Processing of Novel L2 Words and Sentential Contexts [I. E. , TRT (Context)].*



Note. The Y-axis lists the effects of interest, whereas the black dots along with the values on the panel demonstrate the coefficient estimates obtained from the statistical models. FFD: first fixation duration. SPR: second-pass reading time. TRT: total reading time. FXC: fixation count. TRT (context): total reading time spent on sentential contexts.

Regarding the initial learning of novel L2 words, given one exposure, L2 learners obtained an accuracy rate of 69% and 56% for the form and meaning recognition test respectively. L2 learners were not able to gain any productive knowledge of the target novel words (i. e., meaning recall). Surprisingly, abstract novel words were acquired better than concrete ones, in terms of form and meaning recognition. Finally, attentional processing of novel L2 words, as operationalized by total reading time, contributed positively to participants' performance on the form recognition test.

5.2 Initial Processing of Novel L2 Words During Reading

This study revealed that word length influences both early and late processing of novel L2 words, when first encountered during reading. Such results are consistent with current research findings. Lowell and Morris (2014) reported early influences of word length on the first fixation duration, the gaze duration and skipping rates for novel words. Effects of word length for early-processing measures, including the first fixation duration (e. g., McDonald, 2006), the gaze duration (e. g., Calvo & Meseguer, 2002; Hyönä & Olson, 1995; Juhasz & Rayner, 2003; Kliegl et al., 2004; McDonald, 2006) and skipping rates (e.g., Calvo & Meseguer, 2002; Kliegl et al., 2004), as well as for late-processing measures, such as total reading time (e. g., Juhasz & Rayner, 2003; Kliegl et al., 2004) and fixation count (Hyönä & Olson, 1995; McDonald, 2006), have also been found for processing of familiar words. Specifically, researchers have consistently reported that longer words take more time to process than shorter words throughout processing stages. Nevertheless, in this study, we found that the processing disadvantage of long words was reversed for the first fixation duration—novel L2 words composed of more letters received shorter first fixation durations than those with fewer letters—which stands in contrast to what McDonald (2006) reported for familiar words. According to the E-Z Reader model (Reichle et al., 1998), when processing words during reading, readers start with familiarity check of the orthographic form, which is generally indexed by the first fixation duration (Juhasz & Rayner, 2003). Following this, less time may be needed for L2 learners to process the orthographic form for longer unfamiliar words and recognize their novelty, given that words consisting of more letters may be visually more salient (e.g., Behney et al., 2017; Simoens et al., 2017). By contrast, during late-processing stages, in which L2 learners need to link the orthographic forms of novel words to their meanings, more time is needed to encode form-meaning connections for lexical items consisting of more letters.

Interestingly, concreteness was responsible only for late processing of novel L2 words. Compared with abstract novel words, concrete novel words received shorter total reading time and fewer fixations. Previous studies have reported effects of concreteness for processing of familiar words, extending from early measures—such as the first fixation duration (e.g., Loomis, 2010) and the gaze duration (e.g., Juhasz & Rayner, 2003; Loomis, 2010)—to late measures, such as total reading time (Juhasz & Rayner, 2003). Unlike familiar words, novel words do not have lexical semantic representations in the mental lexicon. A possible explanation for the discrepancy in the temporal loci of concreteness effects between familiar and novel words may be as follows. For familiar words, semantic information, including concreteness, can be accessed from early on. However, for novel lexical items, due to the absence of lexical representation, the recognition of word concreteness may not be accomplished until later stages, after referring to contextual information. Our findings replicate the processing advantages of concrete known words widely reported in the current literature (Juhasz & Rayner, 2003; Loomis, 2010; Schwanenflugel, 1991). According to the classic dual-coding theory (Paivio, 1991), concrete words are recognized faster than abstract words because of the activation of an extra image-based processing system. For concrete novel L2 words embedded in sentential contexts, meaning inference may benefit from the access to nonverbal imagery, resulting in less time demanded for the encoding of semantic information than abstract novel items.

The role of contextual support in the processing of novel L2 words is rather intriguing. Contextual support has been shown to influence the processing of familiar words during reading: lexical items following predictable contexts are processed faster than those embedded in less predictable contexts (Calvo & Meseguer, 2002; Kliegl et al., 2004; Staub, 2015). In this study, we found no direct influence of contextual support on the processing of novel words. However, contextual support did exert an influence on the processing of sentential contexts, with informative contexts receiving longer total reading time than neutral contexts. Such results replicated the findings of Chaffin et al. (2001). Compared with studies that reported significant effects of contextual support for word processing (e.g., Calvo & Meseguer, 2002; Kliegl et al., 2004), the current study differed in two important ways. First, novel words—

instead of familiar words—were used. Second, target words were embedded preceding sentential contexts, instead of following them. When processing familiar words following contexts varying in contextual support, phonological, morphological, or semantic properties of the target words may be activated (Brothers et al., 2015), thus facilitating word recognition. When processing novel words that precede sentential contexts, as in the current study and Chaffin et al. (2001), participants may not relate the sentential contexts to the target novel lexical items. This may be particularly true given that L2 learners in this study were instructed to read for comprehension, but not to intentionally figure out the meaning of the novel words. The significant positive effect of contextual support on the processing of sentential contexts, however, suggests that L2 learners did make use of contextual information when trying to understand the reading materials. Finally, the significant interaction between word length and contextual support on the processing of sentential contexts— that is, L2 learners tended to exploit contextual support when encountering longer words—is also worth mentioning. Such a pattern may result from the possibility that L2 learners were more aware of the novelty of longer novel words and relied more heavily on contextual information to understand the sentences. Given that such a result has not been reported by previous studies, additional research is needed to further validate this explanation.

5.3 Initial Learning of Novel L2 Words From Reading

Exposed to novel L2 words only once during reading, second language learners in this study achieved an accuracy rate of 69% and 56% for form and meaning recognition respectively. However, they were not able to recall the meaning of novel L2 words. Since participants in this study were discouraged from intentionally skipping words unfamiliar to them while reading for comprehension, their consciousness about the existence of novel lexical items was raised, resulting in what SLA researchers (e.g., Laufer, 2005) call "focus-on-form" learning. Few focus-on-form studies have examined how much vocabulary knowledge can be acquired from one encounter during reading. However, to a certain extent, retention levels found in the current research are comparable to those reported by studies examining incidental vocabulary learning. Using pseudowords that replaced common L2 concepts (e.g., face, hospital), Webb (2007) found that L2 learners recognized 67% of the orthographic forms of novel words after one encounter. Chen and Truscott (2010) replicated his study while using low-frequency L2 words. They reported that L2 learners' retention rate for form recognition after one exposure was 43%. The retention rates for meaning recognition reported by Webb (2007) and Chen and Truscott (2010) were 48% and 40% respectively, much lower than what was found by us. In terms of meaning recall, L2 learners in this study could barely produce the meaning of target novel words, whereas measurable levels of retention were reported by Webb (58%) as well as Chen and Truscott (10%).

This study also sheds light on the roles of intralexical and contextual variables in the initial learning of L2 words from reading. Generally speaking, longer words are more difficult to acquire than shorter ones, and this has been confirmed by studies both on incidental (e.g., Godfroid et al., 2018) and intentional vocabulary learning (e.g., Papagno & Vallar, 1992). Godfroid et al. (2018) reported that word length moderated L2 learners' performance on form recognition, meaning recall, and meaning recognition of novel words, yet their study included multiple exposures and did not control for semantic and contextual properties of lexical items. In the current research, although word length was found to influence both early and late processing of novel words, it showed no impact on the initial attainment of orthographic and semantic knowledge. Acquiring the orthography and meaning of new words requires the encoding of such information in memory. Papagno and Vallar (1992) hypothesized that longer novel words are more difficult to acquire, because they impose a processing burden during the rehearsal process. Since participants in this study encountered the novel words only once and were immediately tested on vocabulary knowledge after the reading session, rehearsal processes might have been largely squeezed, leading to the absence of word length effects on vocabulary learning outcomes.

In contrast to previous research findings (e.g., de Groot & Keijzer, 2000; Elgort & Warren, 2014; Mestres-Missé et al., 2014), we found that abstract nouns were acquired significantly better than concrete nouns, in terms of both form and meaning recognition (see Figure 1). Such results seem rather counterintuitive. Similar processing advantages for abstract words have been observed in neuropsychological patients (Mestres-Missé et al., 2014), and such intriguing patterns may possibly be understood based on a hypothesis put forward by Plaut and Shallice (1993, as cited in Mestres-Missé et al., 2014). According to these researchers, concrete words differ from abstract words in that the former are endowed with more semantic features. When inferring the meaning for novel nouns from contexts, more candidate meanings can be activated for concrete words than for abstract words, making it more difficult to anchor the form-meaning connections for novel concrete words. Buchanan et al. (2019) collected semantic feature norms for 4,436 concepts. A comprehensive search found that four concrete words (i. e., *canary*, *limousine*, *pliers*, *walrus*) and one abstract word (i. e., *etiquette*) appeared in their norming list, with the number of semantic features being 4, 16, 16, 16 and 11, respectively. It appears that concrete words indeed have more semantic features than abstract words. However, larger norming datasets are needed to fully confirm such a hypothesis. Orthographic and semantic information have to be integrated during the encoding process. As a result, the observed disadvantage for the acquisition of the meaning of concrete words is likely to be extended to the retention of their orthography, leading to abstract words outweighing concrete words in form recognition as well. The acquisition advantage of concrete words as reported in this study is limited to a situation where L2 learners first encounter a novel word during reading. Future studies are needed to replicate our research findings and verify such a preliminary interpretation.

The absence of the effects of contextual support on the initial learning of novel L2 words is also worth mentioning. Novel words embedded in informative contexts are expected to be acquired more easily than those in neutral. Among the studies that reported significant effects of contextual support on word learning contexts (e. g., Ma et al., 2015; Mulder et al., 2019; Webb, 2008), sentential contexts varying in the degree of contextual support preceded target lexical items. However, in our study, sentential contexts followed the target novel words. Therefore, it is possible that L2 learners may not integrate contextual information with the target novel lexical items to the same extent as when sentential contexts precede target words, especially when they are under a reading-for-comprehension condition.

5.4 Linking Real-Time Word Processing to Vocabulary Learning Outcomes

Attentional processing, as measured by the total reading time, predicted L2 learners' performance on form recognition, but not on meaning recognition, with a longer total reading time on novel L2 words leading to better recognition of their orthographic forms. The predictive power of total reading time for form recognition is unsurprising, as this has been reported by Mohamed (2018). The absence of such an effect for meaning recognition, however, is not expected. Assuming that more attentional processing leads to more acquisition (Godfroid et al., 2013), it was hypothesized that attentional processing, quantified by a global eye-tracking measures such as total reading time, should predict L2 learners' achievement in meaning recognition, as it did for form recognition. SLA researchers have found that total reading time summed across encounters had a significant effect on meaning recognition (Godfroid et al., 2018; Mohamed, 2018). Therefore, we believe that the lack of predictive power of total reading time for meaning recognition may be due to the inadequacy of exposure to the target words. In that way, one exposure to novel words suffices for the encoding of orthographic information, yet it may not be enough for the establishment of initial form-meaning connections.

5.5 Limitations

Most studies in the literature have selected the Courier font to present stimuli. Unlike Times New Roman, Courier is a mono-spaced font, with each letter taking the same amount of horizontal space. Consequently, when using Times New Roman, lexical items consisting of the same number of letters are not matched for spatial width. On the other hand, as indicated by Hautala et al. (2011), when the Courier font is used, the spatial width of words perfectly aligns with the number of letters, making it impossible to tell whether effects of word length reflect processing at the visual or orthographic level. Given that we defined word length as the number of letters, using Times New Roman instead of Courier helps disentangle such effects. Bernard and colleagues (Bernard et al., 2002) found that Times New Roman was read significantly faster and was perceptually more attractive than Courier. Moreover, recent studies (Hautala et al., 2011; McDonald, 2006) also suggest that the spatial width of letters only affects the landing location and skipping rates of words during reading. In our study, word length effects were reported only for fixation count and durations. Therefore, such results may not have been compromised by the use of the Times New Roman font. Another limitation lies in the use of real L2 words. The use of low-frequency words—instead of pseudowords—is to ensure the ecological validity of this study, since the latter generally involves the mapping between novel orthographic forms and familiar concepts. However, we cannot ensure that participants' prior knowledge of the target lexical items was fully removed from data analysis, even though multiple rounds of strict screenings were followed.

6. Conclusion

As far as we know, this study is the first that simultaneously examines the influences of orthographic, semantic and contextual influences on the processing and learning of novel L2 words during reading. With a combination of early- and late-processing eye-tracking measures, the time course of the effects of word length, concreteness and contextual support on real-time processing of novel L2 lexical items was delineated. Additionally, attentional processing of novel L2 words, operationalized by total reading time, was linked to the acquisition of initial vocabulary knowledge. Our study adds to the understanding of the mechanisms underlying novel word processing and L2 vocabulary learning. Future studies are needed to further explore the complex interactions among intralexical and contextual characteristics when investigating L2 word processing and learning.

References

Baddeley, A., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 575-589.

Balota, D. (1994). Visual word recognition: The journey from features to meaning. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (Vol. 29, pp. 303-358). Academic Press.

Barclay, S., & Schmitt, N. (2019). Current perspectives on vocabulary teaching and learning. In X. Gao. (ed.), *Second handbook of English language teaching* (pp. 1-22). Springer.

Barton, J. J. S., Hanif, H. M., Björnström, L. E., & Hills, C. (2014). The word-length effect in reading: A review. *Cognitive Neuropsychology, 31*(5-6), 378-412.

Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Behney, J., Spinner, P., Gass, S., & Valmori, L. (2017). The L2 acquisition of Italian tense: The role of salience. In S. Gass, P. Spinner, & J. Behney (Eds.), *Salience in second language acquisition* (pp. 89-106). Routledge.

Bernard, M., Lida, B., Riley, S., Hackler, T., & Janzen, K. (2002). A comparison of popular online fonts: Which size and type is best? *Usability News*, *4*(1), 1-8.

Betancort, M., Carreiras, M., & Sturt, P. (2009). The processing of subject and object relative clauses in Spanish: An eye tracking study. *Quarterly Journal of Experimental Psychology, 62*(10), 1915-1929.

Bisson, M., van Heuven, W., Conklin, K., & Tunney, R. (2014). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics*, *35*(2), 399-418.

Borovsky, A., Elman, J., & Kutas, M. (2012). Once is enough: N400 indexes semantic integration of novel word meanings from a single exposure in context, *Language Learning and Development*, *8*(3), 278-302.

Brothers, T., Swaab, T., & Traxler, M. (2015). Effects of prediction and contextual support on lexical processing: Prediction takes precedence. *Cognition*, *136*, 135-149. http://doi.org/10.1016/j.cognition.2014.10.017

Brusnighan, S., & Folk, J. (2012). Combining contextual and morphemic cues is beneficial during incidental vocabulary acquisition: Semantic transparency in novel compound word processing. *Reading Research Quarterly, 47*(4), 426-426.

Brysbaert, M., Warriner, A., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904-911.

Buchanan, E. M., Valentine, K. D., & Maxwell, N. P. (2019). English semantic feature production norms: An extended database of 4436 concepts. *Behavior Research Methods* 51**,** 1849-1863. https://doi.org/10.3758/s13428-019-01243-z

Calvo, M. G., & Meseguer, E. (2002). Eye movements and processing stages in reading: Relative contribution of visual, lexical, and contextual factors. *Spanish Journal of Psychology, 5*(1), 66-77.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research, 51*(13), 1484-1525.

Chaffin, R., Morris, R. K., & Seely, R. E. (2001). Learning new word meanings from context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 225-235.

Chen, C. T., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics, 31*(5), 693-713.

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology: A Human Experimental Psychology, 33A*(4), 497-505.

Crutch, S., & Warrington, E. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, *128*(3), 615-627.

Davis, C. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods, 37*(1), 65-70.

de Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning, 50*(1), 1-56.

Elgort, I., & Warren, P. (2014). L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning, 64*(2), 365-414.

Fukkink, R. G., Blok, H., & de Glopper, K. (2001). Deriving word meaning from written context: A multicomponential skill. *Language Learning*, *51*(3), 477-496.

Gibson, E., & Levin, H. (1975). On the perception of words: An application of some basic concepts. In E. Gibson & H. Levin (Eds.), *The psychology of reading*. The MIT Press.

Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y. Q., Johnston, S., Johnston, S., Lee, S., Sarkar, A., & Yoon, H. J. (2018). Incidental vocabulary learning in a natural reading context: An eye tracking study. *Bilingualism: Language and Cognition, 21*(3), 563-584.

Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: Gauging the role of attention in incidental L2 vocabulary acquisition by means of eye tracking. *Studies in Second Language Acquisition, 35*(3), 483-517.

Graves, M. (1986). Vocabulary learning and instruction. In E. Rothkopf & L. Ehri (Eds.), *Review of Research in Education* (Vol. 13, pp. 49-89). American Educational Research Association.

Hautala, J., Hyönä, J., & Aro, M. (2011). Dissociating spatial and letter-based word length effects observed in readers' eye movement patterns. *Vision Research*, *51*(15), 1719-1727.

Hiebert, E., Scott, J., Castaneda, R., & Spichtig, A. (2019). An analysis of the features of words that influence vocabulary difficulty. *Education Sciences*, *9*(1), 8. http://doi.org/10.3390/educsci9010008

Hu, C. (2012). Fast-mapping and deliberate word-learning by EFL children. *The Modern Language Journal*, *96*(3), 439-453.

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(6), 1430-1440.

Jiang, N., Novokshanova, E., Masuda, K., & Wang, X. (2011). Morphological congruency and the acquisition of L2 morphemes. *Language Learning, 61*(3), 940-967.

Juhasz, B. J., & Pollatsek, A. (2011). Lexical influences on eye movements in reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 873-894). Oxford University Press.

Juhasz, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(6), 1312-1318.

Kan, P., & Kohnert, K. (2012). A growth curve analysis of novel word learning by sequential bilingual preschool children. *Bilingualism: Language and Cognition*, *15*(3), 452-469.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1-2), 262-284.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General, 135*(1), 12-35.

Laufer, B. (1990). Why are some words more difficult than others?—Some intralexical factors that affect the learning of words. *International Review of Applied Linguistics in Language Teaching, 28*(4), 293-308.

Laufer, B. (2005). Focus on form in second language vocabulary learning. *EUROSLA Yearbook*, *5*(1), 223-250.

Loomis, R. (2010). *The influence of sentence context on reading times for abstract and concrete words* [Unpublished bachelor's thesis]. Wesleyan University, Middletown, Connecticut, USA.

Lowell, R., & Morris, R. K. (2014). Word length effects on novel words: Evidence from eye movements. *Attention Perception & Psychophysics, 76*(1), 179-189.

Lüdecke, D. (2019). sjPlot: Data visualization for statistics in social science. http://doi. org/10.5281/zenodo.1308157, R package version 2.7.2, https://CRAN.R-project.org/package=sjPlot

Ma, T. F., Chen, B. G., Lu, C. M., & Dunlap, S. (2015). Proficiency and sentence constraint effects on second language word learning. *Acta Psychologica, 159*, 116-122.

McDonald, S. (2006). Effects of number-of-letters on eye movements during reading are independent from effects of spatial word length. *Visual Cognition*, *13*(1), 89-98.

Mestres-Missé, A., Münte, T. F., & Rodriguez-Fornells, A. (2014). Mapping concrete and abstract meanings to new words using verbal contexts. *Second Language Research, 30*(2), 191-223.

Mohamed, A. A. (2018). Exposure frequency in L2 reading an eye-movement perspective of incidental vocabulary learning. *Studies in Second Language Acquisition, 40*(2), 269-293.

Mulder, E., van de Ven, M., Segers, E., & Verhoeven, L. (2019). Context, word, and student predictors in second language vocabulary learning. *Applied Psycholinguistics*, *40*(1), 137-166.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, *19*(3), 304-330.

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review*, *13*(1), 45-52.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45(*3), 255-287.

Palmer, S. D., MacGregor, L. J., & Havelka, J. (2013). Concreteness effects in single-meaning, multi-meaning and newly acquired words. *Brain Research, 1538*, 135-150.

Papagno, C., & Vallar, G. (1992). Phonological short-term-memory and the learning of novel words: The effect of phonological similarity and item length. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 44*(1), 47-67.

Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: An eye tracking study. *Studies in Second Language Acquisition, 38*(1), 97-130.

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do Things Fall Apart? *Reading in a Foreign Language, 22*(1), 31-55.

Perez, M., Peters, E., & Desmet, P. (2015). Enhancing vocabulary learning through captioned video: An eye-tracking study. *The Modern Language Journal*, *99*(2), 308-328.

Pirrelli, V., Marzi, C., Ferro, M., Cardillo, F., Baayen, H., & Milin, P. (2020). Psycho-computational modeling of the mental lexicon: A discriminative learning perspective. In V. Pirrelli, I. Plag, & W. Dressler (Eds.), *Word knowledge and word usage: A cross-disciplinary guide to the mental lexicon* (pp. 23-82). de Gruyter Mounton.

Plaut, D., & Shallice, T. (1993). Deep dyslexia: A case study of connectionist neuropsychology. *Cognitive Neuropsychology, 10*(5), 377–500.

Rayner, K., & Liversedge, S. P. (2011). Linguistic and cognitive influences on eye movements during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 751-766). Oxford University Press.

R Core Team. (2019). R: A language and environment for statistical computing (version 3.5.1). R Foundation for Statistical Computing. https://www.R-project.org.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 125-157.

Rice, G., & Robinson, D. (1975). The role of bigram frequency in the perception of words and nonwords. *Memory & Cognition*, *3*(5), 513-518.

Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition, 35*(2), 213-235.

Robitzsch, A., Kiefer, T., & Wu, M. (2019). TAM: Test Analysis Modules (version 3.3-10). Retrieved from https: //CRAN.R-project.org/package=TAM

Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, *48*(2), 281-317.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329-363.

Schwanenflugel, P. (1991). Why are abstract concepts hard to understand? In P. J. Schwanenflugel (Ed.), *The psychology of word meaning* (pp. 223-250). Erlbaum.

Simoens, H., Housen, A., & De Cuypere, L. (2017). The effect of perceptual salience on processing L2 inflectional morphology. In S. Gass, P. Spinner, & J. Behney (Eds.), *Salience in second language acquisition* (pp. 107–130). Routledge.

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and Linguistics Compass, 9*(8), 311-327.

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning, 69*(3), 559-599.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics, 28*(1), 46-65.

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, *20* (2), 232-245.

White, S., Bertram, R., & Hyönä, J. (2008). Semantic processing of previews within compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(4), 988-993.

Willis, M., & Ohashi, Y. (2012). A model of L2 vocabulary learning and retention. *The Language Learning Journal*, *40*(1), 125-137.

About the authors

Wei YI is an assistant professor at the School of Chinese as a Second Language, Peking University. He is interested in exploring cognitive mechanisms underlying second language acquisition, combining theories and techniques of psychological and neurological sciences. E-mail: weiyisla@pku.edu.cn

Shiyi LU is a tenured associate professor at the School of Chinese as a Second Language, Peking University. His research interests include second language acquisition/cognition and second language testing. E-mail: lushiyi@pku.edu.cn

Robert DEKEYSER is an emeritus professor of Second Language Acquisition at the University of Maryland, College Park, in the United States. He served as editor of *Language Learning* from 2005 to 2010, and as associate editor of *Bilingualism: Language and Cognition* from 2013 to 2017. His research interests concern primarily cognitive aspects of second language acquisition, from implicit and explicit learning mechanisms, automatization processes, and age differences in learning, to more applied concerns such as aptitude-treatment interaction, error correction, and the effects of study abroad. E-mail: rdk@umd.edu

Appendix 1. *Form Recognition Test*

Choose the words that appeared in the reading session.

Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
solemnity	supplication	cowardice	**turpentine**	magnesium	**allegory**
enigma	inebriety	**albatross**	heresy	alacrity	jeopardy
regency	sagacity	enamel	**malady**	melancholy	**blasphemy**
bungalow	epitaph	infringement	emporium	**mahogany**	**moccasin**
incursion	heroism	**canary**	**chinchilla**	**carnage**	**delirium**
reprisal	sanctity	rendezvous	dexterity	kerosene	admiral
littoral	comestible	inoculation	incongruity	**soprano**	deity
aurora	compatriot	**etiquette**	clarinet	substratum	iodine
fraternity	**pliers**	malaria	convocation	emerald	sobriety
astrolabe	emulsion	impediment	**caucus**	exactitude	consolation
bequest	firmament	ostentation	panoply	aberration	blunderbuss
monocle	cinnamon	**walrus**	**limousine**	somersault	amity
mackerel	enchanter	carnation	gondola	**mutiny**	alias
alkali	**menace**	precipitate	**equanimity**	medallion	**casket**
apricot	mutilation	**amethyst**	intermission	**satchel**	**panorama**
formaldehyde	gaiety	**oblivion**	**diadem**	asparagus	**forfeit**

Note. The blocks and the words within each block were randomized. The correct answers (target words) are bolded.

Appendix 2. *Meaning Recognition Test*

Choose the best answer that relates to the meaning of each word.

Number	Word	A	B	C	D
01	turpentine	**liquid**	tool	furniture	food
02	chinchilla	**pet**	toy	plant	tool
03	mahogany	**wood**	drink	paint	machine
04	albatross	**bird**	fish	tribe	snake
05	limousine	**car**	balloon	animal	athlete
06	moccasin	**shoe**	tool	watch	jacket
07	mackerel	**fish**	bowl	vegetable	clothes
08	amethyst	**jewelry**	plant	food	bag
09	soprano	doctor	**singer**	writer	student
10	apricot	cookie	**fruit**	perfume	soap
11	canary	plant	**bird**	chair	sculpture
12	diadem	palace	**crown**	sword	ring
13	pliers	document	**tool**	jewelry	weapon
14	casket	machine	**coffin**	animal	basket
15	walrus	person	**animal**	tree	building
16	satchel	card	**bag**	bottle	plate
17	oblivion	anxiety	pain	**forgetfulness**	wealth
18	blasphemy	claim	excuse	**insult**	strategy
19	allegory	cartoon	photo	**story**	comment
20	reprisal	rocket	command	**attack**	law
21	delirium	skill	interest	**disease**	attitude
22	panorama	gift	food	**scenery**	performance
23	equanimity	passion	elegance	**calmness**	patience
24	etiquette	story	recipe	**custom**	policy
25	enigma	concept	disaster	behavior	**mystery**
26	malady	crisis	scandal	discovery	**disease**
27	mutiny	disease	criticism	scandal	**rebellion**
28	menace	leader	hero	victim	**threat**
29	bequest	gift	document	equipment	**inheritance**
30	carnage	courage	love	mercy	**violence**
31	forfeit	reward	interview	surprise	**penalty**
32	caucus	show	game	report	**meeting**

Note. Test items were randomized. The correct answers (target words) are bolded.