Recall of Vocabulary from a Second Language: Picture Naming vs. Word Definition

Huda I. Albalawi & Sumyah A. Alnajashi*
King Saud University, Saudi Arabia

Abstract: The present study investigated the differences between early and late-acquired words from a second language in terms of the method of recall, speed and accuracy. This involved comparing the performance of 24 native Arabic participants on two different recall methods: picture naming vs. word definition. Each participant attended two sessions, with a 24-hour interval: a learning session, in which they learned new words, and a retrieval session, in which they performed the picture-naming and word-definition tasks while their ERP signals were recorded. The main results showed a significant main effect of age of acquisition on both accuracy and reaction time. Additionally, the N400 ERP component showed a significant interaction between age of acquisition and method of recall in the N400 values recorded in the left frontal area. Together, these findings indicate that both accuracy and reaction time measures can detect differences between early and late-acquired words, recalled through either a picture-naming or word-definition task.

استدعاء كلمات من لغة ثانية: تسمية الصور مقابل تعريف الكلمات

هدى إبراهيم البلوي وسمية عبدالله النجاشي*

جامعة الملك سعود، المملكة العربية السعودية

الملخص: هدفت الدراسة الحالية الى معرفة الفروق بين الكلمات المكتسبة في عمر مبكر والكلمات المكتسبة في عمر متأخر من لغة ثانية وذلك من حيث طريقة استدعاء الكلمات، والدقة، وسرعة المعالجة. وتم ذلك من خلال مقارنة أداء 24 مشاركة يتحدثن اللغة العربية باستخدام طريقتي استدعاء مختلفة: تسمية الصور مقابل تعريف الكلمة. كل مشاركة حضرت جلستين، يفصلهما 24 ساعة: جلسة التعلم، والتي تتعلم فيها المشاركة كلمات جديدة، ثم جلسة الاستدعاء والتي تؤدي فيها مهمة تسمية الصور ومهمة تعريف الكلمات. سُجّلت إشارات النشاط المرتبط بالحدث (ERP) اثناء مهمتي تسمية الصور وتعريف الكلمات. ودلت النتائج الرئيسية للتجربة على وجود تأثير رئيس دال لعمر الاكتساب على دقة الاستدعاء وزمن الرجع. أيضاً أظهرت إشارات 400 السالبة (N400) تفاعلا دالا احصائياً بين عمر الاكتساب وطريقة الاستدعاء في المنطقة اليسرى من الفص الأمامي من الدماغ. هذه النتائج تظهر أن أثر عمر اكتساب الكلمات من لغة ثانية يؤثر على دقة وسرعة معالجة الكلمات سواء كان الاستدعاء بمهمات تسمية الصور أو تعريف الكلمات.

الكلمات المفتاحية: عمر الاكتساب، تسمية الصور، الاستدعاء، تعريف الكلمات، الروابط الدلالية

*salnajashi@ksu.edu.sa

Introduction

Word acquisition is a daily activity for second language learners (De Wilde et al., 2021). However, learners might not be able to recall a suitable word when needed (Borodkin et al., 2016; Ellis & Lambon Ralph, 2000; Lu et al., 2017; Morett, 2019); hence, the level of language does not necessarily improve with the increase in vocabulary (Bates et al., 2001). Individuals recall words using various methods (Xue et al., 2015; Hao et al., 2021), such as by naming pictures or objects – tasks that require low-level semantic processing. Other tasks require deep processing of semantic information, such as concept or word definition, wherein the individual is given a description and is asked to provide the corresponding word (Fargier & Laganaro, 2017).

In this field, the effect of age of acquisition on processing words is a hot topic. Although some researchers have uncovered no effect of age of acquisition on word processing (Alario et al., 2004; Zevin & Seidenberg, 2004), others have found differences in processing early vs. late-acquired words using different paradigms (e.g., Belke et al., 2005; Bonin et al., 2004; Chen et al., 2007; Havelka & Tmita, 2006; Juhasz, 2005; Kittredge et al., 2008; Monaghan & Ellis, 2002; Morrison & Ellis, 1995; Raman, 2006; Weekes et al., 2006; Xue et al., 2017; Zevin & Seidenberg, 2002).

For example, Belke et al. (2005) studied the differences in processing speed for early and late-acquired English words among English native speakers. They asked participants to name sets of homogeneous (i.e., belonging to the same category) vs. heterogeneous pictures. Their study revealed that naming homogeneous words was faster for early-acquired than late-acquired words, emphasising the importance of age of acquisition in determining the strength of associations between concepts and linguistic characteristics. This finding is consistent in other languages. A study by Chen and colleagues (2007) found that Chinese speakers processed late-acquired Chinese words more slowly than early-acquired words. This was evident at the lexical level, as the participants were slower when reading unpredicted late-acquired words compared to early-acquired words. It was also evident in their slower reaction times when categorising late vs. early-acquired words. Together, these studies illustrate that early and late-acquired words are processed differently, especially when the explicit linking of concepts with lexicons is required (see also Damian et al., 2001). As the recall of words requires associating concepts with lexicons, comparing early

and late-acquired words is a feasible method for understanding the process of recalling words.

The effect of age of acquisition on the processing of words has also been evidenced for second-language words (Anderssen et al., 2018; Borodkin et al., 2016; Morett, 2019; Wei et al. 2015). For instance, the study by Xu and colleagues (2017) showed that Chinese speakers were faster and more accurate in judging the match of English word pairs when the words were early acquired compared to when they were late acquired, indicating more semantic information for early compared to late-acquired words. The authors used an electroencephalograph (EEG; the electrical signals of the brain) to monitor the participants while they were performing the task, alongside performance outcomes to enrich the findings with deeper insight. Combining the results of performance accuracy and reaction time with an increase or decrease of EEG in brain areas helps to understand cognitive processes. They found a larger pattern of activities, called N400, for early-acquired compared to late-acquired words. An increase in N400 indicates increased access to semantic information (Barber et al., 2013). This finding also supports the conclusion that early-acquired words have more semantic information than late-acquired words.

In a recent study, Fargier and Laganaro (2017) compared two paradigms to examine the differences in processing early and late-acquired words in French speakers. They asked participants to either name pictures or announce words corresponding to written definitions. The pictures and definitions were related to the same sets of words, which were described as either early or late-acquired words. They recorded behavioural and EEG responses during both tasks, and their results revealed the effects of age of acquisition in both tasks, manifested in a longer reaction time for the late-acquired words. Brain activities showed different patterns for early and late-acquired words in the two paradigms during a period of 100–500 s from stimulus onset. This period includes the peak time for N400, which is known to correspond to semantic processing (Barber et al., 2013; Kutas & Federmeier, 2000; Kutas & Federmeier, 2011).

As Fargier and Laganaro (2017) examined the differences between recall methods for early and late-acquired words within a native language context, it would be informative to compare the processing of early and late-acquired words from a second language. Thus, the main question of the current study is whether the difference between early and late-acquired words is revealed at the behavioural and EEG

levels when performing a word-definition task, which requires deep processing of semantic information, and a picture-naming task, which requires a lower level of semantic processing.

The rationale for this is that recent studies have shown the processing to be different in a second language than in a first language (Abdel Latif, 2019; Barbeau et al., 2017; Baumeister et al., 2017; Borodkin et al., 2016; Comesaña et al., 2010). Thus, it would be valuable to examine the effect of age of acquisition on performance in both picture naming and word definition (i.e., whether early-acquired words in a second language are superior to late-acquired words in terms of semantic processing).

Furthermore, the effects of age of acquisition on word processing corresponding to early vs. late-acquired words in a first language were revealed using neural techniques (Fiebach et al., 2003). We should thus examine whether EEG activities can detect the effect of age of acquisition on picture naming and word definition in a second language. Therefore, the current study uses EEG recording to compare the N400 component in picture naming vs. word definition for early and late-acquired English words in Arabic native speakers as there is evidence that the age of acquisition of a second language has a significant impact on brain functions (Wei et al., 2015). Differences in brain activities were also evidenced when comparing tasks requiring either surface processing or deep semantic processing (Xue et al., 2015).

With regards to the method of recall, several studies have explored what brain activities are triggered in picture-naming tasks (Damasio et al., 1996; Perret & Laganaro, 2011), revealing that picture naming activates regions in the left anterior frontal lobe (Kutas & Federmeier, 2000; Taylor et al., 2014; Quinn et al., 2017). This unique pattern of activities might arise from the need to associate concepts with lexicons when naming pictures (Taylor et al., 2014). As mentioned above, Fargier and Laganaro (2017) found differences in N400 corresponding to picture naming vs. word definition. It remains to compare the N400 component in the two methods of recall in a second language. Indeed, both picture naming and word definition require the association of concepts with lexicons (Caramazza, 1997), but word definition requires deeper access to semantic information (Krieger-Redwood et al., 2015; Fargier & Laganaro, 2017), while picture naming relies on lower-level processing (Strijkers et al., 2009).

Based on previous studies (Belke et al., 2005; Chalard & Bonin, 2006; Chen et al., 2007; Damian et al., 2001; Fargier & Laganaro, 2017), this study predicts slower reaction times and lower accuracy for late-acquired compared to early-acquired words for both the picture-naming and word-definition tasks. However, the main prediction is that late-acquired words will correspond to significantly lower performance in word definition than in picture naming. This prediction is based on the idea that the advantage of early-acquired words arises from their deep semantic roots (Bowers & Kennison, 2011). Since the word-definition task requires more semantic processing than picture naming, which relies on surface associations (Chalard & Bonin, 2006), the identification of late-acquired words may suffer from slower reaction times and reduced accuracy. With regards to the N400 component, the researchers expect that, based on the study of Xu et al. (2017), the high activation of the N400 component will be detected in the left frontal lobe for both tasks. Additionally, based on the study of Fargier and Laganaro (2017), it is expected that the activation of the N400 component will be greater for late vs. early-acquired words and when responding to the word-definition task compared to when naming pictures.

Research questions

1-Are there significant differences between the accuracy of recall for early and late-acquired words in both short and long-term recall?

2-Are there significant differences between the accuracy of recall of early and late-acquired words in both picture-naming and word-definition tasks?

3-Are there significant differences between the reaction time of recall of early and late-acquired words in both picture-naming and word-definition tasks?

4-Are there significant differences between the N400 component corresponding to the recall of early and late-acquired words in both picture-naming and word-definition tasks?

Hypotheses of the Study

1-There are significant differences between the accuracy of recall of early and late-acquired words in both short and long-term recall.

2-There are significant differences between the accuracy of recall of early and late-acquired words in both picture-naming and word-definition tasks.

3-There are significant differences between the reaction time of recall of early and late-acquired words in both picture-naming and word-definition tasks.

4-There are significant differences between the N400 component corresponding to the recall of early and late-acquired words in word-definition tasks, showing more activation for late-acquired words.

Aims of the Study

The study aims to explore the differences between early and late-acquired words in terms of accuracy and speed of recall under several conditions, namely short vs. long-term recall and picture naming vs. word definition. Similarly, it aims to explore the brain activities corresponding to the recall of early and late-acquired words in picture naming vs. word definition.

Significance of the Study

Theoretical significance. This study might encourage researchers in the psychology and psycholinguistic fields to continue combining behavioural and electrophysiological methods to obtain a deeper understanding of language learning and teaching. The study also provides researchers with a list of Arabic words with their characteristics.

Applied significance. The findings of the study can increase awareness among educators about the importance of the early acquisition of a second language, giving further scientific support to the decision to start teaching second languages to students in early grades. Additionally, the findings can inform teachers and module designers about the importance of varying the practice and quiz methods to deepen the processing of acquired words from a second language.

Method

Participants

The participants were 24 healthy females aged between 19 and 24 years (M = 21 years, 8 months; SD = 1 year, 7 months) recruited from the Psychology Department at King Saud University. An additional five participants were recruited, but their data were excluded because they either were not able to successfully complete the learning session (three participants) or did not follow the instruction to press the button when responding in the EEG recording session (two participants). All participants were monolingual native Arabic speakers who spoke English as a foreign language. All were right-handed, had normal or corrected-to-normal vision, and had no reported pre-existing neurological condition. A consent form was signed by all participants prior to the experiment, and some participants were given course credit in return for their participation.

Design

The study used a within-group experimental design. All participants performed the same sets of tasks in two sessions. There were three main variables: (a) age of acquisition (early vs. late), (b) time of recall (short-term recall vs. long-term recall), and (c) method of recall (picture naming vs. word definition).

Materials

We created a set of stimuli for this study, which included pictures of concrete objects with their names and definitions. To this end, we randomly chose a set of 100 English words, with their pictures, from a database ("The International Picture Naming Project at CRL-UCSD", 2018). All pictures depicted images of objects to rule out contamination from different categories, because the processing of object names has been found to differ from that of verbs (Bird et al., 2000). As the characteristics of the stimuli were calculated based on the responses of samples from different populations, the researchers performed three preliminary surveys to judge the pictures and their associated English words. In addition, the researchers translated the words into Arabic and created an Arabic definition for each word. The visual complexity of each picture was taken from the same database, and the lengths of the words were calculated based on the number of letters in each word. Other characteristics were taken from the abovementioned preliminary surveys. The first survey was performed by 23 women aged 18–40 years, and the participants were asked to specify the age of acquisition of each English word, its frequency, and the frequency of its corresponding word in Arabic. Responses were given on a five-point scale, in addition to a "do not know" option for the English words. The second survey was a naming agreement survey, performed by 23 women and 1 man from the same study population, who were asked to rate their agreement with the Arabic and English names of each picture on a five-point scale. The third survey rated the word definitions in terms of their capacity to capture the meaning of the words, whereby 20 participants (19 women and 1 man) from the same study population rated the definitions on a five-point scale.

We followed the method of Izura and Ellis (2002) and classified the words into early and late-acquired words according to the participants' ratings and textbooks used to teach English in schools. As a first cut-off point, the researchers chose words for which less than 40% of respondents chose "do not know" as candidates for early-acquired words and words for which more than 60% of respondents chose "do not know" as candidates for late-acquired words. Then, from the early-acquired candidates, only words acquired in the first three years of learning the English language, on average, were included. In contrast, from the late-acquired candidates, only those acquired after the third year of learning the English language were included. This resulted in 33 words for each set. Piloting the main experiment on five participants showed that participants were unable to identify three pictures from each set when they were presented on the monitor. Therefore, these items were removed, leaving 30 stimuli in each of the early and late-acquired word sets. An independent-sample t test showed a significant difference between the average acquisition ages for the two sets of words (t (58) =12.39, p < .001). It was ascertained that the two sets of words, i.e., the early and the late-acquired words, were not significantly different in their length or frequency in Arabic. The characteristics of the pictures, including their visual complexity and the naming agreement, and the definition sentence characteristics, including the agreement on the suitability of the definitions and the frequency of the last word in each sentence, were also equivalent. However, it was not possible to match the frequency of English words because by default words learned later in the process of learning a new language are not as frequent as early-acquired words. Table 1 shows the average and the standard deviation of scores for each characteristic and the values of the independent-sample t tests.

Eight additional pictures and definitions were chosen for the training phase. Additionally, each definition had an adjacent definition for another word, whereby only the final word differed in the two definitions. For instance, one sentence was "the human in his first years of life", with an adjacent definition of "the human in his last years of life" (in Arabic, the words "first" and "last" end these sentences). As it was difficult to assign each definition to an adjacent one, 15 camouflage definitions were added. The adjacent and camouflage definitions were used to prevent participants from guessing the correct answers.

Table 1. Characteristics of Each Set of Pictures and the Corresponding Words and Definitions

Items	Characteristics	Early-acquired words		Late-acquired words		T test	
		M	SD	M	SD	t	P
Pictures	Visual complexity	21098.47	9030.22	18211.3	7688.71	1.333	Ns
	Naming agreement in English	4.70	0.34	4.67	0.19	0.421	Ns
	Naming agreement in Arabic	4.70	0.46	4.87	0.16	-1.859	Ns
Words	Age of acquisition	3.31	0.83	1.14	0.47	12.388	<.001
	Length	5.23	1.45	5.7	1.57	-1.191	Ns
	Frequency in English	2.14	0.63	0.81	0.36	9.903	<.001
	Frequency in Arabic	2.45	0.89	1.75	0.89	3.014	<.01
Definition	Agreement on definition	4.67	0.34	4.59	0.27	0.923	Ns
	Frequency of last word	3.20	0.99	2.80	0.97	1.563	Ns

Procedures

The experiment started with a learning session, in which all participants were asked to perform a recall/know judgement test. This was to ensure that the participants knew more than 50% of the early-acquired words and could recall more than 30% of the words in this category. In addition, the aim was to ensure that participants could not recall more than 15% of the late-acquired words and did not know more than 50% of the words in this category. The words were presented on a 15-inch laptop screen using PsychoPy software (v.1.85; Peirce, 2007), with each word appearing in a black 70-point font on a white screen. The early-acquired words set was displayed first, followed by the late-acquired words set. Arabic words appeared sequentially, each for five seconds, and the participant had to pronounce the English translation of the Arabic word (pre-test recall). The experimenter pressed the right arrow key when the response was correct. After the participant's response, or after five seconds had elapsed, the Arabic word disappeared and its English translation appeared alongside its pronunciation with an automatic voice generated by a website (www.online-audio-converter.com). After that, the participant indicated whether she had known or had not known the word (pre-test know judgement). The experimenter pressed either the right or left arrow based on the participant's response. After the pre-test tasks, the participant was given time to review the words and their meanings and pronunciations. At the end of the learning session, each participant performed the short-term recall task on both groups of words, repeating them until she achieved 90% accuracy. We scored the first round of the short-term recall task for each participant to enable a comparison with the score for the long-term recall, which occurred in the test session (24 hours after the learning session).

The test session, which took place 24 hours after the learning session, began with the long-term recall test, which was similar to the short-term recall test. The participants were given eight training trials for each of the subsequent tasks (picture naming and word definition). The trials involved words that were different from those used in the main tasks in order to prevent the effect of familiarity with the trials in the main task.

After that, the EEG recording phase started. A 16-channel Emotiv EPOC+ BCI was connected wirelessly to the laptop used for the experiment and con-

trolled via EmotivPro software to collect the ERP signal data (BCI recording). The 16 channels included two reference channels at DRL and CMS located above the ears, and 14 recording channels distributed across the scalp: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8 and AF4. Electrodes were attached to saline-soaked felt pads to increase the conductivity of the signals. The sampling rate was 128 per second, and EEG signals were amplified via the Emotiv EPOC+ electrodes and transformed using EmotivPro software (Emotiv, 2019). During the recording of the EEG signals, all the participants performed two recall tasks, namely picture naming and word definition, on each word set. The order of the tasks was counterbalanced across participants. The presentation of the stimuli was controlled by E-Prime software (v2; Psychology Software Tools, Inc., 2016). For the picture-naming task, stimuli were presented twice in a pseudo-random order in two blocks with no break in between. The participants were asked to produce the word corresponding to each picture as quickly and accurately as possible. Each trial started with a fixation cross presented for 500 ms, followed by a blank screen for 200 ms, and then a picture (1280x720 pixels) appeared on a grey screen for 1,800 ms. A blank screen lasting 2,000 ms was displayed before the next trial or until the participant responded.

For the definition task, stimuli were presented once in a pseudo-random order. Each definition sentence was divided into three phrases, which were displayed in a 44-point font on a grey screen. Each trial was initiated with a fixation cross for 500 ms, followed by a delay of 200 ms. Then, the first phrase appeared in a white font for 600 ms. After the disappearance of the first phrase, the second phrase appeared in white font and remained for 1,250 ms. Then, 200 ms afterwards, the last single-word phrase was displayed in yellow font and remained on the screen for 3,000 ms. The next trial started after a 2,000 ms blank screen or after the participant responded. The participants were asked to produce the word corresponding to the definition as quickly and accurately as possible, and were told not to respond before they saw the yellow phrase. For example, the sentence "fat extracted from milk" was divided into three phrases, each appearing on a single screen. The first phrase was "fat", the second phrase was "extracted from", both of which were written in white, and the third phrase was "milk", written in yellow. For both tasks, participants were asked to press the down-arrow button when they

started to pronounce each word. Participants' responses were recorded by the experimenter.

Data Analysis

Behavioural data

To calculate the accuracy, the participants were given a score of 1 for correct responses and 0 for incorrect responses. As there were 30 words in each condition in our experiment, the total score for each condition was 30.

Reaction time (RT) was calculated in the picture-naming condition by calculating the time between the appearance of each picture and the participant's response. Then, the total was divided by the number of correct trials in each condition. For the word-definition task, RT was the sum of the duration starting from the appearance of the third phrase in each sentence until the response, divided by the number of correct answers in each condition.

EEG data

EEG acquisition and pre-processing. Data were recorded via EmotivPro software at a sampling rate of 128 Hz. The two reference channels were attached to the bones behind the ears (for further information about the device settings see Emotiv, 2019). Eight channels corresponding to the frontal and temporal brain regions (AF3, AF4, F3, F4, FC5, FC6, T7, and T8) were selected. The raw EEG data were band-pass filtered between 0.1 and 45 Hz using a 4th order Butterworth filter (zero-phase forward and reverse filter). An epoch was extracted for each correct response. For the picture-naming tasks, the epochs were extracted around the stimulus onset (-100 to $+500$ms). Epochs were similarly extracted for the word-definition task. Extracted epochs were visually inspected and the ones contaminated by heavy artifacts were rejected. All included participants had more than 70% intact epochs. Pre-processing was performed using FieldTrip MATLAB Toolbox (Ostenveld et al., 2011).

ERP analysis. Event-related potentials (ERP), representing brain activities around certain events (the appearance of stimuli in the current experiment), were extracted. To achieve this, epochs were grouped per condition, and the corresponding ERPs were extracted for every subject. The ERPs per subject comprise the average of all the epochs under the condition. Grand ERP is the average of all subject ERPs; in our case, the plain average was used. FieldTrip MATLAB Toolbox (Ostenveld et al., 2011) was used

to calculate the subject ERP and grand ERP signals, per condition.

Statistical Methods

A paired sample t test was used to compare performance in the two conditions, and a repeated measures analysis of variance (ANOVA) was used to compare the levels of multiple factors.

Results

To confirm the differences between the sets of early and late-acquired words, the researchers conducted a paired sample t test to verify whether there were significant differences between participants' knowledge of the early and the late-acquired words. Indeed, there was a significant difference between the pre-test recall of early-acquired words (M = 16.29, SD = 7.04) and the pre-test recall of late-acquired words (M = 3.08, SD = 4.63, t (23) = 12.17, p < .001). Additionally, in order to examine participants' knowledge of the two sets of words, the researchers compared their responses to the pre-test know condition, where they were asked to choose between "know" and "do not know" responses after the presentation of the word's meaning. There were significant differences between the early pre-test know (M = 23.79, SD = 4.48) and the late pre-test know (M = 8.79, SD = 8.78, t (23) = 12.17, p < .001).

With regards to the first question of the study, the researchers compared the participants' ability to recall the words in the short-term recall task (at the end of the learning session) and the long-term recall task (at the beginning of the testing session) (see Figure 1 for the means and standard errors in each condition). A 2x2 (early vs. late acquisition x short vs. long-term recall) repeated measures ANOVA showed a significant main effect of age of acquisition (F (1, 23) = 26.36, p < .001, MSe = 9.73, $\eta^2 p$ = .53). When collapsing across short- and long-term recall, accuracy in the recall of early-acquired words (M = 28.25, SD = 1.41) was higher than in the recall of late-acquired words (M = 24.97, SD = 3.66). Similarly, there was a significant main effect of time of recall (F (1, 23) = 36.14, p < .001, MSe = 5.08, $\eta^2 p$ = .60). When collapsing across early and late-acquired words, accuracy in short-term recall (M = 27.97 SD = 1.99) was higher than long-term recall (M = 25.25, SD = 4.68). Finally, the interaction between age of acquisition and time of recall was significant (F (1, 23) = 16.65, p < .001, MSe = 4.73, $\eta^2 p$ = 042). There was a large difference between the short-term recall (M = 28.71, SD = 1.60) and the long-term

recall (M = 27.79, SD = 2.04) of the early-acquired words (t (23) = 2.22, p < .05). Similarly, there was a significant difference between the short-term recall (M = 27.25, SD = 2.11) and the long-term recall (M = 22.71, SD = 5.21) of the late-acquired words (t (23) = 5.64, p < .001). However, it is obvious that the difference between the means of short and long-term recall in late-acquired words was larger than the difference between the means of short and long-term recall in early-acquired words.

The second question of the study was aimed at exploring the differences between early and late-acquired words in the picture-naming task and the word-definition task; see Figure 2 for the means of accuracy and standard errors in the two tasks. A 2x2 (early vs. late acquisition x picture-naming vs. word-definition recall) repeated measures ANOVA showed a significant main effect of age of acquisition (F (1, 23) = 79.39, $p < .001$, $MSe = 9.07$, $\eta^2p = .77$). The accuracy of the recall of early-acquired words ($M = 24.56$, $SD = 3.93$) was higher than that of the recall of late-acquired words ($M = 19.08$, $SD = 5.33$). In contrast, there was a nonsignificant main effect of method of recall (F (1, 23) = 7.14, p = ns, $MSe = 10.63$, $\eta^2p = 0.03$). The interaction between age of acquisition and method of recall was also nonsignificant (F (1, 23) = 0.81, $p > .05$, $MSe = 3.70$, $\eta^2p = .03$).

To answer the third question of the study, another 2x2 (age of acquisition x method of recall) repeated measures ANOVA was performed on reaction time, showing a significant main effect of age of acquisition (F (1, 23) = 32.85, $p < .001$, $MSe = 63727.69$, $\eta^2p = .58$). The reaction time for early-acquired words ($M = 1437.44$, $SD = 381.08$) was shorter than that for late-acquired words ($M = 1732.78$, $SD = 521.59$). The interaction between the age of acquisition and the method of recall was not significant (F (1, 23) = 1.13, p = ns). See Figure 3 for the means and standard errors of RT in the picture-naming and word-definition tasks.
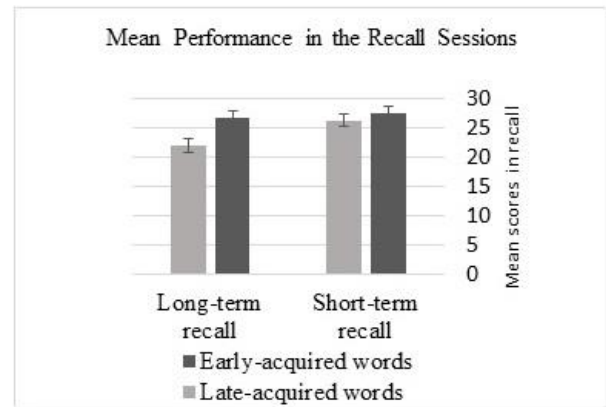

Figure 1. Accuracy in the recall of early and late-acquired words in the short and long-term recall sessions
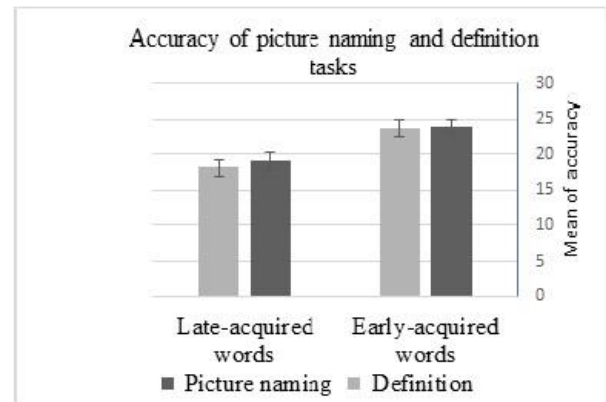

Figure 2. Accuracy in naming pictures using early and late-acquired words in the EEG recording session.
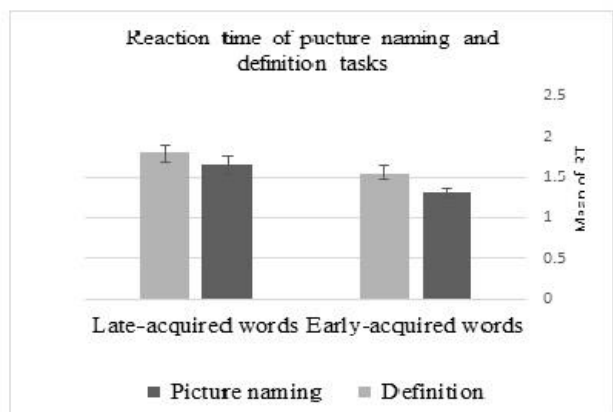

Figure 3. RT in naming pictures with early and late-acquired words in the EEG recording session.

In relation to the fourth question of the study related to N400 corresponding to the recall of early or late-acquired words in picture-naming and word-definition tasks, repeated measures analysis of variance ANOVA was performed to compare N400 across channels within each hemisphere, with age of acquisition and method of recall as the within-subject variables. For the left hemisphere, there was a main effect of channels (F $(3, 69)$ = 4.09, p = .01, MSe = 599.17, $\eta^2 p$ = .15). For the right hemisphere, there also was a main effect of channels (F $(3, 69)$ = 3.65, p = .017, MSe = 411.94, $\eta^2 p$ = .13). Neither the main effect of age of acquisition nor the main effect of method of recall were significant. See Figure 4 for the topographical images of the grand means of the ERPs. Hence, the researchers conducted a repeated measures ANOVA for each single channel, with age of acquisition and method of recall as the within-subject factors. The only significant finding was for channel F3, where a significant interaction was found between age of acquisition and method of recall (F $(1, 23)$ = 9.43, p = .005, MSe = 891.43, $\eta^2 p$ = .29). See Figure 5 for the N400 component across the four conditions in channel F3.

Figure 5. N400 component activities, recorded via channel F3, averaged across all participants within each condition. (p1) shows the picture-naming early-acquired words condition, (p2) shows picture-naming late-acquired words, (d1) shows word-definition early-acquired words, and (d2) shows word-definition late-acquired words.
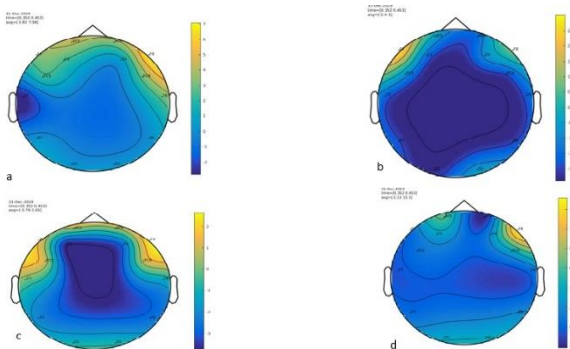
Figure 4. Topographical images of the brain activation in each channel. The values are the grand means of all epochs across all participants for each condition: (a) represents the picture-naming early-acquired words condition, (b) represents picture-naming late-acquired words, (c) represents word-definition early-acquired words, and (d) represents word-definition late-acquired words.
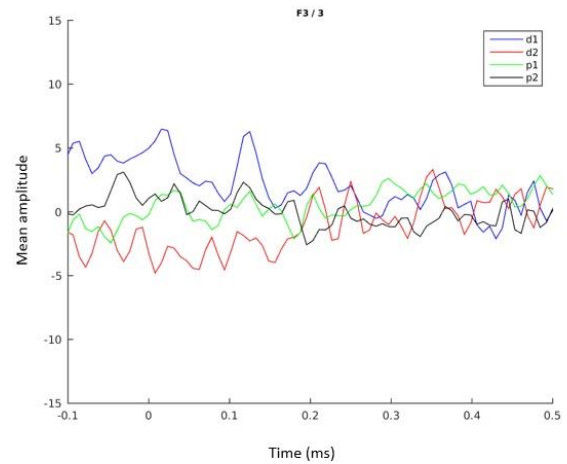
Discussion

The current study aimed to examine the differences between early and late-acquired words in picture naming vs. word definition using both behavioural and electrophysiological measures. The first question of the study focused on the differences between early and late-acquired words in short and long-term recall tasks. It was assumed that the two sets of words (early and late-acquired words) are not equivalent in terms of speed of processing or accessibility. This assumption was partially confirmed by the significant difference between early and late-acquired words in the short-term and long-term recalls.

As mentioned earlier, this research defines early-acquired words as words learned early on in the language learning process rather than words learned early on in life. The advantage of using words learned early on in the language learning process is consistent with previous findings (Izura & Ellis, 2002; Moore & Valentine, 1999). Furthermore, it is important to point to the study of Xu et al. (2017), who found a significant difference in the performance of Chinese adults when judging matches among early vs. late-acquired pairs of English words. The current study supports the superiority of early-acquired words in English for Arabic speakers who learn English as a second language.

Although a lag of 24 hours between the learning and testing sessions was found in a previous study to be adequate to consolidate the learning of new words (Li & Dekeyser, 2019; Quinn et al., 2017), we found an increase in differences between early and late-acquired words in long-term recall. This indicates persistent differences in the processing of the two groups of words.

For the purpose of understanding the differences in processing early and late-acquired words, the second and third questions compared the accuracy and RT of recall in picture-naming and word-definition tasks for both early and late-acquired words. The differences between accuracy and RT in early and late-acquired words were consistent in both tasks. Contrary to our prediction, there appeared to be no significant differences in either accuracy or RT depending on the method of recall. Although the interaction was not significant, the differences between reaction times for the two sets of words were larger in the word-definition task than in the picture-naming task.

As mentioned earlier, naming and definition involve different processes. In the picture-naming task, the participants perform two processes: the perception of the picture and the recognition of the object (Cheng et al., 2010; Fargier & Laganaro, 2015; Glaser, 1992; Kroll & Stewart, 1994). These processes are not as deep as the word-definition task, for which the participants have to recognise the words (semantic processing), then combine the words of the sentences, and finally produce the target words. Hence, early and late-acquired words were expected to show larger differences in recall in the word-definition task, where deeper semantic processing was required. Our findings showed a trend towards larger differences in RT between picture naming and word definition for late-acquired words, wherein participants took more time to respond to late-acquired words in the word-definition task. Further studies could re-examine these differences using a larger number of trials.

With regards to the results of the N400 component, which is related to peak brain activities corresponding to semantic processing, it was expected that it would show higher activation corresponding to late-acquired words, specifically in the word-definition task (Fargier & Laganaro, 2017; Xu et al., 2017). Additionally, the N400 component is believed to refer to the anterior left frontal brain region (Xu et al., 2017). Indeed, the F3 channel, corresponding to regions in the left frontal lobe, induced larger N400 activation for late-acquired compared to early-acquired words

in the word-definition task, whilst no differences between early and late-acquired words were found in the picture-naming task. This finding can be explained by the deeper semantic processes required to solve the word-definition task compared to those required for the picture-naming task (Fargier & Laganaro, 2017). It is harder for participants to access the semantic information of late compared to early-acquired words (Belke et al., 2005; Chen et al., 2007).

The findings of the current study are consistent with the findings by Xu and colleagues (2017), in that the behavioural and ERP data indicated that early-acquired compared to late-acquired words from a second language have faster and deeper access to semantic information. Both picture naming and word definition require associating lexicons to concepts (Fargier & Laganaro, 2017), and the participants in the current study were able to perform both tasks with equivalent levels of accuracy. However, their performance varied according to the type of words, i.e. early or late-acquired, whereby the accuracy for the early-acquired words surpassed that for the late-acquired words. Nevertheless, the N400 component showed an interaction between age of acquisition and method of recall, indicating that even though both picture naming and word definition are suitable tasks for the recall of early and late-acquired words in a second language, they involve different processes. Picture naming requires holistic processing while defining words requires feature processing, which requires a longer time, and this might be one reason for the differences in brain activities while performing the two tasks (Quinn et al., 2017). Hence, one should always consider the optimal method of recall when learning words from a second language as tasks have different demands and require different processes.

The results have some limitations as it was difficult for the Arabic native speakers to distinguish early from late-acquired English words. This is because English is common in daily life and each individual's personal experience of English varies. Additionally, we speculate that we failed to detect a significant interaction between age of acquisition and method of recall in accuracy and RT because of the small number of trials.

Limits of the study

The experiment was performed in the Psychology Department at King Saud University, and all participants were female undergraduates from this department. There are thus limitations on the generalizabil-

ity of the findings because of the small number of participants. Additionally, the findings might lack ecological validity as they are based on a lab experiment. The characteristics of the stimuli used in this study were determined using a small number of participants, and the EEG recording was performed using a 16-channel device. Therefore, we should acknowledge that the results are also limited by the instruments used in the study.

Suggestions and recommendations

1-Further studies should test the ecological validity of these findings in a context where participants have to communicate their thoughts rather than recall a series of words learned in a lab.

2-Future studies might apply the picture naming vs. word definition paradigm to school students, for whom differences in exposure to the English language can be estimated and controlled.

3-Cross-cultural studies are required as Crepaldi et al. (2012) found that age of acquisition has different effects on different languages.

4-Educators and language teachers need to use various methods to enhance the recall of learned words as each method triggers a different level of processing.

5-English language learners can expose themselves to lists of words prior to speaking as this can facilitate the recall of words from memory within a short period.

Declaration

Conflict of interest. The author declares that there is no conflict of interest in publishing the research results.

Ethical approval. All procedures performed in the study were in accordance with the ethical standards of the institutional research committee [Deanship of Scientific Research, King Saud University: 3-19-1439].

Informed consent. Informed consent was obtained from all participants in the study.

Data availability. The data pertaining to this study are available from the author upon reasonable request.

References

Abdel Latif, M. (2019). Eye-tracking in recent L2 learner process research: A review of areas, issues, and methodological approaches. *System, 83*, 25-35. https://doi.org/10.1016/j.system.2019.02.008

Alario, F., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers, 36*(1), 140-155. http://www.doi.org/10.3758/bf03195559

Anderssen, M., Bentzen, K., Busterud, G., Dahl, A., Lundquist, B., & Westergaard, M. (2018). The acquisition of word order in L2 Norwegian: The case of subject and object shift. *Nordic Journal of Linguistics, 41*(3), 247-274. https://doi.org/10.1017/s0332586518000203

Barbeau, E., Chai, X., Chen, J., Soles, J., Berken, J., & Baum, S. et al. (2017). The role of the left inferior parietal lobule in second language learning: An intensive language training fMRI study. *Neuropsychologia, 98*, 169-176. http://www.doi.org/10.1016/j.neuropsychologia.2016.10.003

Barber, H., Otten, L., Kousta, S., & Vigliocco, G. (2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain and Language, 125*(1), 47-53. http://www.doi.org/10.1016/j.bandl.2013.01.005

Bates, E., Burani, C., D'Amico, S. & Barca, L. (2001). Word reading and picture naming in Italian. *Memory & Cognition, 29*(7), 986-999. DOI: 10.3758/bf03195761

Baumeister, J., Foroni, F., Conrad, M., Rumiati, R., & Winkielman, P. (2017). Embodiment and Emotional Memory in First vs. Second Language. *Frontiers in Psychology, 8*. http://www.doi.org/10.3389/fpsyg.2017.00394

Belke, E., Brysbaert, M., Meyer, A., & Ghyselinck, M. (2005). Age of acquisition effects in picture naming: Evidence for a lexical-semantic competition hypothesis. *Cognition, 96*(2), B45-B54. http://www.doi.org/10.1016/j.cognition.2004.11.006

Bird, H., Howard, D., & Franklin, S. (2000). Why Is a Verb Like an Inanimate Object? Grammatical Category and Semantic Category Deficits. *Brain and Language, 72*(3), 246-309. http://www.doi.org/10.1006/brln.2000.2292

Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story?. *Journal of Memory and Language, 50*(4), 456-476. http://www.doi.org/10.1016/j.jml.2004.02.001

Borodkin, K., Kenett, Y., Faust, M., & Mashal, N. (2016). When pumpkin is closer to onion than to squash: The structure of the second language lexicon. *Cognition, 156*, 60-70. http://www.doi.org/10.1016/j.cognition.2016.07.014

Bowers, J., & Kennison, S. (2011). The Role of Age of Acquisition in Bilingual Word Translation: Evidence from Spanish-English Bilinguals. *Journal of Psycholinguistic Research, 40*(4), 275-289. http://www.doi.org/10.1007/s10936-011-9169-z

Caramazza, A. (1997). How Many Levels of Processing Are There in Lexical Access?. *Cognitive Neuropsychology, 14*(1), 177-208. https://doi.org/10.1080/026432997381664

Chalard, M., & Bonin, P. (2006). Age-of-acquisition effects in picture naming: Are they structural and/or semantic in nature?. *Visual Cognition, 13*(7-8), 864-883. http://www.doi.org/10.1080/13506280544000084

Chen, B., Zhou, H., Dunlap, S., & Perfetti, C. (2007). Age of acquisition effects in reading Chinese: Evidence in favour of the arbitrary mapping hypothesis. *British Journal of Psychology, 98*(3), 499-516. http://www.doi.org/10.1348/000712606x165484

Cheng, X., Schafer, G., & Akyürek, E. (2010). Name agreement in picture naming: An ERP study. *International Journal of Psychophysiology, 76*(3), 130-141. http://www.doi.org/10.1016/j.ijpsycho.2010.03.003

Comesaña, M., Soares, A., & Lima, C. (2010). Semantic representations of new cognate vs. noncognate words: Evidence from two second language learning methods. *Procedia - Social and Behavioral Sciences, 5*, 199-203. http://www.doi.org/10.1016/j.sbspro.2010.07.072

Crepaldi, D., Che, W., Su, I., & Luzzatti, C. (2012). Lexical-semantic variables affecting picture and word naming in Chinese: A mixed logit model study in aphasia. *Behavioural Neurology, 25*(3), 165-184. http://www.doi.org/10.1155/2012/178401

Damasio, H., Grabowski, T., Tranel, D., Hichwa, R., & Damasio, A. (1996). Erratum: A neural basis for lexical retrieval. *Nature, 381*(6585), 810-810. http://www.doi.org/10.1038/381810b0

Damian, M., Vigliocco, G., & Levelt, W. (2001). Effects of semantic context in the naming of pictures and words. *Cognition, 81*(3), 77-86. http://www.doi.org/10.1016/s0010-0277(01)00135-4

De Wilde, V., Brysbaert, M., & Eyckmans, J. (2021). Formal versus informal L2 learning. *Studies in Second Language Acquisition, 44*(1), 87-111. https://doi.org/10.1017/s0272263121000097

Ellis, A., & Lambon Ralph, M. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal Of Experimental Psychology: Learning, Memory, And Cognition, 26*(5), 1103-1123. http://www.doi.org/10.1037/0278-7393.26.5.1103

Emotiv. [EPOC+ Emotiv]. (2019). EPOC+ Emotiv user manual. https://emotiv.gitbook.io/epoc-user-manual/

Fargier, R., & Laganaro, M. (2015). Neural dynamics of object noun, action verb and action noun production in picture naming. *Brain and Language, 150*, 129-142. http://www.doi.org/10.1016/j.bandl.2015.09.004

Fargier, R., & Laganaro, M. (2017). Spatio-temporal dynamics of referential and inferential naming: Different brain and cognitive operations to lexical selection. *Brain Topography, 30*(2), 182-197. http://www.doi.org/10.1007/s10548-016-0504-4

Fiebach, C., Friederici, A., Müller, K., von Cramon, D., & Hernandez, A. (2003). Distinct brain representations for early and late learned words. *Neuroimage, 19*(4), 1627-1637. http://www.doi.org/10.1016/s1053-8119(03)00227-1

Glaser, W. (1992). Picture naming. *Cognition, 42*(1-3), 61-105. http://www.doi.org/10.1016/0010-0277(92)90040-o

Hao, S., Liang, L., Wang, J., Liu, H., & Chen, B. (2021). The effects of emotional context and exposure frequency on L2 contextual word learning. *International Journal of Bilingualism, 25*(5), 1280-1296. https://doi.org/10.1177/13670069211016540

Havelka, J., & Tomita, I. (2006). Age of acquisition in naming Japanese words. *Visual Cognition, 13*(7-8), 981-991. http://www.doi.org/10.1080/13506280544000156

IPNP Website (2018). *The International Picture Naming Project at CRL-UCSD."*. Retrieved from https://crl.ucsd.edu/experiments/ipnp/actobj/getpics/getpics.html

Izura, C., & Ellis, A. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicológica, 23*(2), 245–281. https://www.uv.es/psicologica/articulos2.02/4.IZURA%26ELLIS.pdf

Juhasz, B. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin, 131*(5), 684-712. http://www.doi.org/10.1037/0033-2909.131.5.684

Kittredge, A., Dell, G., Verkuilen, J., & Schwartz, M. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology, 25*(4), 463-492. http://www.doi.org/10.1080/02643290701674851

Krieger-Redwood, K., Teige, C., Davey, J., Hymers, M., & Jefferies, E. (2015). Conceptual control across modalities: graded specialisation for pictures and words in inferior frontal and posterior temporal cortex. *Neuropsychologia, 76*, 92-107. http://www.doi.org/10.1016/j.neuropsychologia.2015.02.030

Kroll, J., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory And Language, 33*(2), 149-174. http://www.doi.org/10.1006/jmla.1994.1008

Kutas, M., & Federmeier, K. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463-470. http://www.doi.org/10.1016/s1364-6613(00)01560-6

Kutas, M., & Federmeier, K. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*(1), 621-647. http://www.doi.org/10.1146/annurev.psych.093008.131123

Li, M., & Dekeyser, R. (2019). Distribution of practice effects in the acquisition and retention of L2 Mandarin tonal word production. *The Modern Language Journal, 103*(3). https://doi.org/10.1111/modl.12580

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

Mobaghan, J., & Ellis, A. (2002). Age of acquisition and the completeness of phonological representations. *Reading and Writing, 15*(7/8), 759-788. http://www.doi.org/10.1023/a:1020958722472

Moore, V., & Valentine, T. (1999). *The effects of age of acquisition in processing famous faces and names: Exploring the locus and proposing a mechanism.* Proceedings of the Twenty-First Annual Meeting of the Cognitive Science Society, Mahwah: Lawrence Erlbaum Associates, 416-421.

Morett, L. (2019). The power of an image: Images, Not glosses, enhance learning of concrete L2 words in beginning learners. *Journal of Psycholinguistic Research, 48*(3), 643-664. https://doi.org/10.1007/s10936-018-9623-2

Morrison, C., & Ellis, A. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*(1), 116-133. http://www.doi.org/10.1037/0278-7393.21.1.116

Oostenveld, R., Fries, P., Maris, E. & Schoffelen, J. M. (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Compuationalt Intelligence & Neuroscience, 2011,* 156869. doi:10.1155/2011/156869

Peirce, J. W. (2007). PsychoPy - Psychophysics software in python. *Journal of Neuroscience Methods, 162*(1-2), 8-13. http://dx.doi.org/10.1016/j.jneumeth.2006.11.017

Perret, C., & Laganaro, M. (2011). Comparison of electrophysiological correlates of writing and speaking: A topographic ERP analysis. *Brain Topography, 25*(1), 64-72. http://www.doi.org/10.1007/s10548-011-0200-3

Psychology Software Tools, Inc. [E-Prime 2.0]. (2016). Retrieved from http://www.pstnet.com

Quinn, C., Taylor, J., & Davis, M. (2017). Learning and retrieving holistic and componential visual-verbal associations in reading and object naming. *Neuropsychologia, 98*, 68-84. http://www.doi.org/10.1016/j.neuropsychologia.2016.09.025

Raman, I. (2006). On the age-of-acquisition effects in word naming and orthographic transparency: Mapping specific or universal?. *Visual Cognition, 13*(7-8), 1044-1053. http://www.doi.org/10.1080/13506280500153200

Strijkers, K., Costa, A., & Thierry, G. (2009). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex, 20*(4), 912-928. http://www.doi.org/10.1093/cercor/bhp153

Taylor, J., Rastle, K., & Davis, M. (2014). Distinct neural specializations for learning to read words and name objects. *Journal of Cognitive Neuroscience, 26*(9), 2128-2154. http://www.doi.org/10.1162/jocn_a_00614

Weekes, B., Castles, A., & Davies, R. (2006). Effects of consistency and age of acquisition on reading and spelling among developing readers. *Reading and Writing, 19*(2), 133-169. http://www.doi.org/10.1007/s11145-005-2032-6

Wei, M., Joshi, A., Zhang, M., Mei, L., Manis, F., & He, Q. et al. (2015). How age of acquisition influences brain architecture in bilinguals. *Journal of Neurolinguistics, 36*, 35-55. http://www.doi.org/10.1016/j.jneuroling.2015.05.001

Xu, X., Kang, C., & Guo, T. (2016). Imageability and semantic association in the representation and processing of event verbs. *Cognitive Processing, 17*(2), 175-184. http://www.doi.org/10.1007/s10339-015-0747-0

Xue, J., Liu, T., Marmolejo-Ramos, F., & Pei, X. (2017). Age of acquisition effects on word processing for Chinese native learners' English: ERP evidence for the arbitrary mapping hypothesis. *Frontiers in Psychology, 8*:818, 1-14 http://www.doi.org/10.3389/fpsyg.2017.00818

Xue, J., Marmolejo-Ramos, F., & Pei, X. (2015). The linguistic context effects on the processing of body–object interaction words: An ERP study on second language learners. *Brain Research, 1613*, 37-48. DOI:10.1016/j.brainres.2015.03.050

Zevin, J., & Seidenberg, M. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language, 47*(1), 1-29. http://www.doi.org/10.1006/jmla.2001.2834

Zevin, J., & Seidenberg, M. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory & Cognition, 32*(1), 31-38. http://www.doi.org/10.3758/bf03195818