The Effects of Event Depictions in Second Language Phrasal Vocabulary Learning

Dissertation

zur Erlangung des akademischen Grades

Doktorin der Philosophie (Dr. phil.)

Eingereicht an der Sprach- und literaturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin
von

Huong Thi Thu Nguyen, M.A.

Disputation: 29. März 2021

Gutachter/Innen:
1. Prof. Dr. Pia Knoeferle
2. Prof. Dr. Natalia Gagarina

Prof. Dr. -Ing. Dr. Sabine Kunst
Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr. Stefan Kipf
Dekan der Sprach- und literaturwissenschaftlichen Fakultät

Acknowledgements

My profound thanks go to my supervisor, Prof. Dr. Pia Knoeferle. Approximately five years ago, she gave me an opportunity and said that I would have a big adventure. I have not given up on the adventure because of her. She supported me whenever I met difficulties or needed advice or suggestions. My work benefited hugely from her encouragement, scientific advice in developing my interests, understanding, and humanity.

I am also very grateful to Dr. Katja Maquate for her help in my scientific and daily life. I want to thank Carsten Schliewe for his technical assistance. My next special thanks go to Linh Thi Dieu Nguyen for recording the auditory stimuli in Vietnamese for all experiments. I acknowledge the financial support by the Vietnam International Education Development, Vietnam Ministry of Education and Training (Project 911).

I enjoyed the collaborative and friendly atmosphere in the HU Psycholinguistic Lab, where I could always be sure to receive help when needed, and hence wish to thank the team: Katja Maquate, Dato Abshidze, Camilo Rodríguez Ronderos, Daniela Palleschi, and Aine Ito. I am also grateful to the psycholinguistic student assistants who have assisted with collecting image materials, recording stimuli, and analyzing IQ subtests: Luise Henneberg, Oleksandra Butenko, Melis Odabaş, and Minying Feng.

Finally, I wish to acknowledge my friends, my family, my husband, Viet Anh, and my daughter, Linh Giang, for their support and patience throughout the last five years. They kept me going on, and this work would not have been possible without their input.

Zusammenfassung

In früheren Studien zum L2-Wortschatzerwerb wurden die Auswirkungen des visuellen Kontexts auf das Lernen und die Verarbeitung von Wörtern und Kollokationen in der L2 untersucht. Es wurde festgestellt, dass die Erstsprache einen positiven Transfer auf das Lernen einer Zweitsprache hat, wenn die Wörter Ähnlichkeiten aufweisen. Darüber hinaus wurden die Einflüsse der kognitiven Fähigkeiten der Lernenden und ihres Erwerbsalters (AoA) auf das L2-Vokabellernen unter verschiedenen Bedingungen des L2-Vokabellernens festgestellt. Ziel der vorliegenden Arbeit war es, die Auswirkungen des visuellen Kontexts und des Transfers auf das Lernen von L2-Vokabeln weiter zu untersuchen und zu klären, wie die kognitiven Fähigkeiten und das Erwerbsalter diese Auswirkungen in einem bestimmten L2-Lernkontext beeinflussen. Im Detail wurden Effekte der Ereignisdarstellung (d.h. nicht-sprachlicher visueller Kontext) untersucht sowie Transfereffekte aus der Erstsprache in die Zweitsprache im Bezug auf das Lernen von L2-Phrasenwortschatz (d.h. Verb-Nomen-Phrasen) bei erwachsenen Anfängern. Wir führten Kurzzeitexperimente zum L2-Wortschatzerwerb durch, bei denen wir die Reaktionszeiten maßen. Zwei weitere Forschungsfragen untersuchten, ob es Zusammenhänge zwischen der AoA oder den kognitiven Fähigkeiten der Lernenden und ihrem Lernerfolg beim Vokabellernen in einer kurzfristigen L2-Lernumgebung gibt. Die Ergebnisse zeigten, dass erwachsene L2-Anfänger*innen beim L2-Vokabellernen von visuellen Darstellungen profitierten: Sie waren unter Lernbedingungen mit Ereignissen genauer und schneller als unter Lernbedingungen ohne Ereignisse. Diese Effekte konnten in drei Experimenten nicht nur mit jungen Erwachsenen im Alter von 18 bis 31 Jahren nachgewiesen werden, sondern galten auch für Erwachsene im frühen und späten mittleren Alter von 32 bis 65 Jahren. Die vorangegangene Forschung deutete darauf hin, dass die Ähnlichkeit zwischen L1 und L2 das L2-Lernen beeinflussen könnte, jedoch nicht in diesem spezifiscchen L2-Lernkontext. Darüber hinaus wurde der AoA der Probanden manipuliert, was dazu führte, dass junge Erwachsene in den kognitiven Tests und bei den L2-Lernaufgaben besser abschnitten als die anderen beiden Gruppen. Basierend auf den Ergebnissen unserer Forschung konnten wir herausfinden, welche Faktoren den Erfolg des L2-Wortschatzerwerbs bei erwachsenen L2-Anfängern stark beeinflussen und dass das Lernen von L2-Phrasenwortschatz mit dargestellten Ereignisfotos angewendet werden kann.

Abstract

Previous studies of L2 vocabulary learning presented visual context effects on L2 word and collocation learning and processing. It was found that L1 has a positive transfer in L2 learning when words have similarities. Furthermore, the influences of learners' cognitive ability and their age of acquisition (AoA) in L2 vocabulary learning have been found in diverse L2 vocabulary learning conditions. The present dissertation aimed to further investigate the effects of visual context and transfer on L2 learning, as well as how cognitive ability and AoA influence any such effects in a particular L2 vocabulary learning context. In detail, we investigated event depiction (i.e., non-linguistic visual context) effects and L1–L2 transfer effects on L2 phrasal vocabulary (i.e., verb-noun phrases) learning for adult beginners. We conducted short-term L2 vocabulary learning experiments during which we measured reaction times. Two other research questions examined whether there are relationships between learners' AoA or their cognitive ability and their L2 vocabulary learning success in a short-term L2 learning setting. Results showed adult L2 beginners benefited from visual depictions in L2 vocabulary learning: They were more accurate and faster in event-present learning conditions than in event-absent learning conditions. These effects were not only replicated with young adults aged 18 to 31 in three experiments but they also extended to early and late middle-aged adults aged 32 to 65. The prior research suggested that the L1–L2 similarity might influence L2 learning, but not in our L2 learning context. In addition, the AoA of subjects was manipulated, which resulted in young adults performing in the cognitive test and L2 learning tasks best compared to the other two groups. Based on the findings of our research, we were able to identify which factors strongly influence L2 vocabulary learning success for L2 adult beginners, and whether L2 phrasal vocabulary learning with depicted event photographs can be applied.

Contents

List of Figures

List of Tables

1. Introduction

1.1 Thesis Motivation

Vocabulary is one of the most crucial ingredients of any language, and beginners always focus on building their vocabulary in the early stages of second language learning. Without sufficient vocabulary, language learners cannot understand others or state their thoughts in communication. Thus, vocabulary is fundamental to all language use because “while without grammar, very little can be conveyed, without vocabulary, nothing can be conveyed” (Wilkin, 1972, pp.111–112). Vocabulary is determined as single words, phrases, or chunks of several words that can carry a specific meaning (Lessard-Clouston, 2013). Successfully acquiring an L2 word means that learners know many aspects of the L2 word, such as its meaning, written and spoken forms, frequency, collocations, etc. Which words (i.e., vocabulary selection) and how many words or phrases (i.e., vocabulary size) L2 learners should be taught depend on their needs, their L2 level (i.e., beginner, immediate, or advanced), and their L2 learning goals (Zimmerman, 1997). More importantly, it is believed that learners can remember a word better through the three significant processes of noticing, retrieval, and generative use (Nation, 2001). While explicit learning mechanisms are efficient for the meaning aspects of vocabulary learning, implicit learning mechanisms are more suitable for the form aspects of vocabulary learning (Ellis, 1994). Explicit versus implicit learning is also a crucial research topic for vocabulary learning in second language learning. Other topics include the role of L1 in L2 vocabulary learning, the task effect on L2 vocabulary learning, the relationship between vocabulary knowledge and language proficiency, testing vocabulary knowledge, vocabulary learning strategies, the role of word frequency in L2 vocabulary learning, the construct of vocabulary knowledge, and learning vocabulary via electronic media and computer-assisted techniques (Bogaards & Laufer, 2004).

This research is inspired by previous studies related to L1–L2 transfer in L2 vocabulary learning and processing (Jiang, 2002; Meade et al., 2018). In particular, L2 beginners often have the advantage of L1–L2 correspondence compared to the interferences of L1–L2 differentiation (Ellis, 1985; Gass & Selinker, 2001; Odlin, 1989). Vocabulary learning is also motivated by visual context effects on real-time language processing in adults (Knoeferle et al., 2005; Knoeferle, 2015; Tanenhaus et al., 1995) and children (Münster, 2016; Trueswell et al.,

1999) as well as the important role of visual context in cross-situational language learning (Koehne et al., 2015; MacDonald et al., 2017; Yu et al., 2011). Moreover, individual differences (i.e., age-related differences and cognitive ability) are also included in the research motivation since they might explain differences in L2 learning success among age groups (Muñoz, 2008).

In our study, we investigate whether L1 German adults (aged from 18 to 65) experience facilitation in learning L2 Vietnamese verb-noun phrases (i.e., phrasal vocabulary) from (a) event photograph presence (present vs. absent) and (b) verb mapping (similar vs. different between L1 and L2). We carried out short-term L2 vocabulary learning experiments in the laboratory for different adult groups.

1.2 Thesis Aims

In six reaction time studies, this thesis will investigate the effect of nonlinguistic visual context (event depictions/depicted actions), language similarity effects (L1–L2 language similarity), and their possible interaction during the L2 learning and testing process. We will focus on the question of whether these two visual and linguistic effects can facilitate L2 vocabulary learning. We will also investigate whether the age and the cognitive ability (WAIS scores) of L2 learners modulate the use/application of visual and linguistic cues for L2 learning. Moreover, we will investigate if we can replicate visual and linguistic effects in different learning contexts of L2 learning and assessment.

1.3 Thesis Outline

To further motivate our research, we will first discuss the diverse L2 vocabulary learning paradigms and testing tasks (Section 2). Section 2 will focus on which levels of L2 vocabulary are often learned and how words or phrasal vocabulary are learned. We will focus on a word-referent pair learning paradigm along with corresponding testing tasks.

Section 3 will review the effects of visual context in L2 language learning, processing, and comprehension. Having looked at the critical role of L1 on L2 vocabulary learning, in Section 4, we will examine language transfer in the L2 learning process. We will outline and discuss how the similarities and differences between L1 and L2 affect L2 vocabulary learning and processing. Section 5 and Section 6 will consider how age-related and intelligence-related differences interact with the L2 learning process and L2 learning success.

We present the research questions in Section 7. We will also describe the general design of the experiments, how we collected and analyzed the data, and the predictions of the research before conducting.

Sections 8 to 10 will present the studies grouped by the research questions. We will start in Section 8 with the first experimental design (Experiment 1: learning L2 in visual context and testing by matching a verb audio sound with a suitable object picture to complete a verb-noun phrase) compared to the innovative follow-up design (Experiment 2, 2R, 3, and 4: learning L2 in visual context and testing event depictions of verb-noun phrases directly – hearing a verb-noun phrase sound and then choosing one of two event photographs) to see where and why the effects of experimental factors (i.e., event photographs and language mapping) on L2 vocabulary learning success can be found. Section 9 will look more closely at differences in participants' ages (i.e., young adults, early middle-aged adults, and late middle-aged adults) and their cognitive abilities and their L2 learning success. Section 10 will investigate whether different L2 learning and testing paradigms (e.g., immediate vs. delayed testing) can change the main effects of language similarities and depicted actions on L2 learning success.

Taking previous research into account, we will discuss each of our findings in Section 11. Furthermore, based on the outcomes of our research, we will point out potential suggestions for further research related to adults learning L2 phrasal vocabulary. Lastly, we will give our conclusions and review the importance of our findings for (psycho)linguistic research.

2. Second Language Vocabulary Learning and Assessment

2.1 Second Language Learning and Acquisition

2.1.1 Second Language and Foreign Language

Many scholars (Dulay et al., 1982; Liao, 1996; Skehan, 2000) have applied both terms, *second language* and *foreign language,* to a target language people learned or acquired after their native language (also first language, or mother language, or primary language). The most significant difference between the two mentioned terms (Gass, 2013; Saville-Troike, 2006) is where the target language is acquired. A *second language* is often learned or acquired by non-native speakers in the environment that language is spoken (e.g., German native speakers learning English in the UK). A *foreign language* is studied outside of its natural language environment (e.g., German speakers learning English in Germany). The term *second language* is used in our current research, covering the meaning of the term *foreign language* because we wanted to reference a wide range of relevant studies.

2.1.2 Language Learning and Language Acquisition

Krashen (1981), Kramina (2000), and Galasso (2002) identified *language acquisition* as a subconscious process. It is similar to the process children apply in acquiring their first language. Contrarily, *language learning* is a conscious process that results in knowing about another language in a formally situated learning context. Language learners, *in language acquisition,* can understand messages in the natural communicative form of the target language, and they are not concerned about the utterance forms. Language learners*, in language learning*, proceed from the simple to the complex in an instructed learning context with error correction. According to Gass (2013), second language acquisition (SLA) refers to learning a non-native language. Researchers have been concerned about how learners acquire their L2 knowledge, what form it takes, and why some people are better at SLA than others. Gass also used SLA as a broad term involving both second language learning and foreign language learning.

In our research, we use the term *second language learning* because participants in our experiments learned L2 and were tested in a specific language learning context set up on a computer. Our learning experiment design also refers to a conscious (short-term) learning process in a laboratory.

2.2 Second Language Vocabulary Learning and Testing Tasks

Learning vocabulary is a crucial part of mastering a second language. Generally, L2 learners' lexical knowledge involves knowledge of the morphophonological, semantic, collocational, grammatical, and associational aspects of the word (e.g., Nation, 2001; Ringbom, 1987). Nation (2001) presented the most primary classification of the range of word knowledge aspects such as word form, word meaning, and word usage. To know a word for L2 language beginners, a variety of primary questions should be asked:

- What is the spoken and written form of the word?
- What meaning does the word form signal?
- What words or types of words occur with this word?
- Where, when, and how often would we expect to meet this word?

For L2 beginners, when they learn an L2 word, they want to know what the word looks like and what it sounds like. Thus, a word sound combined with a referent (i.e., a picture) has become a simple and effective way to help them understand the word meaning. As a next step, learners want to know what other words frequently collocate with the word in question, and how often the word occurs in different contexts.

According to Smith (2008), various sorts of word knowledge are essential foundations for researching word acquisition processes and pedagogy. In the same review of L2 vocabulary learning research, Smith (2008) summarized six main points that are useful for L2 vocabulary instruction. One of them was the use of an intentional learning method for L2 beginners. That method helped to build an initial L2 form-meaning link at the beginning of the vocabulary acquisition process. Then, it was necessary to have repeat exposures for the consolidation and improved word knowledge (i.e., collocation) in diverse contexts, including incidental learning. Smith listed examples of L2 vocabulary learning activities for intentional vocabulary learning, such as:

- Use an online database including examples, a dictionary, and a quiz game (Horst et al., 2005)

- Use an Internet chat program, for instance, a picture story task, and a decision-making task (Smith, 2004)
- Use L2–L1 translation and produce a sentence (Webb, 2005)
- Produce new word forms before getting correct materials (Barcroft, 2007)
- Pick words from a text and use them to convey related ideas (Joe, 1998)
- Process words without a context (Princess, 1996)

There are diverse L2 vocabulary learning methods (e.g., translation, reading in context, or keywords) and L2 assessment types (e.g., recall, recognition, or production) in different learning environments (e.g., in natural settings, in classrooms, in specialized language programs, or in a laboratory). In the following sections, we focus on previous studies of L2 vocabulary learning of words (Section 2.2.1) and L2 phrasal items (Section 2.2.2) in visual environments to examine which learning designs and testing methods were applied.

2.2.1 Second Language Word Learning and Testing

A well-known L2 vocabulary instruction method involves the use of pictures (e.g., Nelson, 1979; Paivio, 1991; Paivio & Csapo, 1973). We investigated studies of L2 word-referent learning: learning an L2 word via L2–L1 translation or via an object picture as a referent. In this section, we aimed to know:

- how L2 separate words have been learned and tested
- which L2 word-learning methods were better than others
- which types of L2 learners were tested

Table 1 presents brief summaries of seven recent studies of L2 word-referent learning from 2012 to 2019 before reviewing them in detail. All the studies focused on L2 word learning in visual contexts. Participants learned L2 words successfully, mostly in immediate and delayed testing in different visual learning conditions. We identified some findings that motivated our research. L2 learners had higher accuracy in picture-based instruction (Study 2), but they were faster and had fewer errors in word-based instruction (Study 1). Learning L2 words in auditory-visual conditions was the most helpful and active context for learners compared to auditory-only or visual-only conditions (Studies 3 and 4). Moreover, specific features of referents were important for learners. L2 learners could more successfully learn L2 words in referent-familiar and phonologically familiar learning conditions than in unfamiliar conditions or without referents (Studies 5, 6, and 7).

Table 1: A summary of previous studies of L2 word-referent learning

Authors	**Participants, Learning and Testing methods**	**Main Results Which Learning Method is Better?**
Study 1: Comesaña et al. (2012)	➢ 48 native L1 Portuguese children (mean age = 10.87 years); no prior L2 knowledge ➢ 42 L2 Basque words (cognates vs. non-cognates); picture-based method vs. word-based method ➢ translation recognition task; immediate vs. delayed testing (one week after)	➢ Accuracy: no significant difference between the two learning methods ➢ Reaction times and errors: the word-based method group <[1] picture-based method group
Study 2: Emirmus-tafaoğlu & Gökmen (2015)	➢ 75 native L1 Turkish students (7th grade); L2 English at elementary level ➢ 20 unknown L2 English words; word-based instruction vs. picture-based instruction ➢ providing L2 written words; immediate vs. delayed 1 (one week after) vs. delayed 2 (one month after)	Accuracy was measured in both immediate and delayed: word-based instruction < picture-based instruction
Study 3: Nassaji (2012)	➢ 79 L1 Farsi adults (19–24 years) in three groups ➢ 24 nonsense words of L2 English; auditory only vs. visual only vs. dual modality ➢ recall and recognition tests	Mean scores of both tests: visual only < auditory only < the auditory-visual condition
Study 4: Wang et al. (2017)	➢ 71 L1 English speakers ➢ 24 L2 Mandarin words; familiar vs. novel-object condition ➢ recognition retention tests	Accuracy via mouse clicking: Participants were successful in learning L2 words and

[1] shorter reaction time; fewer errors or lower accuracy/scores between/among learning methods

		retention of the L2 target word-object pairs for both learning conditions
Study 5: Kaushanskaya et al. (2013)	➢ 81 L1 English adults in two groups; prior L2 Spanish knowledge ➢ 48 L2 Spanish words were learned in four different conditions: phonological familiarity (familiar vs. unfamiliar) × referent familiarity (familiar vs. unfamiliar) ➢ forced-choice recognition task	Correct choices: familiar referent and phonologically unfamiliar < familiar referent and phonologically familiar
Study 6: Havas et al. (2017)	➢ 68 L1 Spanish participants (18–36 years: no prior L2 Hungarian knowledge); four learner groups: L1 –sleep, L1 +sleep, L2 –sleep, and L2 +sleep ➢ 144 novel words (L1 phonologically changed and L2); three learning conditions: familiar referent vs. unfamiliar referent vs. no referent ➢ auditory recognition memory test; four alternative forced-choice word-choice matching tasks; semantic priming task	➢ no referent < unfamiliar referents < familiar referents
Study7: Ong & Chan (2019)	➢ 50 bilingual English–Chinese young (mean age: 22.3) and old (mean age: 66.8) adults ➢ 24 disyllabic pseudowords were learned in different conditions: unknown vs. known (objects and names); auditory-visual learning (word-referent pair) ➢ recognition task (six alternative forced-choice tasks); immediate vs. delayed testing	Correct choices: unknown faces < known faces; unknown objects ≈ known objects

It cannot be denied that L2 word learning and testing paradigms could significantly affect L2 learning output. Hence, we wanted to review the studies in Table 1 in detail to determine what we can learn from each study.

Study 1: Comesaña et al. (2012) tested 48 native Portuguese children (mean age = 10.87) learning Basque as an L2 without previous L2 knowledge. The study examined the semantic interference effects of two learning methods (picture-based method vs. word-based method) and two types of words (cognates vs. non-cognates). Forty-two Basque words (21 words for each type of word) were selected to be learned, and each L2 word was mapped to three types of L1 word such as a correct translation (i.e., *zeru*/sky-*céu*/sky), a semantically related word (i.e., *zeru*/sky-*azul*/blue), and an unrelated word (i.e., *zeru*/sky-*marca*/mark). In the learning phase, 24 participants in each method group individually learned L2 words in four word lists in a quiet room, and they learned L2 words paired with the equivalent L1 translations or paired with the corresponding pictures. Participants had 9 minutes to memorize each word list (words were seen in written form or as picture referents) while an experimenter was reading L2 words aloud four times in the same order. They then had another 9 mins to revise all the L2 words, and the experimenter read all 42 L2 words again. Next, the participants answered a vocabulary test (i.e., the translation of L2 words) with the experimenter. The percentage of accuracy was measured by learning method group, and there was no significant difference between the two groups (89.46% for the word-based method and 93.25% for the picture-based method). In testing, there were three different randomly created groups of children following three experimental lists. The children performed a backward translation recognition task from the L2 to the L1 word. An L2 word was displayed on a computer screen for 250 ms after a fixation point was presented for 1000 ms. Then, an L1 word was shown on the computer screen until a response was made. In the absence of a response, the L1 word disappeared after 2500 ms. By pressing one of two buttons (yes vs. no), participants had to decide as quickly (reaction time measured) and accurately (accuracy measured) as they could whether the second word (L1) was a correct translation of the previously presented L2 word. The task was conducted in two different kinds of testing consisting of an immediate testing condition (10 mins after learning) and a delayed-testing condition (one week after learning). The following main results were found:

(a) For reaction times, participants were faster in the delayed test than in the immediate test; the responses in the picture-based method group were slower than in the word-based method group.

(b) For the error analysis, children made more errors in related pairs (vs. in unrelated ones), in the picture-based method (vs. word-based method), in the delayed testing (vs. immediate testing), in non-cognate unrelated pairs (vs. cognate related pairs).

From the learning situation, we see the effects of the word-based method (compared to the picture-based method) in L2 vocabulary learning because participants in the word-based method group were not only faster in the translation recognition task but also had fewer errors than those in the picture-based method group. However, it was necessary to examine whether the results would be replicated in the other L2 learning design.

Study 2: Emirmustafaoğlu and Gökmen (2015) also investigated the effectiveness of two instruction methods of L2 vocabulary learning (word-based instruction vs. picture-based instruction). Seventy-five L1 Turkish students in the seventh grade, with L2 English at elementary level, learned 20 unknown L2 words referring to concrete objects. They were divided into two learning method groups. In each group, they learned during two vocabulary sessions in which 10 L2 written words were presented in PowerPoint form in each session together with corresponding L1 written words or object pictures. L2 spoken words were read aloud by an experimenter for each learning trial. After each learning session, participants performed an immediate test, and they were asked to write down L2 equivalent words after seeing L1 written words or object pictures. One week after, they took the first delayed test with the same task as in the immediate test. One month later, they took the second delayed test. In the test, 10 pictures and 10 L1 written words were given to each group, and participants were expected to provide L2 written words. Data from 31 students in the L1 word instruction group and 29 students in the picture instruction group were included in the analysis using SPSS 15.0 to see whether there were significant differences in the correct answers between the two L2 learning method groups in the various tests. The results indicated that the participants in the picture-based instruction group performed tasks significantly better than the participants in the word-based instruction group in the first session, in the immediate test, and in the first delayed test. In the second delayed test, the difference was not statistically significant when participants were tested with both congruent and incongruent items compared to their instruction method group. However, in the second delayed test, the word-based instruction group provided L2 written words significantly better than the picture-based instruction group did when they saw the written L1 word. Meanwhile, no significant difference was found when both groups saw pictures.

In Study 2, participants with previous L2 knowledge learned some novel L2 words in two different visual learning conditions. With L2 written words as the test task, participants benefited from the picture-based instruction much more than from the word-based introduction.

In Study 1 and Study 2, subjects learned an L2 word through two kinds of learning methods that we summarized in Figure 1. Subjects learned L2 words in visual contexts combining both linguistic (i.e., written words) and nonlinguistic (i.e., spoken words and object pictures) cues. Study 3 was reviewed as a comparison of different L2 vocabulary learning contexts, namely auditory context, visual context, and auditory-visual context.

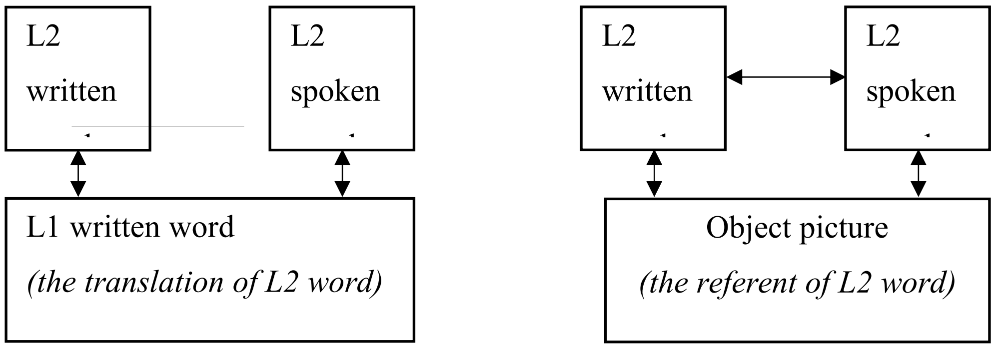


Figure 1: The word-based instruction and the picture-based instruction in L2 word learning

Study 3: Nassaji (2012) examined three L1 Farsi adult groups learning nonsense words in L2 English in three different associated word-referent conditions: (i) auditory only (a spoken word paired with a referent, n = 26); (ii) visual only (a written word paired with a referent, n = 28); (iii) dual modality (a spoken and written word paired with a referent, n = 25). Seventy-nine university students from 19 to 24 years old enrolled in English classes as a foreign language participated in the experiment. They learned 24 unknown labels as names (new words) for specific known objects (e.g., object-*camelus* and name-*TEV*). Three experimental blocks (eight word-referent pairs per block) were presented in the learning phase. After each leaning block, subjects were tested with recall and recognition tests for each half of the items learned. In the recognition test, for condition (i), an object picture was displayed on the computer screen when subjects heard four spoken words. An object picture appeared together with four written names on the screen for condition (ii). In condition (iii), subjects viewed an object picture while simultaneously listening to the spoken word and seeing the written word. Subjects had to indicate which name referred correctly to the object picture (e.g., object-*camelus*) by circling a, b, c, or d (e.g., *a. LEB; b. SOJ; c. TEV; d. KAG*) on their answer sheet. For the recall test, subjects saw a picture on the computer screen, and they were asked to recall the name by saying it into a microphone in condition (i), writing it on their answer sheet in condition (ii), or performing both tasks in condition (iii).

The results (see Table 2) showed that the mean scores for both the recognition and recall tests in the auditory-visual condition (7.04 and 4.88) were significantly higher than the mean scores in either the auditory condition (5.42 and 4.03) or the visual condition (5.25 and 3.82). Moreover, multiple comparisons among learning conditions using a post hoc Scheffé test indicated no significant difference between the auditory and visual conditions, but significant differences ($p < 0.5$) between single modality (auditory or visual) and dual modality (auditory and visual). Also, a significant main effect of measurement types was found. Participants in all three conditions were much more accurate at recognizing than recalling words ($p < 0.0001$). However, the study did not control how individual differences can influence differences in L2 learning success.

Table 2: Means and Standard Deviations of immediate recognition and recall (Nassaji, 2012; p. 47)

Measures	Conditions	m	SD	n
Recognition test (0–12)	Auditory	5.42	1.39	26
	Visual	5.25	1.35	28
	Auditory/Visual	7.04	1.36	25
	Total	5.87	1.57	79
Recall test (0–12)	Auditory	4.03	1.14	26
	Visual	3.82	1.24	28
	Auditory/Visual	4.88	1.05	25
	Total	4.22	1.22	79

Study 3 examined the effects of modality when adults learned L2 words under different learning conditions. It found that subjects learned L2 vocabulary most successfully in a thoroughly mixed combination of the written L2 word, the spoken L2 word, and the referent. Although no significant difference between the auditory condition and the visual condition was found, we noticed that subjects learned L2 words better in the auditory condition (a spoken word paired with a referent) than in the visual condition (a written word paired with a referent). In Study 3, the crucial role of referents in L2 vocabulary learning was also validated in all three learning conditions.

From Study 4 to Study 7, we reviewed how different features of referents in L2 word-referent learning paradigms influenced subjects' learning outcomes.

Study 4: Wang et al. (2017) examined an L1–L2 word-learning paradigm combined with cross-modal working memory binding. The focus was on the word-learning task, together with the retention performance only. Twenty-four Mandarin words were learned as L2 words by 71 L1 English speakers in a familiar-object condition (known-object and L2 word) and a novel-object condition (unknown object and L2 word) in separate blocks. In each condition, participants were first presented with 12 spoken word-object pairs in 12 sequential trials (each trial contained an object picture and an audio sound) before learning blocks. In each learning trial, participants heard a spoken L2 word and saw L2 object pictures immediately after clicking on the fixation cross on the computer screen. Their task was to identify the target object by mouse clicking. They got feedback on the next screen with the correct object picture displayed before moving on to the next item. They learned 12 items within each learning block, and the display location of object pictures changed randomly across trials. Participants had the opportunity to learn the same 12 word-object pairs in four other learning blocks. The proportion of correct choices was measured to evaluate immediate learning outcomes for each learning condition. Participants also performed two delayed recognition retention tests on Day 1 and Day 2. In both retention tests, 84 spoken word-object pairs were presented, and participants were asked to choose one of four responses (e.g., *intact, rearranged, new, or single*) to describe each pair precisely. Results for learning and retention performance showed that participants were successful in learning L2 words and retaining the L2 target word-object pairs for both learning conditions. In detail, the percentage of correct responses gradually increased from the first to the fifth learning block, for instance, 21%, 36%, 46%, 62%, and 67% in the novel-object condition, and 34%, 47%, 60%, 69%, and 80% in the familiar-object condition. Moreover, participants accurately recognized 75% to 87% of all intact pairs across two delayed tests.

In Study 4, participants successfully learned and recognized L2 words in auditory-visual contexts. Also, there was no significant difference between familiar-object conditions and novel-object conditions when participants learned a spoken L2 word together with its referent. Another subsequent study (Study 5) focused on examining the benefits of not only referent familiarity (familiar vs. unfamiliar) but also phonological familiarity (familiar vs. unfamiliar) in L2 word learning.

Study 5: Kaushanskaya et al. (2013) carried out a study to investigate whether phonological familiarity can benefit novel word-referent learning. Eighty-one L1 English adults with a knowledge of Spanish learned phonologically familiar novel words (+P: using English sounds) or phonologically unfamiliar novel words (-P: non-English and non-Spanish sounds) together with familiar (+R) or unfamiliar referents (-R). Forty participants in a group learned +P, and 41 others learned –P. Participants had two different learning sessions. In both groups, they were taught 24 new words paired with familiar referents (pictures of animals), and 24 other words paired with unfamiliar referents (pictures of aliens). In each learning trial, participants heard the spoken word twice via headphones and inspected the referent displayed on the computer screen for 6 seconds. In testing, they performed a forced-choice recognition task. Four referents (object pictures) were presented together, and participants were asked to choose one of them referring to the spoken new word. Accuracy rates were above chance performance level (0.25) for four learning conditions ($p < .0001$). Further analysis in learning familiar referents as novel words showed that participants recognized more accurately phonologically familiar novel words than phonologically unfamiliar novel words. The effect was not only significant by subject ($t1(80) = 2.12$, $p < .05$) but also by item ($t2(47) = 2.77$, $p < .05$). However, no significant effect of phonological familiarity ($p > .05$) was found when participants learned unfamiliar referents as novel words.

The study documented the effect of phonological familiarity on learning to recognize familiar referents only when designated by novel words. It also suggested that L1 adults can more successfully learn L2 words with familiar referents despite the phonologically unfamiliar sounds compared to learning L2 words with unfamiliar referents.

The next research study applied an audio-visual context to L2 word learning to investigate sleep effects (i.e., overnight consolidation) when subjects learned both novel L1 modified and L2 words with three different types of referents: familiar referents, unfamiliar referents, and no referents. The researchers also wanted to determine the differences in L2 word-learning success among the three learning conditions.

In Study 6, Havas et al. (2017) examined semantic and phonological effects on learning L1 and L2 spoken words in an initial acquisition (–sleep) and with overnight consolidation (+sleep). L1 Spanish participants without L2 Hungarian knowledge learned new words in spoken forms in L1 Spanish and L2 Hungarian. Sixty-eight L1 speakers, aged from 18 to 36, were divided into four experimental groups including L1 –sleep, L1 +sleep, L2 –sleep, and L2 +sleep (i.e., –sleep groups were trained in the morning and tested in the evening, while +sleep groups were trained in the evening and tested on the morning of the following day). They learned 144 novel spoken words of L1 or L2 paired with familiar pictures, unfamiliar pictures, or no picture (nothing to see). L1 words were created by changing an L1 original word (e.g., *casco* => *cosco*), while L2 words were Hungarian words (e.g., *golyó*). Importantly, L1 speakers did not have any knowledge of half of the phonemes in the L2 words because of phonological differences between the two languages. Participants in the training phase learned spoken word-picture pairs in three different learning conditions (familiar pictures, unfamiliar pictures, or no picture). Each word-picture pair was presented five times in a total of four training runs, and learners were instructed to learn as many of them as possible.

Participants took (a) an auditory recognition memory test after each training run. They were asked to judge whether they learned each item of 18 learned items and 18 unlearned items. For delayed testing, 12 hours after training, participants had to perform three other tasks:

(b) an additional auditory recognition memory test: Participants performed an old-new judgment by pressing a button for 72 trained and 72 new spoken words.

(c) a four-alternative forced-choice word-choice matching task: They were required to choose one of four familiar or unfamiliar pictures that can be paired with a spoken word they heard.

(d) a semantic priming task: participants were tested with 24 primes (24 L1 or L2 spoken words). After a 500-ms fixation cross, the prime auditory stimulus was presented before a written target word was visually displayed on a computer screen. Each prime was presented four different times, once with a Spanish translation related to the prime, once with a real Spanish word unrelated to the prime, and twice with Spanish pseudowords). A lexical decision task was completed by participants to compare reaction times between related and unrelated prime trials.

The findings of the study are listed as follows. For the auditory recognition memory test in training (a), no significant time-of-day effects on initial learning were found ($F(1,61) = 0.02$, p

= 0.885). However, a main effect of the picture condition was shown, which means participants learned novel spoken words much better in familiar picture conditions than in unfamiliar picture conditions or no picture condition ($F(2,122) = 15.55$, $p = 0.0001$). They also found a significant effect of training run ($F(3,183) = 25.71$, $p = 0.0001$) in which participants improved their recognition throughout the training. Also, the main effect of language ($F(1,61) = 24.38$, $p = 0.0001$) indicated that it was more difficult for L1 learners to learn novel words in a phonologically different L2. In the second auditory recognition memory test (b), the authors investigated all three main effects of three factors (i.e., picture, sleep, and language). Participants recognized words much better in the familiar pictures compared to two other conditions ($F(2,120) = 22.25$, $p = 0.0001$)). They were also more successful at recognizing L1 words than L2 words ($F(1,61) = 6.06$, $p = 0.017$), and sleep benefited their recognition ($F(1,61) = 4.58$, $p = 0.036$). Post hoc analyses demonstrated the beneficial effect of sleep in the group learning only L2, and the effect of language was only present for the L2 –sleep group. The results of the third task (c) indicated a significant main effect of the picture; significant two-way language by picture and language by sleep interactions. Post hoc analyses showed the main sleep effect for L2 learning groups, but not for L1 learning groups. Also, a main effect of language was found for +sleep groups (i.e., L2 +sleep performed better than L1 +sleep). For the picture by language interaction, the main benefit of familiar pictures (compared with unfamiliar pictures) was seen in the L1 learning groups, and not in the L2 learning groups. In the last task, participants responded significantly more quickly to a Spanish translation related to the prime than to a real Spanish word unrelated to the prime. However, no main priming effects were found when the semantic priming task results were examined with trained item primes.

To summarize the study, three main results were found: (i) L2 spoken words could be recognized better because of sleep effects. (ii) Familiar objects/pictures could enhance recognition memory for both L1 and L2 spoken words in immediate testing and delayed testing. (iii) Due to the phonological familiarity, L1 spoken words were learned and remembered much better than L2 spoken words in the –sleep condition. In contrast, L1 spoken words were learned and recognized as well as L2 spoken words in the +sleep condition because of overnight phonological consolidation.

This study suggested that people can successfully learn a novel L2 word (with different phonological features from their L1) when L2 spoken words were paired with familiar (vs.

unfamiliar) objects. Learners can learn L2 phonological features better if there is a consolidation phase.

Study 7 tested whether referent features can affect L2 learning success; referents were not only objects referring to nouns but also faces referring to names.

Study 7: Ong and Chan (2019) investigated a word-learning paradigm to determine the influence of referent features on word-referent mapping. Fifty bilingual (English–Chinese) young (M age = 22.32) and older adults (M age = 66.80) learned 24 word-referent pairs (i.e., names of objects or nicknames of people) with two different referent types (face vs. object) and their familiarity (known vs. unknown). These words were disyllabic pseudowords (e.g., *Zoohee, Kepfi, Pigga,* and *Famdae*) spoken in a robotic voice, which makes their sounds different from both English and Chinese. They were learned in four different referent conditions: known-face, unknown-face, known-object, and unknown-object (see Figure 2). In each learning trial, participants saw a referent presented on the computer for 4500 ms before hearing its name via headphones in 2000 ms. Each referent-name pair was presented twice during the learning phase.

	Face	Object
Unknown		
Known		

Figure 2: Four learning conditions in Study 7 (adapted from Ong and Chan, 2019)

After learning, participants were tested (immediate vs. delayed) with a recognition task (i.e., six alternative forced-choice tasks). In every testing trial, participants had to choose one of six referents on the computer screen corresponding to a word they had heard before. The complete procedure (approximately 1 hour for young adults and 1.5 hours for older adults) included learning, immediate test, cognitive test (nonverbal intelligence test and English receptive vocabulary test), delayed test, and post-task questionnaire.

The results for the name learning tasks in the immediate and delayed tests showed that both young and old learners successfully learned the names of referents in all learning conditions because their performance on all the referent conditions was significantly above chance ($p < .01$). More importantly, known referents for faces were learned much better than unknown items ($p < .001$), and there were no significant differences between known and unknown referents for objects ($p = .881$) when they were learned. There were other findings of interest: the main effect of age (young adults performed significantly better than older adults in both tests); the main effect of the testing session (performance was better in the immediate test than in the delayed test); and some two-way interactions (age × testing session, age × familiarity, and familiarity × type) were observed. In conclusion, the findings of Study 7 confirmed that L2 learners successfully named known and unknown referents (i.e., faces or objects), and the suggested characteristics of referents influenced word-learning outcomes and learners' memory.

All studies in Section 2.2.1 focused on L2 word learning in different visual contexts. Participants learned L2 words successfully (i.e., mostly in immediate and delayed testing) in various visual learning contexts, especially in auditory-visual learning conditions with familiar or known referents. However, it is not known whether L2 beginners would learn L2 phrasal vocabulary in a visual context. In the next section (2.2.2), we will review studies of learning and processing L2 phrasal vocabulary because we want to know how L2 phrases have been learned, tested, or processed in different contexts.

2.2.2 Second Language Phrasal Vocabulary Learning and Processing

A high level of L2 vocabulary learning is required to learn L2 phrasal vocabulary or L2 collocation. L2 learners should learn L2 phrases (i.e., collocations) as a higher vocabulary level (not grammar) because of the many benefits of L2 phrases in L2 usage. We did not find any study of L2 collocation learning in visual contexts. Most studies focused on examining L2 learners' knowledge of L2 collocations when they had been learning L2 for a long time. Study 8 and Study 9 reviewed how L2 speakers with different L2 proficiency levels processed L2 collocations. Hence, this study could offer some suggestions for L2 phrasal learning for L2 beginners.

Study 8: Wolter and Gyllstad (2011) examined the influence of L1 intra-lexical knowledge on forming L2 collocations in native and non-native speakers by using the Lexical Decision Task (LDT). L1 Swedish participants (n = 30, ages 18–61) learning L2 English and native English (n = 35, 19–65 years) participated in the study. Non-native speakers of English reported their self-report proficiency scores of L2 with means for speaking, 6.8: listening, 7.4; reading 6.8; and writing, 6.1compared to 10 for near native-like. Two participant groups were tested with three lists (33 items per list) of English collocations consisting of an L1–L2 collocation list (*ge ett svar*$_{(Swedish)}$ ⇔ *give an answer*$_{(English)}$), an L2 collocation list only (*betala ett besök*$_{(Swedish)}$ # *pay a visit*$_{(English)}$), and an unrelated list. Also, 121 filler items were added as an additional list. After seeing the prime (a verb: *pay*) and the target (a noun: *VISIT*), participants were asked to respond as fast as possible by pressing the YES or NO key to indicate whether the string (verb-noun: *pay-VISIT*) is a real word (see Figure 3).

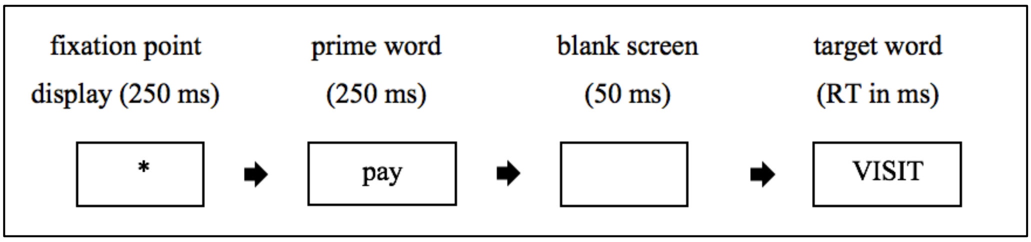


Figure. 3. Sequence of presentation for items on Lexical Decision Task (adapted from Wolter and Gyllstad, 2011).

The goal of the study was to investigate the possible activation of connections in participants' mental lexica. Three hypotheses were tested, and the following results were expected:

- A significant priming effect of L1–L2 collocation compared to unrelated collocation for L1 Swedish speakers
- Less significant priming effect of L2 collocation only than L1–L2 collocation for L1 Swedish speakers
- A significant priming effect of both L1–L2 collocation and L2 only for native English speakers, and no significant difference between them

Results showed a significant effect in reaction times both for English (F: $p < .001$, F2: $p = .001$) and Swedish (F: $p < .001$, F2: $p = .002$) speakers, but no significant differences in the error rates across the three lists were found for both groups. Pairwise comparisons for the reaction times indicated significant differences for the English group between the L1–L2 list and an unrelated list (F1: $p < .001$, F2: $p = .003$), also between the L2 only list and the irrelevant list (F1: $p < .001$, F2: $p = .006$), but not between L1–L2 list and L2 only list (F1: $p = 1.0$, F2: $p = 1.0$). Consequently, in the native English group, there was a priming effect for both collocational conditions; however, there was no significant difference between them as expected. For the non-native group (Swedish speakers), there was only the comparison between the L1–L2 list, and the unrelated list showed significance for both the subject analysis and item analysis (F1: $p < .001$, F2: $p = .001$).

The study focused on investigating the influence of L1 on the development of L2 collocational knowledge for speakers with high L2 proficiency, and a clear significant difference between the L1–L2 condition and the unrelated condition for L2 speakers was found. This statement suggested that L2 verb-noun phrases or collocations can have a more central position in L2 vocabulary learning. The best learning situation for L2 learners is that L1 collocation knowledge is fully transferable to L2 (i.e., L1–L2 correspondence).

Another study (Study 9) also examined how different native and non-native speakers processed English verb-noun collocations.

Study 9: Yamashita and Jiang (2010) investigated how three different participant groups (English speakers, n = 20; Japanese English as a second language speakers, n = 24; and Japanese English as foreign language learners, n = 23) processed English verb-noun collocations. All participants were tested with 24 congruent collocations (L1⇔L2: *make lunch, heavy stone*), and 24 incongruent collocations (*kill time*$_{(L1\text{-}English)}$ = *break time*$_{(L2\text{-}Japanse)}$, *slow learner*$_{(L1\text{-}English)}$ = *a person who learns slowly*$_{(L2\text{-}Japanese)}$), and 48 implausible word combinations (*abstract fruits, begin a bed*). The participants' task was to press a YES or NO key (i.e., Is it a collocation?) as quickly and as accurately as they could when they saw each item displayed on a computer screen after a fixation of 500 ms (see Figure 4). The task was called the "Phrasal Decision Task."

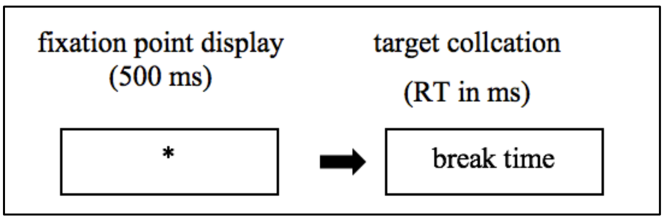


Figure. 4. Sequence of presentation for items (adapted from Yamashite and Jiang 2010).

The results for reaction times and error rates indicated no significant differences between congruent and incongruent conditions for native speakers (differences: 6 ms in reaction times, and 2.6% in error rates). However, for Japanese groups, the differences were statistically significant in error rates in that participants made more errors in incongruent collocation conditions than in congruent conditions (ESL: 11.48%, and EFL: 24.24%). For reaction times, a significant difference was found in the EFL group only insofar as participants responded more slowly ($m = 55$ ms) in incongruent conditions compared to congruent ones. Also, for the congruency effect (congruent vs. incongruent), no significant differences were found in the error rate and reaction times for native English and ESL speakers both for participant analysis ($p > .05$) and item analysis ($p > .05$). However, the researchers found a congruency effect in the EFL group because these participants made more errors in incongruent conditions than in congruent conditions both in participant analysis ($p < .001$) and in item analysis ($p < .001$).

For reaction times in participant analysis, participants responded significantly more slowly in incongruent collocations as compared to congruent ones ($p < .05$). There was a marginally significant difference in item analysis ($p = .06$), which suggested that L2 collocation might be processed via L1 mediation at an early L2 learning stage.

An analysis of group differences under each collocation condition was included. For the congruent condition, the EFL learners made significantly more errors than native speakers and ESL users ($p < .05$), and native speakers were faster than both ESL and EFL users. For the incongruent condition, the error rates and reaction times of the three groups were significantly different, except for the item analysis of error rates for ESL users and native speakers. The differences in English proficiency among groups might explain the above results, and they indicate the difficulties of learning and acquiring L2 collocations for both non-native groups.

We mention the study because it examined L2 speakers with different L2 proficiency levels processing several types of L2 collocations/phrases. L2 learners with lower L2 proficiency processed congruent collocations (existing in L1 and L2) more successfully than incongruent ones (differences between L1 and L2 or only existing in L2). The results in L2 collocation processing also suggested that L2 beginners can quickly learn L2 basic phrases or collocations corresponding to L1.

Overall, L2 collocation (i.e., L2 phrasal vocabulary) processing has been considered in some studies mentioned in section 2.2.2 when learners had intermediate or advanced L2 knowledge. However, we found no empirical research on L2 phrasal vocabulary learning for L2 beginners.

We can summarize the section (2.2) in the following significant points:

- L2 intentional vocabulary learning is essential for L2 beginners, and learning becomes much more accessible when using a word-referent learning paradigm, especially with familiar referents.
- After the learning phase, many diverse tasks in both the immediate and delayed tests were used to see whether learners could maintain or improve their learning success or which learning conditions could support learners better than others.
- In previous L2 learning studies, participants at beginner level learned L2 words separately or following specific categories, not linked with a higher level of vocabulary units (i.e., noun phrases or verb-noun phrases). In L2 processing studies, advanced L2 learners processed L1–L2 congruent collocations most successfully.

Therefore, we were motivated to research whether L2 beginners can successfully learn L2 phrasal vocabulary items in a visual context (phrasal-referent conditions) with different kinds of language mappings (i.e., L1–L2 congruent vs. L1–L2 incongruent phrases).

In the next section 3, we reviewed how language (both L1 and L2) was learned, processed, and comprehended visually to understand more about the critical role of visual context. More specifically, we wanted to find more evidence to research the question, “How do L2 beginners learn L2 phrasal vocabulary with or without nonlinguistic visual cues?” when language comprehenders and language processors experience many benefits in nonlinguistic visual cues.

3. Depicted Objects and Depicted Actions: From Language Processing and Language Comprehension to Language Learning

Visual context is an indispensable factor in language processing and comprehension. Knoeferle and Guerra (2012), in a review, summarized studies indicating that various kinds of visual context could immediately and incrementally influence language comprehension. Language listeners' visual attention and online language understanding can be directly affected by nonlinguistic information from the immediate context (e.g., *objects* and *events*), from a recently inspected visual context, from the speaker (e.g., *eye-gaze* and *gesture*). These can happen through referential (e.g., *gardens-gardening*), simple lexico-semantic associative (e.g., *drink-wine*), or functional relationships (e.g., *subject-object*). The primary role of a nonlinguistic visual context has been emphasized in visual-world eye-tracking studies. Comprehenders can speedily and efficiently use the linguistic input linked to nonlinguistic visual information about objects and depicted action events in the visual world. Thus, the real-time processing of linguistic input, such as clarifying a sentence in an ambiguity, would be facilitated. Study 10 could provide evidence of language processing in a visual context containing both linguistic and nonlinguistic information.

In Study 10, Münster (2016), in her eye-tracking studies, investigated the effects of emotional facial expressions (happy vs. sad) and depicted action events (present vs. absent) on situated language processing in a visual context (i.e., seeing an emotional face before hearing a spoken sentence and seeing an event picture presented simultaneously) at various life stages (German children vs. German younger adults vs. German older adults). In Experiment 3 for the young adult group (18–30 years), Experiment 5 for the children's group (4–5 years), and Experiment 6 for the older adult group (60–80 years), participants experienced the same experimental procedure (see Figure 5). First, they saw a naturally emotional female face (happy vs. sad or smile vs. no smile) presented for 5500 ms as a prime. Then, participants inspected an event picture (with vs. without action from potential agents to the "patient") for 2000 ms before spoken sentence onset (e.g., *Den Marienkäfer kitzelt vergnügt der Kater*). Finally, to finish the experimental trial, they had to answer a comprehension question (e.g., *Wer wird hier gekitzelt?*).

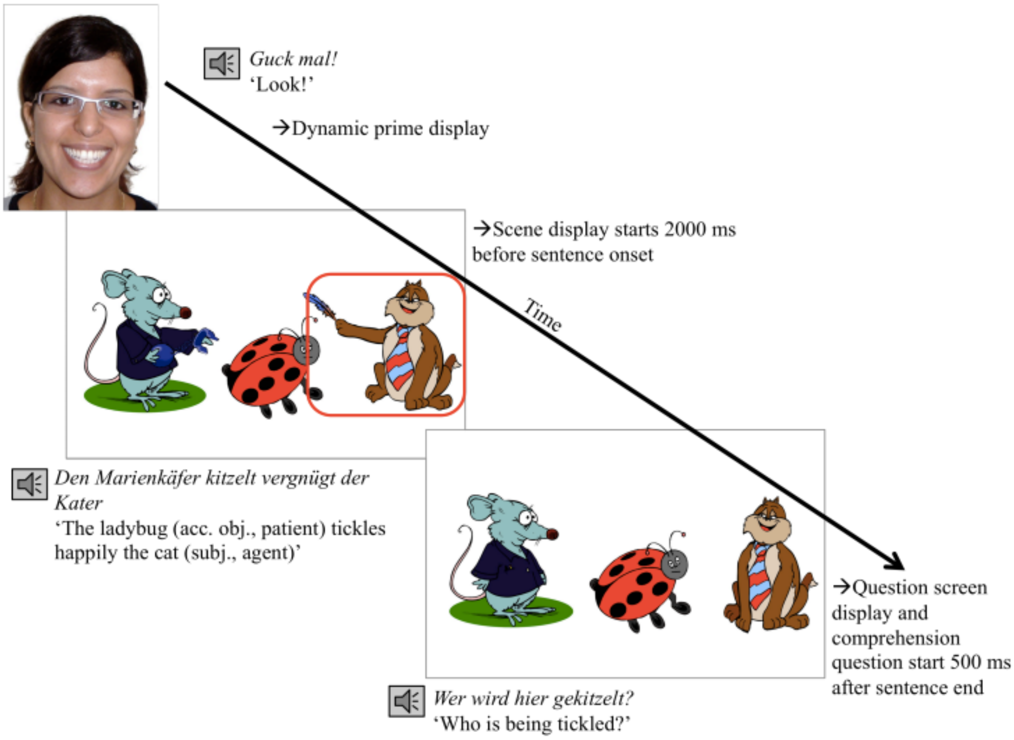


Figure 5. The procedure of the eye-tracking experiment (Münster, 2016)

There were four experimental conditions: (a) happy face and present action; (b) happy face and absent action; (c) sad face and present action; (d) sad face and absent action. In the study, we only paid attention to the effects of direct cues (depicted action: present vs. absent) on language processing in a visual context, and whether these effects were different among children, young adults, and older adults.

The study examined the action effect, the positive prime effect, and the cumulative visual context effects (i.e., direct cue and indirect cue interaction). Münster found robust action effects across all age groups. The main result showed that in language processing, participants were more strongly affected by direct cues (i.e., depicted actions) than indirect cues (i.e., emotional facial expressions). In Table 3, we summarized the similarities and differences of depicted action effects among three different age groups both in online language processing and in responding to the offline comprehension questions.

Table 3: Depicted action (present vs. absent) effects on language processing across various life stages (Münster, 2016)	
Similar	All three age groups benefited from present action (vs. absent one) to determine exactly "who does what to whom."
Different	The action effect - Mean log ratio of looks for the main effect of action in the Verb-Adverb region + Younger adults (present > absent) + Children and older adults (present ≈ absent) - Accuracy (offline) + Children (much higher accuracy in present action conditions than absent action) + Younger and older adults (a little higher accuracy in present action conditions than absent action)

We learned from Study 10 that depicted actions were very helpful for children and adults when they processed sentences and decided "who does what to whom." Therefore, we were inspired to think more about how visual context (e.g., present vs. absent actions/events) effects would be shown in L2 language learning, for instance, when people learn L2 verb-noun phrases (e.g., *drink-tea)* with and without event picture referents. Another question is whether we can see age-related differences in the learning situation.

In the area of language learning, Knoeferle (2015), in a review study, highlighted that children, for instance, can learn about object-based regularities (nouns), actions (verbs), and events (verb-noun phrases). They can map nouns rapidly onto objects, and actions onto verbs after extracting them from supportive streams of auditory and visual events. Importantly, learners can learn nouns most successfully in visual contexts when the visual objects and language input referenced have a regular pattern. Furthermore, cross-situational word-learning experiments have provided evidence of visual context (i.e., object referents paired with spoken words) effects on language learning (Studies 11 and 12).

In Study 11, Smith and Yu (2008) examined how infants can learn under uncertainty via cross-situational statistics (across word-referent pairs and co-occurrences across time). The most straightforward case of cross-situational learning is that a learner encounters two trials comprising two unknown spoken words (e.g., Trial 1: *bosa* and *gasser;* Trial 2: *manu* and *gasser*) and two images corresponding to these object names per trial. On the first trial, learners cannot identify which object is *bosa* and which one is *gasser* when they listen to these spoken words and look at the object pictures. However, for the second trial, they might precisely map the spoken word *gasser* to the visual referent (e.g., *a star shape*) because *gasser* and the star shape appeared in both Trials 1 and 2 while the other object and its name were changed. The results showed that infants learned word-referent mappings rapidly via cross-situational learning, as evidenced by reliably more looks at the target when it was named (e.g., at the star vs. another object when hearing *gasser*). A previous study by Yu and Smith (2007) experimented on how adults learned uncommon words (e.g., *rasp, facial sauna,* and *canister*) via cross-situational learning. Subjects learned words under uncertainty in different learning conditions such as 2 × 2 (two words and two referents), 3 × 3 (three words and three referents), and 4 × 4 (four words and four referents). Eighteen word-referent pairs were trained per condition. After training in each condition, the participants were asked to indicate which one of four pictures presented was named by a word they were hearing. Results showed that adults learned words paired with their referents much better than would be expected by chance in all three conditions; specifically, they could learn more than 16, more than 13, and almost 10 of the 18 pairs in the 2 × 2, 3 × 3, and 4 × 4 conditions, respectively. Many other studies have shown that people in different groups can learn new words successfully in various scenarios of cross-situational word learning (Chiders & Paik, 2008; Kachergis, Yu & Shiffrin, 2010; MacDonald, 2017; Monaghan & Mattock, 2009; Vouloumanos, 2008; Vouloumanos & Werker, 2009). These studies concentrated on investigating the nature of word-learning mechanisms, underpinned by object-word co-occurrence frequencies across situations. Researchers found that infants and adults benefit from the word-learning mechanisms in many experimental designs; their word-learning scores were often higher than expected by chance. In these studies, we recognized significant effects of nonlinguistic visual cues (i.e., object referents) as objects were beneficial for learners to extract correct pairs (i.e., spoken words and objects) among many learning trials. Subsequently, they could remember specific pairs better during learning.

The following study (Study 12) investigated the word-learning mechanism at a more challenging level. Leaners learned verbs and nouns and then processed simple L2 Subject-Verb-Object (SVO) sentences.

In Study 12, Koehne and Crocker (2015) carried out three experiments to estimate word learning based on cross-situational and sentence-level constraints. They evaluated whether linguistic contexts (i.e., a verb's argument selection restrictions impose constraints on upcoming direct object nouns) could, across trials, benefit word learning. They further considered the interaction of cross-situational and sentence-level constraints on word learning. In Experiment 1, cross-situational and sentential constraints were complementary; in Experiment 2, they were redundant; in Experiment 3, they were conflicting. Native German-speaking adults learned a set of verbs, and then novel nouns (from a miniature semi-natural language) in linguistic contexts situated in visual environments, or they learned in strictly linguistic contexts. Post-learning, the authors examined participants' noun knowledge (via accuracy and eye movements). The results supported the view that learners draw on multiple mechanisms (Experiment 1) and that both cross-situational and verb constraints influenced learning.

Moreover, the results indicated that cross-situational word learning was not employed when the sentence-level restriction was completely disambiguating. By contrast, cross-situational word learning was employed when verb constraint resulted in insufficient information for classifying a referent (Experiments 1 and 2). The study also showed that L2 beginners could learn not only words (nouns) but also verb-noun phrases and simple sentences in visual contexts. That suggested a crucial point for us to pursue in our research concerning a visual context for training L2 beginners in verb-noun phrases.

In summary, language comprehenders could effectively use nonlinguistic visual cues (i.e., depicted objects and actions) for their online language understanding. Also, adult beginners could acquire second language vocabulary in specific situated learning conditions. They were able to learn not only single words but also phrases and simple sentences via different learning mechanisms. Visual contexts play a significant role in language learning because they permit learners to relate language to objects or actions. This state of the art motivated us to conduct experiments on how successfully adults (L1 German) learn L2 (Vietnamese) verb-noun phrases with/without supporting nonlinguistic visual cues (i.e., events/actions). We wanted to examine the effects of visual contexts (i.e., present vs. absent action) to facilitate L2 phrasal vocabulary learning for adults with no prior knowledge of the L2 (Vietnamese).

In Section 4, we examine in more detail the indispensable role of L1 in L2 vocabulary learning or processing in prior studies because understanding more about language transfer (L1–L2 transfer) can explain why people can successfully learn some specific L2 vocabulary items but not others.

4. Language Transfer in Second Language Learning

When L1 speakers acquire a second language, their language knowledge in the first language is a principal factor affecting the process. It acts as “a major factor in second language learning” (Ellis, 1990, p. 297). Learning an L2 means that learners must form a new, at least partially separate habit set for the L2 (Lado, 1957), though not independent from their L1. Since the 1960s, the term “language transfer” has been applied in linguistics. It references influences from experience in an L1 when an L2 is learned (Gass, 1984; Sajavaara, 1987). Language transfer is understood as “the influence resulting from the similarities and differences between the target language and any other language that has been previously acquired” (Odlin, 1989, p. 12). Researchers commonly believed that the similarities between L1 and L2 could promote L2 learning (“positive transfer”), whereas the differences can lead to interference (“negative transfer”) and associated errors in L2 performance (Ellis, 1985; Gass & Selinker, 2001; Odlin, 1989). Contrastive analysis (CA) is described as a linguistic comparison of the structures of two or more languages to demonstrate their differences and similarities (Figure 6). Consequently, that suggested that when the L1 and L2 are very different, L2 learning might be relatively unsuccessful or stressful.

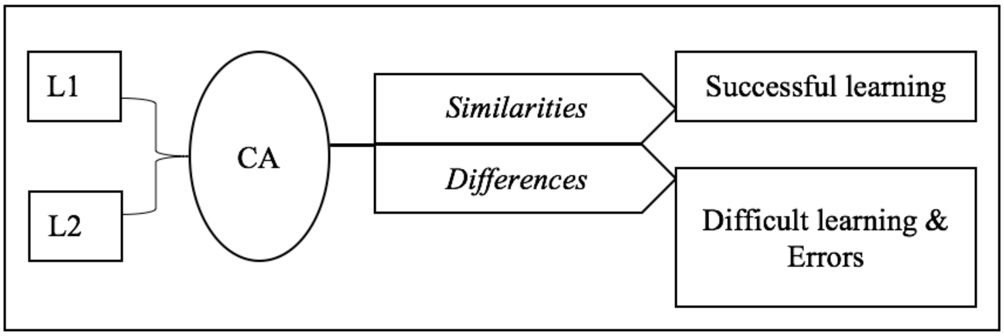


Figure 6. Contrastive analysis and L2 learning (adapted from Al-khresheh, 2016)

Many previous studies investigated the benefits and interferences of L1 transfer in L2 learning not only in many linguistic areas (phonology, lexicology, semantics, syntax, and pragmatics) but also regarding language skills (speaking, reading, writing, and translation). The transfer has also been examined for vocabulary acquisition, given its central role in L2 learning (Lewis, 1993), and we will further investigate lexical transfer.

4.1 Lexical Transfer: Definition and Types

Jarvis (2009, p. 99) described the lexical transfer as “the influence that a person’s knowledge of one language has on that person’s recognition, interpretation, processing, storage and production of words in another language.” There can be both positive and negative lexical transfer. If lexical similarities between L1 and L2 could support learners in L2 vocabulary learning and usage, that would be positive transfer. Conversely, negative lexical transfer could happen if applying L1 vocabulary knowledge leads to error in L2. In other words, lexical differences between L1 and L2 can prevent people from learning L2 vocabulary successfully. Ringbom asserts that “it is difficult to determine exactly the extent of positive influence, compared with negative influence, since the only tangible signs of cross-linguistic influence are negative ones, errors.” (1986, p. 160).

Specific types of lexical transfer have been categorized differently. From the viewpoint of error analysis, Ringbom (1987, 2001) believed that transfer has always been negative. The reason is that recognizing an L2 learner error has been easier than perceiving the positive effect of L1 when acquiring an L2. He considered *formal transfer* linked to morphonological errors and *semantic transfer* related to L2 word usage errors, especially in L2 word meanings.

> The ability to access a word in one’s mental lexicon (language activation), knowledge of how the word is pronounced and spelled in its various forms (morphophonology), and semantics: knowledge of the meanings of the word, the word combinations in which it conventionally occurs (collocations) and the word’s associations with other words and notions. Morphophonological errors are called *formal transfer* and include the use of a false cognate, lexical borrowings or lexical inventions, and *semantic transfer* can be characterized as the use of a target language word with a meaning that reflects the influence from the semantic of a corresponding word in another language. (Ringbom 2001, p. 64)

Jarvis and Pavlenko (2008) and Jarvis (2009) thought that lexical transfer occurs through two distinct processes in the mental lexicon. They differentiated between lexemic transfer and lemmatic transfer. Lexemic transfer (i.e., *formal transfer* in Ringbom, 1987) involves the phonological and the graphemic structure of a word (e.g., knowledge of spoken and written forms of the word *eat* such as *eat, eats, eating, ate,* and *eaten*). Lemmatic transfer refers to syntactic features (i.e., noun or verb) and semantic properties (i.e., synonyms and antonyms) of a word. Table 4 presents all sub-types from two main types (i.e., formal transfer and semantic transfer) of lexical transfer that we summarized from previous studies.

Table 4: Types of lexical transfer summarized from Ringbom (1987) and Javis (2009)

Types	Sub-types	Examples of negative lexical transfer/ errors in L2 usage
Formal transfer/ Lexemic transfer	False friends/Cognates (L1 speakers think L1–L2 share forms together with semantic similarity)	- "fabric" instead of "factory" - e.g., "At the time he works in a *fabric*" Ringbom (1987, p. 117) because L1 Swedish: *fabrik* ⇔ L2 English: *factory*
	Unintentional language switches (L1 speakers apply L1 word form in L2 context)	- "pigg" instead of "refreshed" - e.g., "I'm usually very *pigg* after the diet" Ringbom (1987, p. 117) because L1 Swedish: *pigg* ⇔ L2 English: *refreshed*
	Coinages (L1 speakers create non-existent L2 words)	- "*piggy*" instead of "*refreshed*" - e.g., "In the morning I was tired, and in the evening I was *piggy*" Ringbom (1987, p. 117) because L1 Swedish: *pigg* ⇔ L2 English: *refreshed*
Semantic transfer/ lemmatic transfer	Semantic extensions	- "*language*" instead of "*tongue*" - e.g., "He bit himself in the *language.*" (Ringbom 1987: p. 117) because L1 Finnish: *kiele* ⇔ L2 English: *language* and *tongue*
	Calques/loan translations (L1 speakers apply L1 combinations in L2 usage)	- "*fire sticks*" instead of "*matches*" (Ringbom 1987, p. 115) because L1 Finnish: *tulitikut* ⇔ L2 English: *matches,* but literally, *fire sticks*
	Collocational transfer (L1 speakers apply L1 collocational links in L2 usage)	- "*admit discount*" instead of "*allow discount*" (Hasselgren 1994, p. 251) because L1 Norwegian: *tillate* ⇔ L2 English: *admit* and *allow*

	Subcategorization transfer (L1 speakers select the wrong complement for a word)	- a noun phrase "*his mother*" after a verb instead of a prepositional phrase "*about his mother.*" - e.g., "He was *thinking his mother.*" (Javirs 2009, p. 117)

More recent research also divided lexical transfer into formal and semantic transfer (Bardel, 2015). Formal transfer includes borrowing, foreignizing, and spelling transfer, while semantic transfer includes lexeme matching, semantic extensions, and direct translations. In L2 vocabulary acquisition, according to Ringbom (1987), the stage of learning can affect L1–L2 lexical transfer:

> *Stage of learning*: The role of the L1 in L2 learning is most important at the early stages of learning and decreases as learning progresses. This is due to the fact that L2 knowledge of a beginning learner is so limited that he needs to believe that the L2 will be in many, or at least in some, respects work in a similar way to his L1. While intermediate and advanced learners will show a complex interaction of L1- and L2-influence, with the former decreasing as he reaches L2 proficiency. (Ringbom, 1987, p. 63)

Lexical transfer among L2 beginners, L2 immediate, and L2 advanced learners is not consistent. L1 lexical knowledge could strongly influence the L2 vocabulary learning process and the results at the early stage of L2 learning. According to many researchers (Bardel & Falk, 2007; Falk & Lindqvist, 2014; Navés, Miralpeix, & Celaya, 2005; Pfenninger & Singleton, 2016; Williams & Hammarberg, 1998), at lower levels of L2 proficiency, form-based transfer tends to dominate, whereas meaning-based transfer dominates when learners have high L2 ability.

Positive or negative lexical transfer depends on the degree of similarities and differences between L1–L2 words or phrases in specific learning situations. Whenever the lexical similarity between L1 and L2 has a facilitating lexical effect on learners' L2 usage, it is regarded as positive. By contrast, L2 learning errors resulting from subjects' L1 vocabulary usage are considered to be negative.

In the next section (4.2), we focused on specific studies on L1–L2 lexical transfer in language learning and language processing when there is a form or meaning correspondence or differentiation between the L1 and L2 words or phrases.

4.2. Lexical Transfer: First Language–Second Language Vocabulary Correspondence and Differentiation

Many researchers found that similarities and differences in word forms and word meanings play a significant role in L2 vocabulary learning success. Hence, the more formal and semantic the similarities between L1 and L2, the more learners will benefit from their L1 in learning L2 vocabulary (i.e., lexical transfer). "Lexical items which are cross-linguistically similar to L1-items already stored will be understood best of all by learners learning closely related languages." (Ringbom, 1987, p. 35).

The role of L1 in the formation of L2 lexical representations has been addressed via different models of L1–L2 lexical processing. In the following sub-sections, we reviewed some models applied in the early stage of L2 vocabulary learning and the later stage for learners with high L2 proficiency.

4.2.1 Lexical Transfer in Second Language Learning

To examine the role of L1, Stockwell et al. (1965a, 1965b), based on the results of cross-language comparison (i.e., contrast analysis theory), summarized key differences between languages and established a hierarchy of learning difficulty consisting of five categories/levels (see Table 5).

In Table 5, we can see that the differentiation between L1 and L2 has been the most challenging category in L2 learning. For instance, if a word has only one form in the L1 (English: *to know*), but it has two or more possible forms in the L2 (e.g., Italian*: (i) sapare* – 'to know,' e.g., a fact/to get knowledge of something/how to do something*; (ii) conoscere* – 'to be familiar with something'), learning the distinction between the two verbs of knowing might be challenging for English speakers learning Italian as an L2. By contrast, correspondence between L1 and L2 is expected to greatly facilitate learning (e.g., *plurality* in English L1 and Italian L2). Three other levels of learning difficulty include coalescing (e.g., *he/she* in English L1 ⇔ *su* in Spanish L2), absent category (e.g., tone system exists in Vietnamese L1, but it is absent in

English L2), and new category (e.g., there is no article system in Vietnamese L1, but there is an article system in English L2).

Table 5: Five categories of L1–L2 mapping			
Type of differences	Description	Example	Difficulty Order
Differentiation	An item in L1 becomes two or two more items in L2	L1 English: *to know* L2 Italian: *sapare* and *conoscere*	most difficult
New	An L2 item is not found in L1	L1 Vietnamese: no article system L2 English: article system	
Absent	An item in L1 is absent in L2	L1 Vietnamese: tones L2 English: no tones	
Coalesced	Two items in L1 coalesce into one item in L2	L1 English: *she, he* L2 Spanish: *su*	
Correspondence	No difference or contrast is present between L1 and L2	L1 English and L2 Italian: *plurality*	no difficulty

The following studies focused on L2 learners' avoidance of L2 words or phrases when L1 and L2 had a differentiation (i.e., one L1 item ⇔ two L2 items). We wanted to know how the L1–L2 lexical transfer was evident when learners had a higher L2 proficiency (i.e., learners were not absolute L2 beginners).

In Study 13, Dagut and Laufer (1985), for instance, selected 15 English phrasal verbs and their single counterparts. They gave native English speakers a cloze test using one of four verbs (e.g., Sentence: *We didn't believe that John could ever ____ his friend.;* Choices: *a. let down, b. solve, c. disappoint, d. carry on*). They observed that native English speakers preferred a phrasal verb (e.g., *let down*) over a semantically equivalent one-word verb (e.g., *disappoint*). They further examined whether and to what extent L1 Hebrew speakers (who had been learning L2 English for 7–8 years) would also prefer phrasal verbs to their single-word counterparts. The multiple-choice test, a verb translation test, and a verb memorizing test indicated that L1

Hebrew subjects preferred L2 English single verbs (e.g., *enter*) over semantically related phrasal verbs (e.g., *come in*). Also, they avoided L2 English figurative phrasal verbs (e.g., *let down* ⇔ *disappoint*) compared to literal (e.g., *come in* ⇔ *enter*) and completive verbs (e.g., *cut off* ⇔ *disconnect*). The reason that L1 Hebrew speakers chose single-word verbs in L2 English is that every L1 Hebrew verb comprises one word only, and there is no category of phrasal verbs in L1 Hebrew.

In Study 13, in the case of a Hebrew verb corresponding to two English items (e.g., L1 Hebrew: *להיכנס* and L2 English: *to enter, to come in*), it was difficult for native Hebrew speakers to learn and to remember the two L2 items. Native Hebrew speakers learned and preferred to use the single-word English verb (e.g., *to enter),* which has similarities of meaning and form compared to the L1 Hebrew word. Learners with higher L2 proficiency could successfully acquire the two L2 items referring to one L1 item. However, they preferred the item that is more similar to their L1 word.

In Study 14, following up on the results of Study 13, Hulstijn and Marchena (1989) tested the hypothesis that "Dutch learners of English would tend not to avoid English phrasal verbs (e.g., *turn up* ⇔ *appear*), since phrasal verbs (e.g., *op komen dagen*) also exist in Dutch" (p. 241). They tested six independent groups of Dutch intermediate and advanced English learners via three tests consisting of multiple-choice, memorization, and translation tasks. They found that L1 Dutch subjects who were advanced English learners did not avoid L2 English phrasal verbs (e.g., English *turn up* ⇔ Dutch *op kommen dagen*). Although intermediate learners tended to avoid phrasal verbs to some extent, they did not avoid them as a category.

Thus, the evidence indicates that L1 Hebrew speakers avoided L2 English phrasal verbs categorically, while L1 Dutch speakers did not. Hulstijn and Marchena (1989) accommodated the findings by Dagut and Laufer (1985) in terms of semantic considerations because avoidance of phrasal verbs differed for literal, completive, and figurative phrasal verbs. They also noted that Dutch speakers adopted a "play-it-safe" strategy of using generic one-word verbs more frequently than highly specific and idiomatic phrasal verbs (p. 241). Nonetheless, the similarities and differences between L1 and L2 (i.e., L1 Hebrew: single verbs only, L1 Dutch: single verbs and phrasal verbs as synonyms, and L2 English: single verbs and phrasal verbs as synonyms) might have influenced the avoidance of phrasal verbs in L2 English.

In Study 15, Laufer and Eliasson (1993) conducted another study to investigate the relative contribution of L1–L2 differences and similarities and the complexity of L2 features on the avoidance of English phrasal verbs. They tested 87 native Swedish adult university students, all advanced learners of L2 English, in the three following tests.

- In the multiple-choice test, subjects had to choose one of four alternatives (a correct phrasal verb, a correct single-word verb, and two distractors) to fill in a blank in sentences (e.g., Sentence: *He hoped he'd be _____ by his friends in the argument.*; Choices*: a. driven out, b. backed up, c. suggested, d. supported;* Correct choices: b and d).
- In the translation test, participants were asked to translate a Swedish verb into English to insert in a blank in sentences (e.g., Sentence: *He hoped he'd be _____ by his friends in the argument.;* The Swedish verb "*uppbackad*" was given to translate into English.).

- In the comprehension test, subjects had to translate a capitalized expression in an English sentence into Swedish (e.g., sentence*: He hoped he'd be BACKED UP by his friends in the argument.*).

The results established that Swedish speakers, in general, did not avoid English phrasal verbs, and they used phrasal verbs much more than Hebrew speakers did (Dagut & Laufer, 1985) in both the multiple-choice and the comprehension tests. Another difference was that Hebrew-speaking learners mostly avoided figurative (vs. literal) phrasal verbs, while Swedish-speaking learners did not avoid them specifically. These findings could be accommodated by L1–L2 differences. For instance, forms of phrasal verbs are present and corresponding in L1 Swedish and L2 English, but not the same between L1 Hebrew and L2 English.

The three studies on L2 avoidance agreed that an L1–L2 difference is the best predictor of phrasal-verb avoidance in L2 learning. However, we did not think that this was because of something new in L2 that was absent in L1. We thought that there were differentiation and correspondence between two languages. We could understand why L1 Hebrew speakers avoided using L2 English phrasal verbs. There are two different L2 English forms (i.e., single-word form and phrasal form) corresponding to a single Hebrew verb. L1 Hebrew speakers preferred the L2 English item, which corresponds to the L1 form and L1 meaning, and they avoided using the other one. For the L1 Dutch and L1 Swedish speakers, L2 English phrasal verbs were not much avoided in L2 usage (compared to L2 single verbs) because both forms of phrasal and single verbs are synonyms in both L1 and L2.

We can also see how L1–L2 correspondence and differentiation affect the L2 learning process with the Revised Hierarchical Model. The Revised Hierarchical Model (RHM) of Kroll and Stewart (1994) assumed that at the early stages of L2 acquisition, an L2 word was directly linked to the learners' corresponding L1 translation (i.e., learners' L1 conceptual representation in their mind). In other words, the L2 word was, at the beginning stage of L2 learning, associated with a concept via a powerful lexical connection to its L1 translation equivalent. However, when learners had higher L2 proficiency, they developed a direct link between the L2 word and its concept (Kroll & Sunderman 2004; Kroll & Tokowicz, 2005; Sunderman & Kroll, 2006).

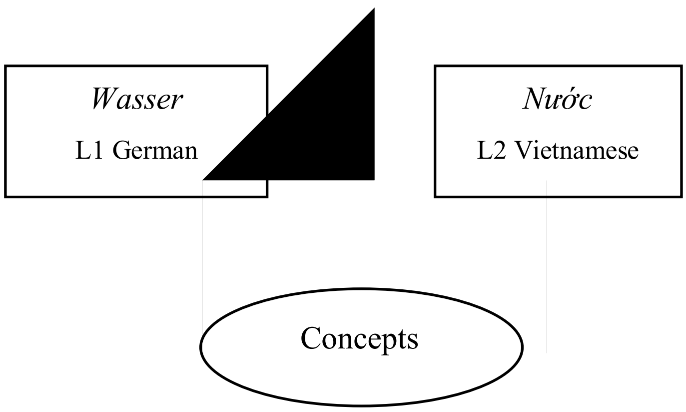


Figure 7: The Revised Hierarchical Model (adapted from Kroll and Stewart, 1994)

Pavlenko (2009) developed anfigother model – the Modified Hierarchical Model (MHM), which built on the RHM by Kroll and Stewart (1994). She considered that L2 learning perhaps involves reorganizing conceptual storage together with particular L1 concepts, particular L2 concepts, and shared L1–L2 concepts. For example, L1 Russian learners of L2 English will establish a new conceptual category for the concept of *privacy* in English because there is no conceptual equivalent in Russian (Pavlenko 2009, p. 138–140). In another example, L1 English learners of L2 Finnish will restructure their concept for *fall* because, according to Jarvis and Pavlenko (2008, p. 80), there are two Finnish equivalents including *pudota* (to fall from a higher to a lower altitude) and *kaatua* (to fall from a vertical to a horizontal position). The MHM assumes that differences in conceptual equivalence relationships between L1 and L2 result (i.e., fully shared, partially shared, or non-shared) in the three following different types of learning processes.

- L1–L2 conceptual equivalence: L2 beginners can quickly form the link between L1 and L2 words with the same existing concepts.
- L1–L2 partial conceptual equivalence: L2 learners have to restructure their existing L1 concepts and create new links between L2 words and the concepts that are not known in L1.
- L1–L2 conceptual non-equivalence: L2 learners need to build new concepts for the L2 words they learn.

As a consequence, L2 learners, at the early learning stage, can benefit from L1–L2 conceptual equivalence (i.e., L1–L2 correspondence). However, they have many more difficulties in L2 learning in two other cases, such as partial L1–L2 conceptual equivalence (i.e., L1–L2 differentiation) and conceptual L1–L2 non-equivalence (i.e., something new or absent in L2 compared to L1).

4.2.2. Lexical Transfer in Second Language Processing

Learning L2 vocabulary in L1–L2 correspondence conditions might be not difficult for L2 beginners when compared with differentiation conditions because learning new L2 word mapping in an L1 translation is much easier than learning two more L2 word mapping words for only one L1 translation. For non-native speakers with good L2 knowledge, how they processed these two language cases was investigated in two studies by Jiang (2002, 2004) examining the central claim of the adult L2 vocabulary acquisition model—"the concept onto which an L2 word is mapped is an L1 concept" (Jiang, 2004, p. 419).

Studies 16 and 17: Jiang (2002) compared how non-native young adults (L1 Chinese, advanced L2 English, and university students) versus native English speakers interact with two kinds of word pairs having a related meaning (e.g., *problem* and *question*) or a word unrelated in meaning (e.g., *matter* and *value*). For the related meaning, *problem* and *question* are translated into the same Chinese word *wenti*). By contrast, *painter* and *artist* are semantically related in English but translated into two different words *huajia* and *yishujia* in Chinese. Both the related and unrelated English word pairs were presented at the center of a computer screen. Participants were asked to judge whether the two words in a pair were related in meaning by pressing one of two buttons (negative vs. positive). Participants' accuracy and reaction times were measured and analyzed. The results showed: a main effect of participant type (native speakers performed the task better than non-native speakers); a main effect of pair type (related pairs responded faster than unrelated pairs); a significant interaction between participant type and pair type in reaction time. More importantly, non-native English speakers responded to related L2 pairs sharing the same L1 Chinese translation (L1–L2 differentiation) significantly faster than to pairs with different translations (L1–L2 correspondence) in L1. Jiang's (2004) study used the same research method as in Jiang (2002). He tested Korean–English bilinguals as non-native English speakers. Performing a semantic judgment task, the L1 Korean speakers of L2 English learning were significantly faster when responding to the same translation pairs (i.e., two English words share the same Korean translation – *behavior* and *action* ⇔ *행동*) than to the different translation ones (i.e., two English words do not share the same Korean translation – *ability* and *capacity* ⇔ *능력* and *용 량*). The same-translation effect was replicated from the previous study (Jiang, 2002), which confirmed the positive effects of semantic transfer from L1 to L2 lexical processing. More specifically, adults with an advanced L2 level processed word pairs of L1–L2 differentiation (e.g., an L1 Korean word: *행동* ⇔ two L2 English words: *behavior* and *action)* much better than L1–L2 correspondence word pairs (e.g., two L1 Korean words: *능력* and *용 량* ⇔ two L2 English words: *ability* and *capacity*).

From the L2 processing studies, the author thought that the reason that the same-translation effect was found in L2 adult learners was that their L1 semantic structure was present in L2 lexical representations during L2 processing. That can be explained by the adult L2 vocabulary acquisition model (Jiang, 2000). Jiang proposed a three-stage model based on the belief that "the existing L1 linguistic and conceptual systems are actively involved in the L2 learning process" (Jiang, 2004, p. 417).

- Stage 1 – Lexical association stage: An L2 word is recognized in an orthographic or/and phonological form. In other words, a lexical entry is registered in the learners' mental lexicon. Adults associate the L2 word with its L1 translation, and, based on this association, they can understand the meaning of the L2 word. In this stage, there is activation and mediation of L1 influence on L2 lexical processing because there is no connection between the L2 word and its concept. As a result of continued coactivation, a transfer of lemma information (i.e., syntactical and semantic) from L1 to L2 was assumed to occur.

- Stage 2 – Lemma mediation stage: The L2 word, in this stage, is linked directly to its conceptual representation. The semantic and syntax features transferred from the L1 translation now reside in L2 form. Other novel L2 meanings and L2 lemmas, which may include both L1 and L2 specifications, can be developed via contextualized inputs.
- Stage 3 – Full integration stage: The L2 word can be used with little or no influence from the L1 translation when the integration of the L2 lexical knowledge in its entry is correctly done.

Because Jiang applied the model of L2 vocabulary acquisition in the classroom learning environment with formal instruction and limited L2 inputs, the model did not address L2 conceptual development and restructuring as they might occur in cultural contact situations (Pavlenko, 2009).

Summarizing Section 4, it is clear that learners can learn or process L2 specific words or phrases better when these words or phrases are similar (vs. different) to their L1 counterparts, or, in other words, L1–L2 word correspondence (compared to L1–L2 word differentiation) can facilitate L2 vocabulary learning or processing. Also, cross-linguistic influences differ for L2 beginners, intermediate, and advanced learners in different situations of language learning and language processing. In word-level studies, previous research has mostly examined how L1 speakers with prior L2 knowledge learn and process an L2 or how L1 speakers learn L2 words that share a close relationship with the L1 word. In the current research, we wanted to investigate how novices/beginners can learn basic L2 phrasal vocabulary when there is (and when there is not) L1–L2 correspondence. Would the novices learn L2 verb-noun phrases better when these phrases share features (e.g., verb-argument relations) with their L1 counterparts?

From the background of Sections 2, 3, and 4, this current state of the art motivated us to conduct experiments in the laboratory on how successfully L2 adult beginners (i.e., L1 German) learn L2 verb-noun phrases (i.e., L2 Vietnamese) with/without supporting visual cues when there is a correspondence or a differentiation of L1–L2 verb mapping.

In Section 5, we will discuss the influences of AoA in L2 vocabulary learning. The age factor has been suggested as one of the highly influential factors in L2 learning. We thought it would be advantageous to know how the age factor was tested with different tasks, types of stimuli, languages, and participant populations or ages range in previous studies.

5. Age of Acquisition Related Differences in Second Language Learning

More generally, two important distinct age-related conclusions for long-term L2 learning have been widely accepted (Krashen et al., 1979):

- younger children generally make less rapid progress in the first stages of L2 acquisition process than older ones, adolescents, and adults
- the younger the L2 starter, the more likely they are to achieve native-level proficiency

In this section, we mainly concentrated on age effects on L2 vocabulary learning, especially for L2 beginners in different methodological designs.

Study 18: Muñoz (2008) reviewed many studies on the effects of age on L2 learning in foreign language settings. She focused on studies with some standard-specific variables (i.e., the L2 learners' age of testing, the amount of exposure to the L2, and the intensity of the L2 exposure) in the methodological design. She pointed out the effects of age on L2 learning following three main research categories:

(i) same age at time of testing, different amount of exposure
(ii) different age at time of testing, same amount of exposure
(iii) same age at time of testing, same amount of exposure, and different intensity

For the (ii) research category, a significant finding from 22 reviewed studies is that older learners showed higher L2 learning efficiency than younger ones when the amount of exposure was identical. The age of testing in those comparisons is mostly within or between child groups (2–10 years old) and adolescent groups (11–19 years old) in a longitudinal L2 learning situation (e.g., 18 weeks, 200 hours, or a year) in a school or a program.

Study 19: One of three experiments conducted in a short-term training duration concerned L2 Russian listening comprehension (Asher & Price, 1967). Four L1 English participant groups (8, 10, 14 year-olds, and adults) had three short training units during which they listened to taped L2 Russian commands and watched a person responding to them in a video (e.g., *walk to the chair*). Their task in a retention test was to listen and obey the L2 Russian trained commands and new commands. Results indicated that adults outperformed the adolescents and the children, and both the 14-year-olds and the 10-year-olds significantly outperformed the 8-year-olds. The study claimed that older children perform better than younger children, and adults perform better than older children when all of them are learning a second language under the same conditions.

Study 20: Another short-term L2 learning study (Olson & Samuels, 1973) tested the L2 English pronunciation of L1 German subjects in three age groups (9.5–10.5, 14–15, and 18–26) after subjects had been taught 33 phonemes in 13 sessions over two weeks. In training, subjects heard a recorded German voice. Then, they learned to pronounce the words as accurately as possible. In testing, they were pretested and post-tested and given the Raven Progressive Matrices Test. In the pretest, no significant difference in pronunciation was found. However, in the post-test, the older groups (14–15 and 18–26) were significantly ($p < .01$) better at German pronunciation than the younger group (9.5–10.5). The evidence of the study suggested that adults perform much better than children in foreign language pronunciation.
Related studies focused on a different age at the time of testing, or the same amount of exposure: Studies 21, 22, 23, and 24 focused on age-related effects on L2 or L3 vocabulary learning.

Study 21: Yamada et al. (1980) tested L1 Japanese children in three different groups of 10 (7 years old, 9 years old, and 10 years old) learning 40 L2 English words. Participants had no previous L2 knowledge. Each of them learned four English words with one syllable or two syllables (e.g., *leaf, snail, hippo,* and *ladder*) via a word-picture learning paradigm. There were two learning sessions (day 1 and day 2) for each participant. In the first learning session, participants had 10 experimental trials of (spoken) word-picture pairs. They were asked to repeat the names of the object pictures after the experimenter named them. If participants gave an incorrect pronunciation of the object name, the experimenter would tell them the correct response. Then, the participants had to repeat the correct word. In the second learning session, the same subjects were given the same items as in the first session with the same learning procedure to investigate whether the experience of the first session improved the participants' learning success in the second session. The accuracy of the participants' responses was measured in both sessions (i.e., a correct response scored 1, while an incorrect response or no response scored 0). The results with a $3 \times 2 \times 2$ (i.e., age: 7 or 9, or 10; session: day 1 or day 2; syllable: one or two) analysis of variance design (two factors, session and syllable, had repeated measures) showed no main effect of age, but a three-way age × session × syllable indicated significant interaction. A trend analysis was also performed showing that the mean learning scores declined significantly with age ($F = 13.53$, $df = 2.27$, $p < .001$) in which older children had lower scores than younger ones. Two abilities, including rote memory (i.e., storing the meaning of words when seeing pictures) and motor ability (i.e., involving accurate pronunciation when hearing spoken words), were applied to explain why young children learned L2 words faster and more successfully than older children. These abilities and their interaction were believed to play an essential role in learning L2 words because younger children supposedly had better rote memory and better motor ability than older children. The findings of the study seemed to be the opposite of the finding of Asher and Price (1967, see above). However, this could be understandable because of differences in the L2 participants (L2 knowledgeable learners vs. L2 beginners) and different L2 learning tasks (L2 listening comprehension vs. L2 word learning).

Study 22: Vilke (1988) taught and tested 120 L1 Croatian children learning L2 English following two different groups (early starters: aged 9 vs. late starters: aged 17–19). In the project, many linguistic aspects such as pronunciation, grammar, vocabulary, and language awareness were taught and examined for a year. Some main results summarized by Muñoz (2008) showed that early L2 starters were only better in L2 pronunciation than late L2 starters, while they were not as successful as late L2 starters in L2 grammar, vocabulary, and language awareness. The finding for L2 pronunciation acquisition is also in line with the study by Yamada et al. (1980).

Study 23: The study by Miralpeix (2006) focused on age effects on vocabulary acquisition. Two groups of Catalan/Spanish bilingual individuals learning English as a third language (L3) participated in a long-term study. They started learning English at different ages (early starters: 8 vs. late starters: 11), and they were tested at average ages (16.3 vs. 17.9) after learning English with the same amount of instructed language exposure (726 hours = 7–9 school years). Learners performed tasks: (i) oral tasks: an interview with a researcher (i.e., talking about hobbies), telling a story with six pictures presented together, and a role-play oral test (i.e., a mother talking about organizing a party); (ii) written tests: a free productive vocabulary (i.e., participants wrote a composition in a maximum time of 15 minutes), and controlled productive vocabulary (i.e., a cloze test). The author used *VocabProfile* (Nation, 1995) to obtain and compute the number of types, tokens, and word-families of both oral and written tests. All measures were calculated into non-standardized and standardized (50 tokens) text length.

Also, *t*-tests for independent samples were performed to establish any significant differences between early starters and late starters.

In short, after 726 hours of formal exposure, early English learners did not outperform late English learners in the free productive vocabulary task. Differences between the two groups were not significant, although late starters achieved higher scores in tasks like storytelling and interviews than early starters. In the controlled productive vocabulary task, the older group had a clear advantage because they performed significantly better than the younger group. The findings of the study were also in line with previous studies (Burstall et al., 1974; Lasagabaster & Doiz, 2003; Oller & Nagato, 1974; Singleton, 1999) with evidence in favor of late starters.

Study 24: Miralpeix (2007) carried out another study investigating the effects of age of onset (AO), Age at Testing (AT), and Amount of Exposure (AE) on producing vocabulary of learners of English as a foreign language. Three groups (A1, B1, and A2) learning English at different ages (A1, 8 vs. B1, 11 years old vs. A2, 8 years old) were tested with differences of average age (A1, 16.9 years vs. B1, 17.9 years vs. A2, 17.7 years) after acquiring a different amount of language exposure (A1, 726 hours vs. B1, 726 hours vs. A2, 800 hours). They performed various tasks, comprising oral tasks (an interview, storytelling, and role-play) and one written task (a composition). The main results showed:

- no significant difference in lexical gains between the two different AO groups (A1, early starters and B1, late starters) when the AE was constant.
- also no significant differences between A1 and A2 groups; however, A2, as an earlier starter group having 800 hours of English exposure, performed similarly to the B1 later group with 726 hours of exposure. The results suggested further research questions about the highest AO of foreign language start and the amount of foreign language exposure in formal settings for learners' foreign language success.

In short, the above studies, including the summary table in the review by Muñoz (2008, p213-217), and studies with the same methodological design (i.e., same specific variables: L2 learners' age of testing, the AO of L2, and the intensity of the L2 exposure) in the category *different age at time of testing, same amount of exposure* had the following main features:

- Researchers compared different L2 learners groups such as younger children vs. older children (age range at AT was from 8 to 18) or children (AT 7–18) vs. early and young adults (AT 18–38). Participants having vs. not having previous L2 knowledge were tested.
- L2 mostly learned or taught in long-term conditions, both formal and informal learning environments, and the amount of L2 exposure was equal among the L2 learner groups.
- L2 learning and testing tasks were different among studies, and many linguistic skills (i.e., pronunciation, listening, reading, speaking, and writing) and aspects (i.e., vocabulary and grammar) were examined.
- Main findings: The main finding of this line of research is that late L2 starters (adults and older children) generally learned or processed L2 more successfully than early L2 starters.

To some extent, our methodological design was influenced by these earlier studies and findings:

- A comparison of adults' L2 learning success should be performed to see what would happen if early adults, young adults, middle-aged adults, and older adults have the same amount of L2 exposure when they begin to learn an L2. Testing people with no prior exposure to the examined L2 is suitable to investigate different influences of L2 learning conditions and testing paradigms on their L2 learning success across the life span.
- There were three studies of L2 vocabulary short-term learning (i.e., learning and testing in a day or several hours) in the 22 studies reviewed. One could argue that more empirical short-term studies would be necessary and helpful in researching L2 vocabulary learning.

We also tried to examine other recent studies about the relationship between learners' AoA and their L2 vocabulary learning. The following points are noteworthy:

AoA effects are not restricted to items learned before any putative critical period but should be observed for items learned at any age. The AoA effect is due to the relative order in which items are acquired within a language. An early AoA, regardless of the learners' L2 proficiency, is crucial in finding the L2–L1 priming effects (Hirsh et al., 2003).

The AoA of a concept can exert an influence on processing, independent of the AoA of the word form. Different training methods may lead to qualitative differences, like the links formed between words and concepts during the earliest stages of L2 learning (Palmer & Havelka, 2010).

AoA—not bilingualism—is the primary determinant of ultimate L2 attainment, irrespective of the phonetic, grammatical, and lexical acquisition. However, "the sooner, the better" in relation to AoA does not necessarily hold for language development. It has also been found that even very young L2 learners might not catch up with native speakers in a set time frame. Factors other than age may play a role in determining individual variation in L2 child learners' long-term outcomes with English morphology. In fact, for the correct use of specific linguistic aspects, it does not matter when children start to learn a language. To learn words quickly in a new language, it may be helpful to first build a substantial vocabulary in the first language before learning the new language (Blom & Bosma, 2016).

Age is considered to be an internal factor. However, whether learners' internal or external factors account for more of the variance in acquisition outcomes is still a controversial issue. It has been found that the development of vocabulary and of convoluted syntax were affected by internal and external factors. In contrast, external factors did not contribute to the development of tense morphology (Paradis, 2011).

To summarize this section, AoA would be an essential factor or variable in the short-term experimental design of L2 vocabulary learning. Also, the role of AoA should be observed in the relationship with learned L2 items. Furthermore, how AoA of L2 adult beginners, 18–65 years old, for instance, can influence their short-term L2 vocabulary learning, is one of our research questions. In Section 6, we mentioned another factor—learners' cognitive performance, which has also been shown to affect L2 vocabulary learning.

6. Cognitive Performance of Second Language Learners

Researchers suggest that adult cognitive development is a complex process (Fischer et al., 2003). Intelligence can be measured by a test that rates individual cognitive abilities in comparison with the general population. Cognitive abilities are brain-based skills, and mental processes are often estimated using standardized intelligence tests and verified measures. The psychometric method has classified two categories of intelligence that illustrate distinct rates of change across the lifespan (Schaie & Willis, 1996). Fluid intelligence relates to information processing abilities, such as logical reasoning, remembering lists, spatial ability, and reaction time. Crystallized intelligence involves abilities that use experience and knowledge. Tests of crystallized intelligence include vocabulary tests, solving number problems, and understanding texts. Next, we briefly reviewed studies concentrated on how learners' intelligence and cognitive ability influence their L2 learning.

Study 25: Rouhi and Mohebbi (2013) examined whether different glosses (i.e., pictorial, pictorial + sound, and video) and spatial intelligence (i.e., low and high spatial ability) influence L2 vocabulary learning and retention in a multimedia context. Sixty-two pre-university students (18–20 years), L1 bilingual in Azari-Turkish and Persian, were randomly assigned to four different groups (i.e., pictorial glosses, n = 17; pictorial + sound glosses, n = 17; video glosses, n = 16; and no glosses, n = 12). They had learned L2 English with formal instruction for six years. In the study, they consulted one of the different glosses and read L2 English sections for six sessions. They then performed sentence completion tests (i.e., immediate post-test vs. delayed post-test 25 days later). Participants also completed a 10-item questionnaire using a five-point Likert scale measuring spatial intelligence (McKenzie, 1999). Results showed a significant positive effect of multimedia glosses on L2 vocabulary learning, but no significant difference between learners having a high or low spatial ability.

Study 26: Shakouri et al. (2017) investigated whether learners' linguistic intelligence could affect recalling lexical items in SLA. Forty L1 Iranian participants (from 16 to 23 years) with intermediate L2 English participated in the study. Their linguistic intelligence was measured via McKenzie's (1999) questionnaire and an L2 vocabulary recall test. Next, participants were given a recall test with two different lists (i.e., categorically related and semantically related), including 40 L2 lexical items per list. Results demonstrated a significant correlation between subjects' linguistic intelligence and their recall scores of L2 lexical items.

Study 27: Tajeddin and Chiniforoushan (2011) examined 91 low-proficiency L2 English learners (L1 Iranian, 15–30 years) with the visual intelligence section of the Multiple Intelligences Developmental Assessment Scales (Shearer, 1996), a multiple intelligence test, and a test measuring receptive and productive L2 vocabulary acquisition. After a learning phase, participants performed a multiple-choice test of L2 vocabulary reception (20 items) and a cloze test of vocabulary production (also 20 items). Results indicated a marginally significant correlation ($p = .04$) between visual intelligence and vocabulary reception and production.

Study 28: Morvay (2015) investigated how L1 reading ability, L2 proficiency, and learners' nonverbal intelligence influence their L2 reading comprehension. Sixty-four Hungarian high-school students from 17 to 19 years old learning English as a foreign language in Slovakia participated in the study. They learned L2 English for eight years on average. In the research, participants took four tests, including the Michigan Listening Comprehension Test (L2 proficiency test), the Gates-MacGinitie Reading Test (L2 reading comprehension test), the Hungarian National Reading Competency Measure (L1 reading ability test), and the Naglieri Nonverbal Ability Test (a nonverbal intelligence test). A significant correlation between nonverbal intelligence and L2 reading comprehension was found in the study. Also, L1 reading ability was transferred to L2 reading comprehension when the participants had high L2 proficiency.

In short, Studies 25 to 28 consider many aspects of intelligence and how they might influence different aspects of L2 acquisition. The findings indicate that spatial intelligence might not influence vocabulary learning. On the other hand, participants' linguistic intelligence was significantly correlated with the participants' recall of lexical items in L2. However, it has been postulated that spatial and linguistic intelligence might not correlate with the phonological, semantic, and orthographic aspects of words in L2 vocabulary learning. Visual intelligence also plays a role in L2 acquisition, and it can be activated by instructional consciousness-raising procedures. Lexical input enhancement through marginal glosses plus pictures can be more facilitative in improving vocabulary reception than vocabulary production as the former entails less demanding cognitive and memory operations. Nonverbal intelligence moderately influenced L2 reading comprehension. Intelligence also has a different influence on different aspects of L2 learning. Performance on L2 reading and L2 usage tests was significantly correlated with the subjects' intelligence level. However, performance on the tests of listening comprehension and interpersonal communication skills did not correlate with the intelligence level.

With age, systematic declines are observed in cognitive tasks needing self-initiated, effortful processing, without the assistance of supportive memory cues (Park, 2000). Older adults tend to perform less successfully than young adults in memory tasks that require recall of information. Individuals must retrieve the information they learned previously without the aid of a list of feasible choices. When we age, working memory, or our ability to concurrently store and use information, becomes less effective (Craik & Bialystok, 2006). The ability to process information quickly also declines with age. This slow speed may explain age differences in many different cognitive tasks (Salthouse, 2004). Finally, it is well established that our hearing and vision decrease as we age. Longitudinal research has proposed that deficits in sensory functioning explain age differences in various cognitive abilities (Baltes & Lindenberger, 1997). Fewer age differences are found when memory cues are available, such as for recognition memory tasks, or when people can draw upon gained knowledge or experience. Older adults often perform as well if not better than young adults on tests of word knowledge or vocabulary.

In some language processing studies, cognitive tests were administrated to participants to control for key cognitive variables. Then, participants' scores for cognitive tests, for instance, were computed and analyzed to test for differences between various age groups (i.e., younger vs. older adults; children vs. adults) on these variables. That analysis also contributes to explaining the results of their language processing.

Study 29: Carminati and Knoeferle (2013) carried out two visual-world eye-tracking experiments to investigate effects of (native German) speakers' emotional facial expression (happy vs. sad) as well as listener age (younger vs. older adults) on sentence (negative vs. positive) processing in a visual context (i.e., hearing a spoken sentence and seeing two pictures presented simultaneously). Additional cognitive tests were conducted before the eye-tracking study. The tests chosen were subtests from the Wechsler Intelligenztest für Erwachsene (WAIS, Wechsler, 1981). Cognitive performance was estimated using the picture completion, the digit symbol mapping, the digit span, and the similarities subtests (Wechsler, 1981). The latter asked participants to list as many examples of a category (animals) as possible and as many words with an initial letter (the letter "L") as possible within one minute (i.e., vocabulary test). All subjects were administered the same tests. The cognitive tests ensured that the online and offline results of the eye-tracking study were not confounded by between-subject differences in cognitive performance, such as verbal, working, and spatial memory. Conducting the cognitive tests took from 15 to 20 minutes. The results of the cognitive tests showed that younger adults (18–30 years) performed significantly better than older adults (60–80 years) in the picture completion and the digit symbol tasks of the WAIS test. In other tests (digit span, similarities, and vocabulary), the two groups did not differ significantly. The eye-tracking experiment's main findings indicated that enhancement of participants' eye fixations on pictures from the early stages of the referential disambiguation was modulated by age; it was more frequent with negative faces for younger adults, and with positive faces for older adults.

Study 30: Additional cognitive tests were added in Münster's (2016) study mentioned in Section 3. These tests were performed by the children's group (4–5 years) and the older group (60–80 years). Münster used the cognitive tests (the digit span, the word order, and the spatial memory test) of Kaufmann Assessment Battery for Children (K-ABC) and WAIS tests for older adults as in Carminati and Knoeferle (2013). The results of the cognitive tests were linked with participants' accuracy in comprehension questions in each group to determine correlations. The children performed worst in the digit span test, and there was no reliable correlation between accuracy and their cognitive scores. The older adults performed best in the similarities task and worst in the digit symbol mapping task. The mean for the words generated in the verbal fluency task was 38.825. The author computed Spearman's Rho to test whether the accuracy results for the comprehension questions and the cognitive test scores correlated. The result yielded a medium-size but significant correlation ($r_s = .372$, $p < .05$).

Cognitive ability, as one of the individual differences, clearly plays a crucial role in language processing and language learning. Our L2 learning study will also test native German speakers, and we will also use cognitive tests to control differences within or between L2 learner groups in their cognitive scores (ability). That might help to explain the differences in participants' L2 learning success.

As a background for our current research, we reviewed different L2 vocabulary learning and testing contexts (Section 2), effects of visual learning contexts (Section 3), L1 influences when people learn a L2 (Section 4), learners' age-related differences of L2 learning (Section 5), and learners' cognitive performance related to their L2 learning ability (Section 6). We learned from previous studies that people learned L2 single words successfully in various visual learning contexts, especially in auditory-visual learning conditions with familiar or known referents. However, no empirical research on L2 phrasal vocabulary learning in various visual learning conditions for L2 beginners has been found. The next point is that L1-L2 similarity effects on L2 vocabulary learning has been investigated for the word-level, but not phrasal level. Also, learners' age and cognitive ability have been considered as internal factors which could strongly affect their language learning or processing success. We would furthermore examine influences of these factors in short-term L2 phrasal vocabulary learning studies with continuous age groups of adults. The thesis would examine L2 beginners from 18 to 65 years-old learning L2 verb-noun phrases in different visual and linguistic learning conditions. In Section 7, L2 phrasal learning design, research questions, and predictions would be mentioned in detail.

7. Research Questions

As Sections 2, 3, 4, 5, and 6 have demonstrated, L2 vocabulary learning is influenced not only by learners' internal factors (e.g., AoA and cognitive ability) but also by their external factors (e.g., visual context, L1–L2 similarity, learning and testing paradigm). We wanted to conduct a set of studies on L2 short-term phrasal vocabulary learning, combining all the mentioned factors. Thus, the present studies aim to answer the following main research questions (RQ).

RQ1. How do visual context (e.g., depicted actions) and language context (e.g., L1–L2 similarity) affect L2 adult beginners' accuracy (i.e., correct responses) and speed (i.e., reaction times) when they learn L2 phrasal vocabulary and are tested in different testing parts (i.e., immediate vs. delayed testing)?

RQ2. What are the age-related differences in adults' second language learning success?

RQ3. Are there relationships between adults' cognitive performance and their second language learning success?

RQ4. How do different learning and testing paradigms (i.e., delayed testing with vs. without learning repetition) affect second language learning success?

More specifically, six reaction time studies investigated the effect of present events and similarity in L1–L2 mappings on the L2 Vietnamese phrasal vocabulary learning of L1 German adults. To investigate how L2 adult learners' characteristics affect their ability in L2 vocabulary learning, we trained and tested different adult age groups, i.e., young adults, early middle-aged adults, and late middle-age adults.

We chose adults with an age range of 18–65 years old. Researchers, in many studies of language learning and language processing, tested children (under 18), young adults (often from 18 to 30 or 35), and older adults (more than 65 years). However, middle-aged adults (around 35 to 65) have not been investigated as a group. In our research, we would identify similarities and differences among L2 adult age groups in L2 vocabulary learning.

7.1 Design and Conditions

In all six experiments,[2] participants learned L2 Vietnamese phrasal vocabulary by listening to spoken words and phrases and inspecting objects. Event photographs were present in two learning conditions and absent in two other learning conditions. They also learned both types of language mapping (i.e., L1–L2 similarity and L1–L2 difference).

A 2 (present event vs. absent event) × 2 (similar language mapping × different language mapping) design yielded four learning conditions:

- Present event + similar language mapping
- Present event + different language mapping
- Absent event + similar language mapping
- Absent event + different language mapping

To estimate how successfully participants could achieve in L2 vocabulary learning, immediate testing and delayed testing were used in all six experiments. The binary forced-choice task was identical in testing. We measured the accuracy and reaction time of participants**.**

7.2 High-level Predictions

We predicted the following key results:

- RQ1: If events are useful in learning, then participants could give more accurate and faster responses in total via a button press in the "event-present" condition compared with the "event-absent" one.
- RQ1: If verb mapping makes a difference, then participants could respond more slowly and less accurately in different verb-noun mappings compared to similar verb-noun mappings.
- RQ1 – a follow-up prediction after Experiment 1: If there are strong effects of testing part, then participants could give more accurate and faster responses via a button press in the immediate testing part than they do in the delayed testing part.

[2] Six experiments: Young adults (Experiment 1, 2, 2R—the replication of Experiment 2, 2N—a new experiment design), early middle-aged adults (Experiment 3), and late middle-aged adults (Experiment 4).

- RQ2 - a follow-up prediction for Experiments 2R, 3, and 4: If age can affect L2 learning, then there are significant differences of L2 learning results among age groups (young adults, early middle-aged adults, and late middle-aged adults).
- RQ3 – a follow-up prediction for Experiments 2R, 3, and 4: Within each experiment, we predicted that participants with higher scores in the cognitive test would respond better to the testing task (higher accuracy and shorter reaction time) in the learning experiment than the other participants.
- RQ4 – a follow-up prediction for Experiments 2R and 2N: We predicted that learners would perform more accurately and faster in Part 3 (delayed testing) of Experiment 2R because of a learning repetition than in Part 3 of Experiment 2N.

In Sections 8 to 10, we describe and analyze the results of all six experiments by grouping them:

- Experiments 1 and 2: Effects of visual contexts and language similarity depending on L2 vocabulary learning and testing.
- Experiments 2R, 3, and 4: Does the AoA or cognitive performance strongly affect L2 vocabulary learning success?
- Experiments 2R and 2N: Effects of visual context in learning with single exposure vs. learning with/without repetition.

8. Second Language Vocabulary Learning with Suitable Testing

8.1 Experiment 1

8.1.1 Participants

Thirty-two adults (ages 18–31) participated in the experiment conducted in the response-time laboratory of the Psycholinguistics group at the Humboldt University, Berlin, Germany, supported by a laboratory ethics vote of the German Society for Linguistics (2016–2022). All participants had German as their only first language (L1), had acquired no knowledge of Vietnamese, and had not acquired a second language (L2) before age 6.

8.1.2 Materials

We will now describe the material construction in detail. The remaining five experiments use the same stimuli. Therefore, we will only report the changes made to the construction for the remaining five experiments.

8.1.2.1 Stimuli

In the first experiment, 32 basic Vietnamese verb-noun phrases (e.g., *chụp-ảnh* ⇔ *take-photo*) were selected from an online Vietnamese corpus—SEAlang[3]. Then, they were also checked for frequency in other online Vietnamese corpura such as Vietlex[4] and VLSP.[5] Each language stimulus included a noun (N), a verb (V), and a verb-noun phrase (P). For instance, a noun: *ảnh (*Vietnamese*)* ⇔ *photo* (English) ⇔ *Foto* (German); a verb: chụp (Vietnamese) ⇔ *take (English)* ⇔ *machen* (German); a verb-noun phrase: *chụp-ảnh* (Vietnamese) ⇔ *take-photo (English)* ⇔ *Foto machen* (German). Please note that the order of a Vietnamese verb-noun phrase in the infinitive form (verb-noun) is the opposite of a German verb-noun phrase (noun-verb).

---

[3] http://sealang.net/

[4] http://www.vietlex.com/kho-ngu-lieu

[5] https://vlsp.hpda.vn/demo/?&lang=en

A native Vietnamese speaker recorded audio sounds for all 32 phrases consisting of single verbs, single nouns, and verb-nouns phrases. The speaker used a recorder—Philips Voice Tracer DVT2710 DNS Diktiergerät, and she sat in a silent room. During the process of recording materials, the recorder was placed at about 10 cm from the mouth of the speaker. Then, we used Praat to edit the audio files.

We searched online, and we selected photographs referring to each of the nouns (henceforth "object depictions") and phrases (henceforth "event depictions": depictions of the noun referents undergoing an event). Next, we made a 245 × 245-pixel resolution, 100–200KB, in .bmp format file for each photograph to ensure a bright display for two or four pictures on a computer screen using E-Prime and Presentation programs. All stimuli in Experiment 1 are listed in Appendix A.

8.1.2.2 Counterbalancing

There were 16 Vietnamese verbs in 32 Vietnamese verb-noun phrases, which meant that each verb existed in two phrases. For instance, two phrases such as *đọc-sách/read-book* and *đọc-báo/read-newspapers* had the same verb *đọc/read*. Moreover, there were two types of verb mappings between Vietnamese and German (one-one mapping and one more – one mapping). Because the learning experiment was designed for L2 beginners, participants would learn L2 phrasal pairs per training trial. In each learning trial, a Vietnamese verb was not repeated (e.g., *mặc-áo/wear-shirt* and *đội-mũ/wear-hat*), but the verb might be the same in German (e.g., *tragen*).

Each participant would learn 32 phrases in four different learning conditions and would be tested on only half of them. Thus, we created 32 random experimental lists for 32 participants. Each list contained four pairs per four learning conditions, as will be outlined more clearly in the next section.

8.1.3 Experimental Design and Procedure

A 2 × 2 experimental design investigated how two factors influence German participants' Vietnamese phrase learning. Event depiction (i.e., absent vs. present) was the first factor. For instance, some subjects learn a phrase as *đọc sách*/*read-book* by listening to a spoken phrase and inspecting a photograph referent (i.e., event present). In contrast, others learn the phrase by listening to the spoken phrase only without the photograph referent (i.e., event absent). Each participant saw the same number of event-absent and event-present trials but only one level per item. The second experimental factor was verb mapping (i.e., similar vs. different between Vietnamese and German). A Vietnamese and a German verb either corresponded in a verb-argument mapping (e.g., *đọc* ⇔ *to read* ⇔ *lesen*), or differed (e.g., *mặc/đeo/đội* as verbs for wearing a *dress/ring/hat* respectively; German, by contrast, has one verb for wearing*: to wear* ⇔ *tragen*). Testing parts (Part 1 vs. Part2) were to evaluate learning across time, but they were not an experimental factor (see Table 6 and Figure 8)for an overview of the design.

In the initial learning phase (see Table 6 below), participants heard the recorded sounds, either with depicted objects and their corresponding event depictions or only with the depicted objects. Following the presentation of an object depiction and its naming via the previously recorded noun sound *(*e.g., *sách/book*), a verb-noun phrase was heard (e.g., *đọc sách/read-book*). Only in the event-present condition, the verb-noun phrase was additionally accompanied by an event photograph (e.g., *a person reading a book*).

Table 6: The stimuli for one example item set in relation to the manipulated factors event presence and language mapping, yielding four learning conditions[6]		
	(IV) Event depictions	
(IV) Language mappings	*Event-present*	*Event-absent*
Similar		

[6] Due to copyright restrictions, we cannot provide the original photographs in Table 1 and in the whole dissertation. Instead, the photographs illustrating the stimuli are alternatives for the object referents and events in the experiments. They come from free sources for downloading and using photographs (https://unsplash.com/). The original materials are available from the corresponding author upon request.

Different		

Participants were tested immediately after learning on each trial (each experiment part contained n = 16 trials). In testing, participants re-encountered the two objects of the preceding learning trial and heard a Vietnamese verb (e.g., *read*) building a phrase (e.g., *read-book*) with only one of the two objects presented (e.g., *book* and *table*, Figure 8). They had a preview of two pictures for 1000 ms. Then they heard the verb sound and indicated via button press as quickly and accurately as possible which of the two objects (e.g., *book* vs. *table*) formed the correct phrase with the verb heard (e.g., *read*, timeout = 2000 ms). Participants would press the Q button on the keyboard to select the left object picture. They would press the P button to select the object picture on the right side. Participants did not press any button when there was no response. Participants got feedback for each answer on the computer screen after every response. They had another opportunity to make their choice if they had either given no response the first time or made an incorrect first choice.

Figure 8: Example of testing display in Experiment 1

The procedure of the first experiment (for more detail, see Figure 9 illustrating steps 3, 4, and 5) included the following steps:

1. Reading the experimental information and filling out the required forms
2. Completing a short practice session
3. Performing the first part of the learning experiment
4. A short break
5. Performing the second part of the learning experiment (a repetition of the learning-testing regime, see Part 1)
6. Answering post-experiment questions

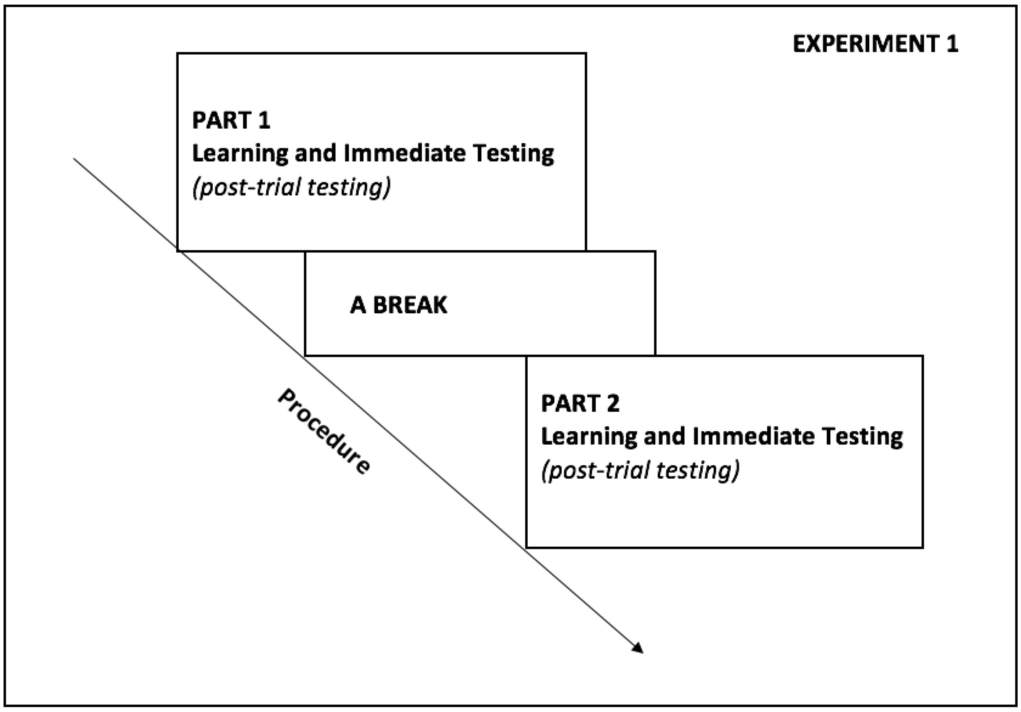


Figure 9: Procedure in the main learning experiment (Experiment 1)

8.1.4 Predictions and Analysis Methods

Previous studies have shown significant visual aids and L1 transfer effects when people were learning an L2 (see Section 2). Based on these previous results, we predicted the following: (a) If events are useful in learning, then participants should give more accurate and faster responses in the “event-present” than the “event-absent” condition. Participants may clearly understand the meaning of the noun, the verb, and the verb-noun phrase while inspecting both the depicted object photograph and the event photographs, and while listening to the sounds. By contrast, if no event photographs appear in the learning phase, participants will have little information when they are choosing the correct answer. We reason that without the depicted events, they cannot grasp the meaning of the verbs, resulting in lower accuracy and slower response times in the event-absent condition. (b) If language (L1–L2) similarity matters for learning success, participants will experience difficulties with verb-noun phrases that differ from L1 to L2. For instance, participants should struggle to learn verb-noun phrases in which one English or German verb corresponds to several distinct Vietnamese verbs in the phrases (e.g., wear ⇔ *mặc*—the Vietnamese verb collocates with *clothes*; *đội*—the verb collocates with *hats*; *đeo*—the verb collocates with *watches, rings or glasses*). By contrast, similar mapping from German to Vietnamese should enable transfer and improve learning (higher accuracy, faster response times). Alternatively, if language similarity and positive transfer do not matter for learning success, then response accuracy and speed should be identical for similar vs. different verb mapping.

We used R software (R Core Team, 2015) to analyze participants' response times (measured in milliseconds) and their accuracy (how many correct choices they made after learning) in the four different learning conditions (see Table 7). Data for both reaction time and accuracy were analyzed using a linear mixed-effects model in the lme4 package (Bates et al., 2015). Only correct trials were included for reaction times, and outliers 2.5 SD above the condition mean were excluded. Models were run with fixed factors: event presence (present vs. absent), language mapping (similar vs. different). Post hoc models were also run with testing part (as a fixed factor) included and excluded.

We also ran models for each separate testing part with two fixed factors (event presence and verb mapping) to investigate their effects. Random slopes for items and participants were included given model convergence. Then, the most parsimonious model was chosen following Bates et al. (2015), and the p-values were obtained using Satterthwaite approximations following Luke (2017). For accuracy, the generalized linear mixed model (GLMM) with family set to binomial (the outcome is binary) was run with the same factors and settings for the reaction time analysis.

Table 7: The analysis model for participants' accuracy and reaction time

	Accuracy (Acc)
Accuracy (Acc)	Without part ***glmer*** *(RT ~ Event presence * Language mapping + (1+ Event presence * Language mapping | Participant) + (1+ Event presence| Item), data = Experiment1), family="binomial",control=glmerControl(optCtrl=list(maxfun=1000000))))*

	With part **glmer** *(RT ~ Event presence * Language mapping * Part + (1+ Event presence * Language mapping * Part | Participant) + (1+ Event presence * Part | Item), data = Experiment1), family="binomial",control=glmerControl(optCtrl=list(maxfun=1000000))))*
Reaction time (RT)	Without part **lmer** *(RT ~ Event presence * Language mapping + (1+ Event presence * Language mapping | Participant) + (1+ Event presence | Item), data = Experiment1), REML=FALSE , control=lmerControl(optCtrl=list(maxfun=1000000))))* With part **lmer** *(RT ~ Event presence * Language mapping * Part + (1+ Event presence * Language mapping * Part | Participant) + (1+ Event presence * Part | Item), data = Experiment1), REML=FALSE , control=lmerControl(optCtrl=list(maxfun=1000000))))*

8.1.5 Results and Discussion 1

8.1.5.1 Accuracy Results

The main results for the accuracy analysis in Experiment1: Results yielded neither a significant effect of event depictions nor language mappings on the accuracy of participants' responses (see Figure 10) since correct answers were equally frequent in the four learning conditions of the two testing parts (i.e., more than 97% of correct responses per condition).

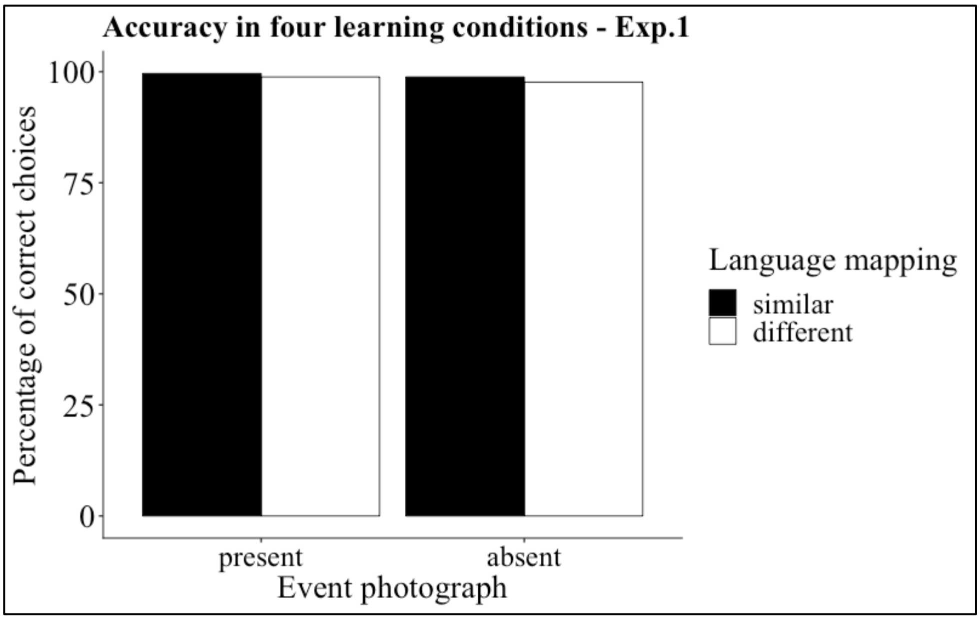


Figure 10: Accuracy in per learning condition in Experiment 1

8.1.5.2 Reaction Time Results

We found no significant main effects of event depictions or language mappings on reaction times when the part was not included in the analysis model. Figure 11 indicates that participants were a little bit faster for correct responses in present-similar and absent-similar conditions than they were in present-different and absent-different conditions. However, the main effects and significant interactions related to two experimental factors (i.e., event depiction and language mapping) w

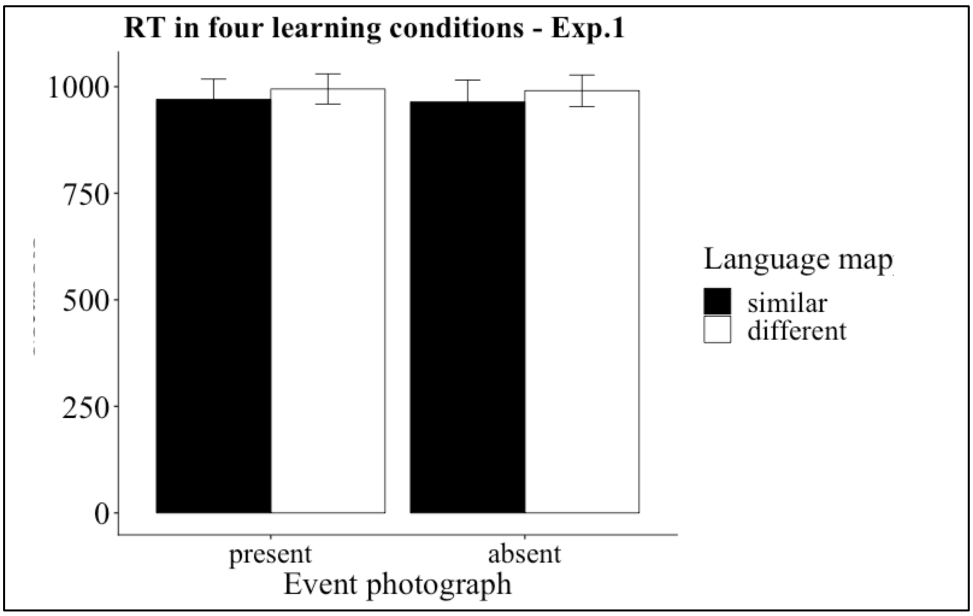


Figure 11: Reaction time in per learning condition in Experiment 1

When part was included as a factor in the analysis model (see more in Table 8), we observed a significant effect of part (*SE* = 0.009 , = 10.07, $p < 0.001$) and a significant language mapping by part interaction (*SE* = 0.008, = 3.83, $p < 0.001$).

Table 8: Results from linear mixed effects model for reaction time in Experiment 1 when testing part was included as a experimental factor in the analysis model

Term	**Estimate**	**SE**	**df**	**|t|**	**p**
Intercept	6.848	0.028	46.5	242.86	<0.001***
Event presence	0.003	0.010	22.8	0.307.5	0.76
Language mapping	-0.021	0.015	31	-1.40	0.16
Testing Part	0.095	0.009	32	10.07	<0.001***
Event presence* Testing Part	-0.010	0.007	837.6	-1.434	0.15
Language mapping* Testing Part	0.0 29	0.007	840.4	3.832	<0.001***
Event presence*Language mapping	-0.0001	0.010	29.8	-0.014	0.98
Event presence*Language mapping* Testing Part	-0.010	0.007	837.8	-1.342	0.17

In detail, results indicated a significantly shorter reaction time for all four conditions in the second testing part than the first testing part (see Figure 12). That convinced us because everything in the second part was repeated from the first part, making it very easy for learners.

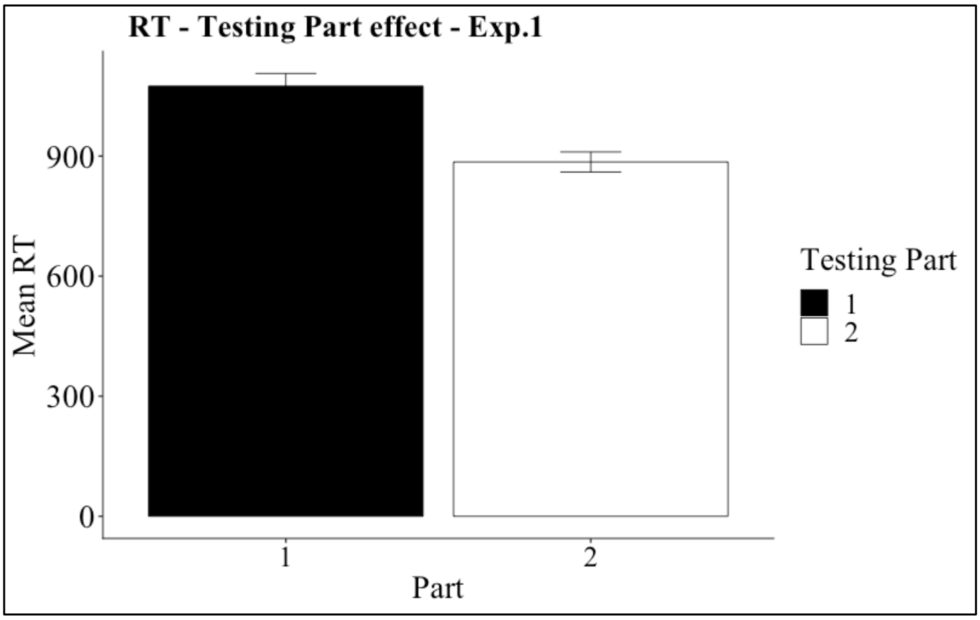


Figure. 12: Part effect in Experiment 1

For the primary interaction between language mapping and part, participants in Part 1 took longer for their correct responses in similar language conditions than in different ones. In contrast, in Part 2, they were faster in similar language conditions than in different language conditions (see Figure 13).

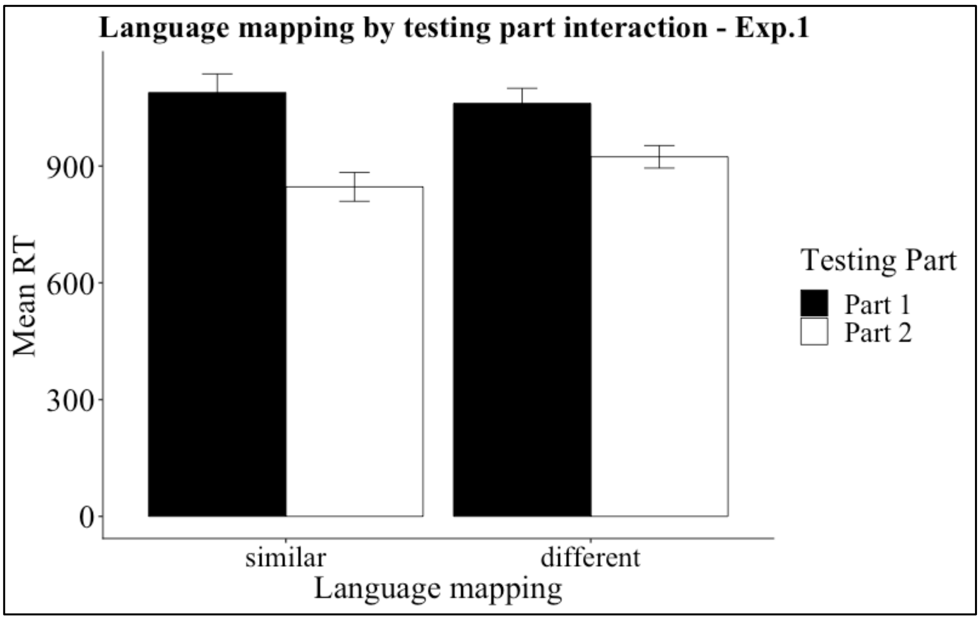


Figure 13: The interaction between language mapping and part for RT in Experiment 1

8.1.5.3 Discussion for Experiment 1

In Experiment 1, participants learned two nouns *(sách/book* and *bàn/table)* by inspecting two referential object photographs and listening to their sounds. In this way, they could, in principle, directly map the names to their referents. Next, they heard sounds of verb-noun phrases *(đọc-sách/read-book; lau-bàn/clean-table)*. The testing task was to match one of two verb sounds they heard *(đọc/read)* with one of two object photographs they inspected *(sách/book; bàn/table)*. Participants mostly selected the verb-associated object photograph (*sách/book*) in testing, and accuracy was not modulated by the absence or the presence of event photographs.

It is possible that the testing task was too easy for the participants (accuracy was close to 100%), thereby precluding the facilitative effects of the event photographs and language mapping differences (instantiating positive vs. negative transfer): Participants did not learn better when the L1 verb and the L2 verb were similar (*đọc/read*) than when they were different (*mặc, đeo, đi/wear*), and results were similar with (vs. without) the event photos. Results further revealed an unpredicted main effect of the part, which means participants overall responded more quickly in the second testing part compared with the first one (perhaps because testing in Part 2 re-used the materials from the first testing part).

We noticed some shortcomings in Experiment 1:

We chose 16 verbs to create 32 phrases. Therefore, one verb related to two phrases, and each verb sound was played three times per part (e.g., the sound of the verb *"lau/clean"* was played three times for the sounds of the phrases *"**lau** bàn," "**lau** nhà"* and the verb sound *(lau/clean)* in testing part . In addition, we used the same photograph materials in the second testing part, meaning that it was not challenging for participants to generalize to new materials.

In each testing part, participants had two response opportunities per trial. If their first response was incorrect, they received the feedback "incorrect" or "no response" (for timeouts) on the computer screen. Their post-feedback response (only) was recorded and it approximated 100% accuracy in all learning conditions of Experiment 1 (post-feedback responses).

We did not record all of the participants' responses (e.g., if participants had given an incorrect first response, and then responded again, only their second response was recorded). Thus, we did not record their performance on their first attempt at a response; this could mask the effects of the manipulated factors.

In the post-experiment debriefing, participants reported that the testing task was not difficult, and that they used strategies. For instance, after the first few items, they focused on extracting and remembering sounds of verbs when the audio sounds of the verb-noun phrases were played per trial. Also, they knew that they would have a second chance to respond after feedback.

Because of these considerations, we adapted the materials, the testing design, and the experimental procedure in Experiment 2. The aim of Experiment 2 was to examine whether the effects of events and language mappings could be shown in a revised testing task. Given the effects of part that exploratory analyses had corroborated, Experiment 2 would assess whether learning success improved across the experiment.

8.2 Experiment 2

8.2.1 Participants

In Experiment 2, we tested a further 32 German native speakers (ages 18–31). All participants had German as their only L1, no knowledge of Vietnamese, and had not learned a second language before age 6.

8.2.2 Materials

In Experiment 2, we changed the language materials used in Experiment 1. Sixteen Vietnamese verbs corresponded to 32 verb-noun phrases (two phrases related to a verb in Experiment 1). To avoid repeating the sounds of verbs during learning, we added 16 new Vietnamese verbs in Experiment 2. That meant that in Experiment 2, 32 verbs in 32 verb-noun phrases were learned. New event photographs were added, and new sounds were recorded. Table 9 presents an overview of the differences between Experiments 1 and 2.

Table 9: The difference in learning regime between Experiment 1 and Experiment 2, illustrated in one item set

Language mapping	Experiment 1	Experiment 2
Similar mapping	*lau-bàn* and *lau-nhà* *(clean-table) and (clean-floor)*	*lau-bàn* and ***quét**-nhà* *(clean-table) and (sweep-floor)*
Different mapping	*mặc-áo* and *mặc-váy* *(wear-shirt)* and *(wear-shirt)*	*mặc-áo* and ***giặt**-váy* *(wear-shirt)* and *(wash-shirt)*

8.2.3 Experimental Design

The same 2 × 2 design, as in Experiment 1, was used in Experiment 2. The difference concerning learning in Experiment 2 was that twice as many verbs in the Vietnamese verb-noun phrases were learned in Experiment 1.

Learning in Part 1 of Experiment 2 was the same as in Part 1 of Experiment 1. Moreover, participants re-learned all items in Part 2 of Experiment 2 without testing. In Experiment 2, testing differed from Experiment 1 (see Figure 4). Immediate testing in Part 1 also took place after each learned pair/trial. However, instead of depicting only objects in a testing trial (as in Experiment 1), in Experiment 2, event pairs (e.g., a person reading books and a person choosing/looking for books in a book store) were depicted next to each other. Participants listened to a full verb-noun phrase (e.g., *read-book*) instead of only a verb (e.g., *read*) as in Experiment 1 (see Figure 1). Furthermore, re-learning of all 16 object pairs (Part 2) was followed by a delayed testing part (Part 3). In this testing part, participants were tested on 16 items with new event depictions (see Figure 14).

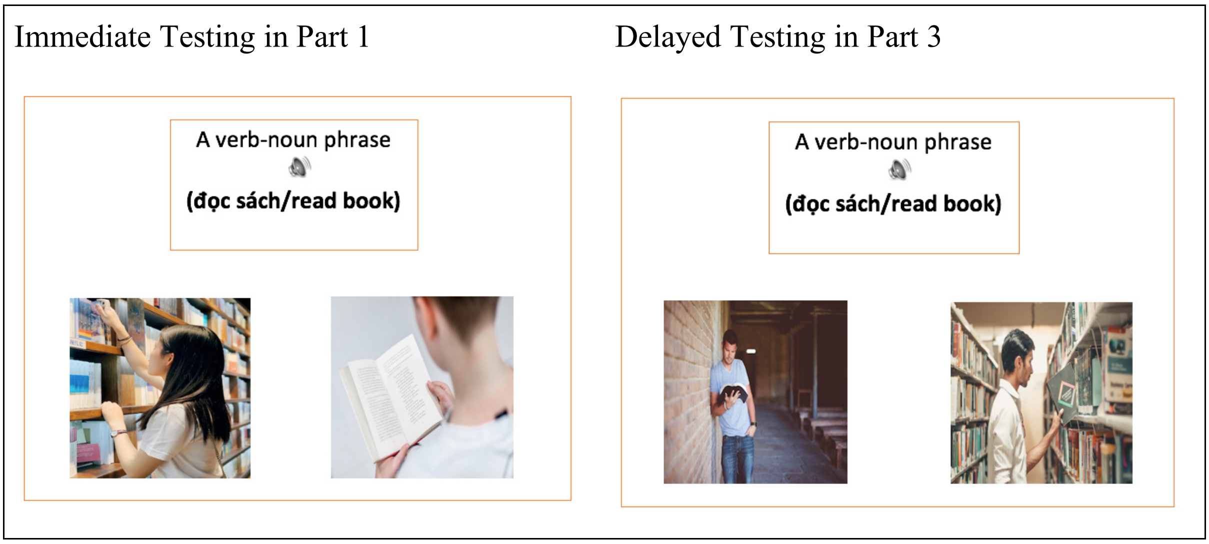


Figure 14: Examples of testing display in Part 1 and Part 3 of Experiment 2

To summarize, the main differences in testing between the two experiments were:

The testing task in Experiment 1 was to match a verb sound with one of two presented object photographs (to complete a verb-noun phrase). That task was changed in Experiment 2 in which we tested a full phrase in an event photograph selection task.

In Experiment 2, participants had only one chance to respond (no feedback was given).

There were also two testing parts in each experiment. The first testing part of the two experiments was an immediate post-trial test (with feedback in Experiment 1; without feedback in Experiment 2). For the other testing part in Experiment 2 (unlike for Experiment 1), after a separate second learning stage, all events were tested using new photographs (delayed testing).

Compare Figures 1 and 2, together with Figures 4 and 5, for the differences between Experiment 1 and Experiment 2.

8.2.4 Procedure and Predictions

The procedure for Experiment 2 included the following steps. Figure 15 illustrates step 3 to step 7.

1. Reading the experimental information and filling out the required forms
2. Completing a short practice run
3. Performing the first part of the experiment (learning and post-trial testing, Figure 5)
4. A short break
5. Performing the second part of the experiment (re-learning all items in a
6. pseudo-randomized order; no testing)
7. A short break
8. Performing the third part of the experiment (testing only)
9. Participants answered post-experimental questions and were debriefed.

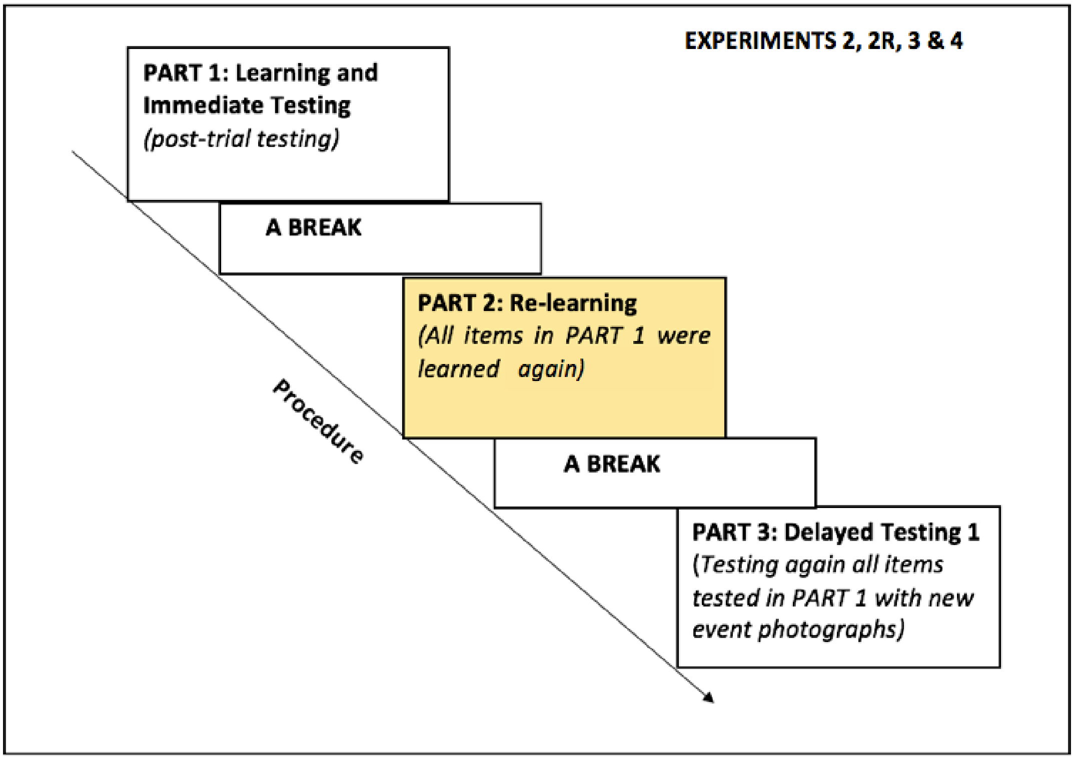


Figure 15: The procedure of Experiment 2 (also in Experiments 2R, 3 and 4)

Based on the above changes in Experiment 2, the revised predictions were:

a. If events are useful in learning, then participants should give more accurate and faster total responses via a button press in the "event-present" compared with the "event-absent" condition.

b. If verb mapping makes a difference, then participants should respond more slowly in different than in similar verb-noun phrase mappings (between Vietnamese and German/English).

c. If there is a strong effect of the testing part, then participants should give more accurate and faster responses via a keyboard key (i.e., P or Q) in the immediate testing in Part 1 compared with the delayed testing in Part 3. Both the delay of testing and the use of new event photographs (still referring to the verb-noun phrases learned in Part 1 and Part 2) make testing in Part 3 more challenging.

8.2.5 Results and Discussion

Using the identical models of the same analysis methods as in Experiment 1, we obtained the following results.

8.2.5.1 Accuracy Results

When the testing part was involved in the analysis model as a factor (see Table 10), a significant effect of event presence emerged in both testing parts ($p < .001$). That means that beginners learned L2 phrases much more successfully with event photographs than without them. Also, a significant effect of the testing part ($p < .001$) on learners' accuracy emerged, with higher response accuracy in the immediate testing than in the delayed testing.

Table 10: Results from generalized linear mixed model fit by maximum likelihood in Experiment 2

Term	Estimate	SE	z value	p
Intercept	1.101	0.172	6.4	<0.001***
Event presence	1.232	0.107	11.5	<0.001***

Language mapping	-0.186	0.167	-1.11	0.27
Testing Part	0.545	0.099	5.5	<0.001***
Event presence*Testing Part	0.528	0.099	5.33	<0.001***
Language mapping*Testing Part	0.084	0.099	0.85	0.4
Event presence*Language mapping	-0.011	0.104	-0.11	0.92
Event presence*Language mapping* Testing Part	0.013	0.099	1.13	0.9

The interaction of visual context by the testing part (see Figure 16) was significant ($p < .001$). When the event photographs were present, learners' accuracy was much higher (95.7%) in the immediate testing compared with the delayed testing (74.21%). In the event-absent conditions, their accuracy was approximately equally low in both testing parts (i.e., immediate testing: 48.05% and delayed testing: 47.26%).

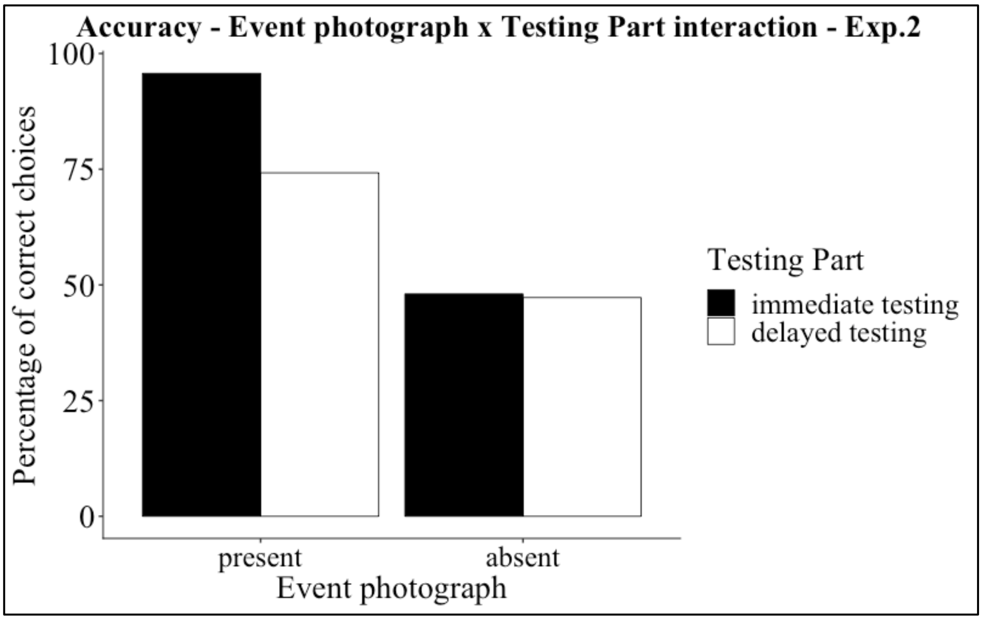


Figure 16: The interaction between event photograph and testing part in accuracy data in Experiment 2

The analyses without testing part in the model corroborated a significant visual context effect ($p < .001$, higher accuracy in testing when events were present than absent, see Figure 17).

Across both parts, accuracy was 84.96% when events were present in training (total: 512 trials). In contrast, accuracy was only 47.66% when no events were present (total: 512 trials).

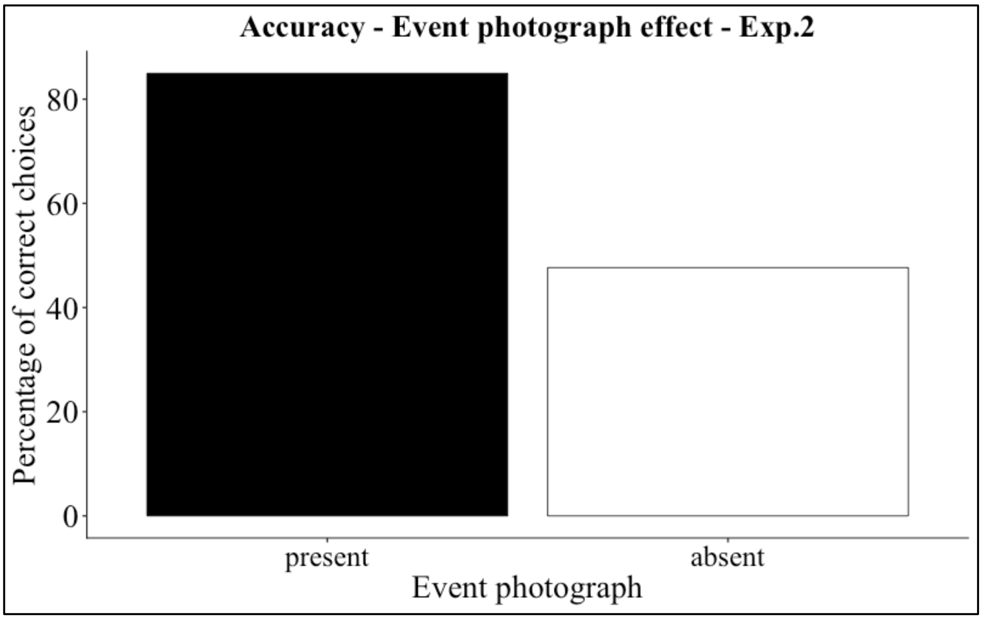


Figure 17: The effect of event photograph on the accuracy in Experiment 2

8.2.5.2 Reaction Time Results

Including the testing part as a factor in the analysis model (see Table 11), results revealed a main effect of event presence (SE = 0.012, = 11.5, p < 0.001) with faster responses for the event-present than absent condition. A significant effect of the testing part suggested that participants responded faster in immediate testing than in delayed testing (SE = 0.021, = 1.25, p < 0.001).

Table 11: Results from linear mixed effects model for reaction time in Experiment 2

Term	**Estimate**	**SE**	**df**	**|t|**	**p**
Intercept	7.378	0.045	38	163.85	<0.001***
Event presence	-0.140	0.012	589	11.5	<0.001***
Language mapping	0.024	0.019	32	1.25	0.22
Testing Part	-0.078	0.021	33	3.63	<0.001***

Event presence* Testing Part	-0.104	0.011	567	9.03	<0.001***
Language mapping* Testing Part	0.030	0.011	561	2.66	0.008**
Event presence*Language mapping	0.029	0.012	591	2.35	0.02**
Event presence*Language mapping* Testing Part	0.015	0.011	564	1.31	0.19

Another significant interaction between event presence and testing part (SE = 0.011, = 9.03, p < 0.001) confirmed that in event-present conditions, participants responded more quickly in immediate testing than in delayed testing. However, in event-absent conditions, they responded more slowly in the immediate testing than in delayed testing (see Figure 18a).

Language mapping also significantly interacted with the testing part (SE = 0.011, = 2.66, p < 0.01) in an unexpected direction. In immediate testing, participants responded faster in different than in similar language mapping conditions. By contrast, they took an approximately equally time to respond for the two language mapping conditions in delayed testing (see Figure 18b).

We also found a significant interaction between event presence and language mapping (SE = 0.012, = 2.35, p = 0.02). Learners were as fast in similar language mapping conditions as in different ones when events were present. However, they were slower in similar mapping conditions than in different mapping conditions when events were absent (see Figure 18c). In other words, with/without supportive events, learning L2 phrasal vocabulary in L1–L2 different verb mapping conditions was not more difficult for beginners than in L1–L2 similar verb mapping conditions.

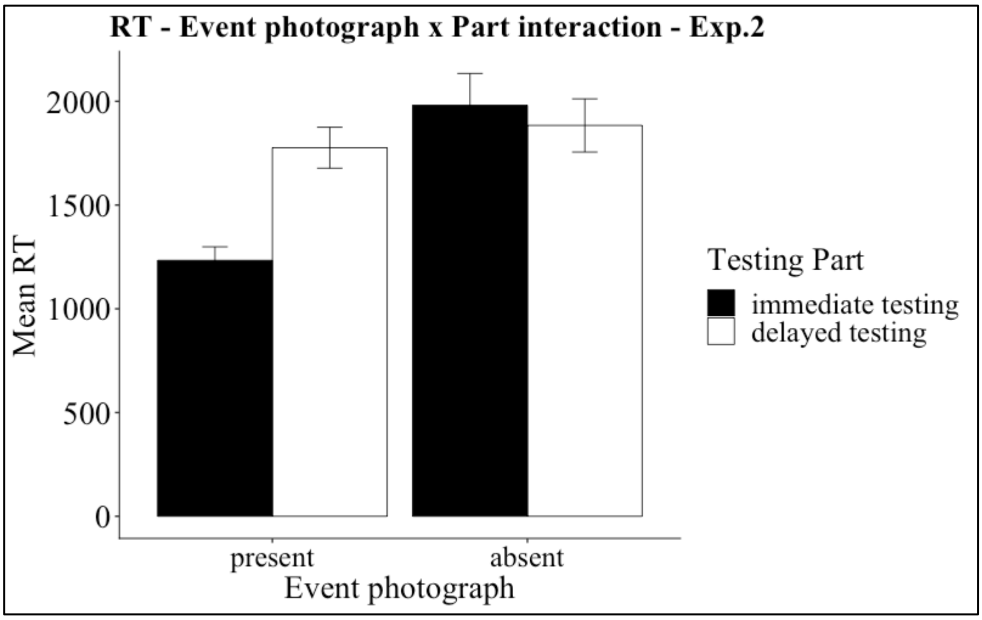


Figure 18a. : The interaction between event photograph and testing part on the reaction time in Experiment 2 (error bars represent 95% CI)

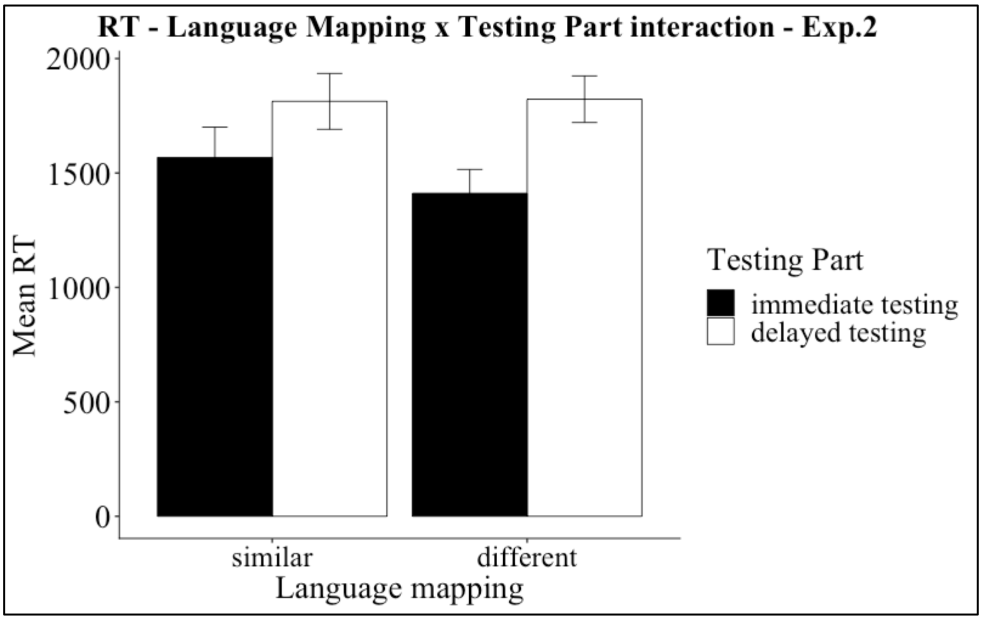


Figure 18b.: The interaction between language mapping and testing part on the reaction time in Experiment 2 (error bars represent 95% CI)

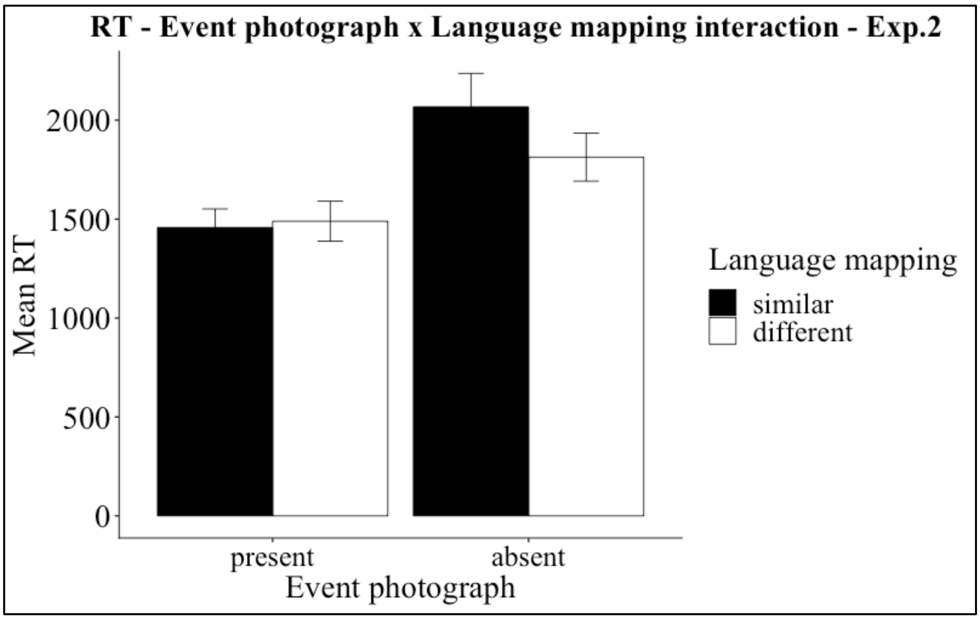


Figure 18c. : The interaction between event photograph and language mapping on the reaction time in Experiment 2 (error bars represent 95% CI)

We also looked at the significant interaction in each testing part (see Figures 19a and 19b). The same pattern for event-absent conditions was investigated in both testing parts indicating the slow speed of responding for similar language mapping items compared to different ones. For event-present conditions, the directions of the interaction were opposite between two testing parts. Participants were slower to respond to similar than different language mapping items in immediate testing. However, they were faster for present-similar items than all three other kinds of items in delayed testing.

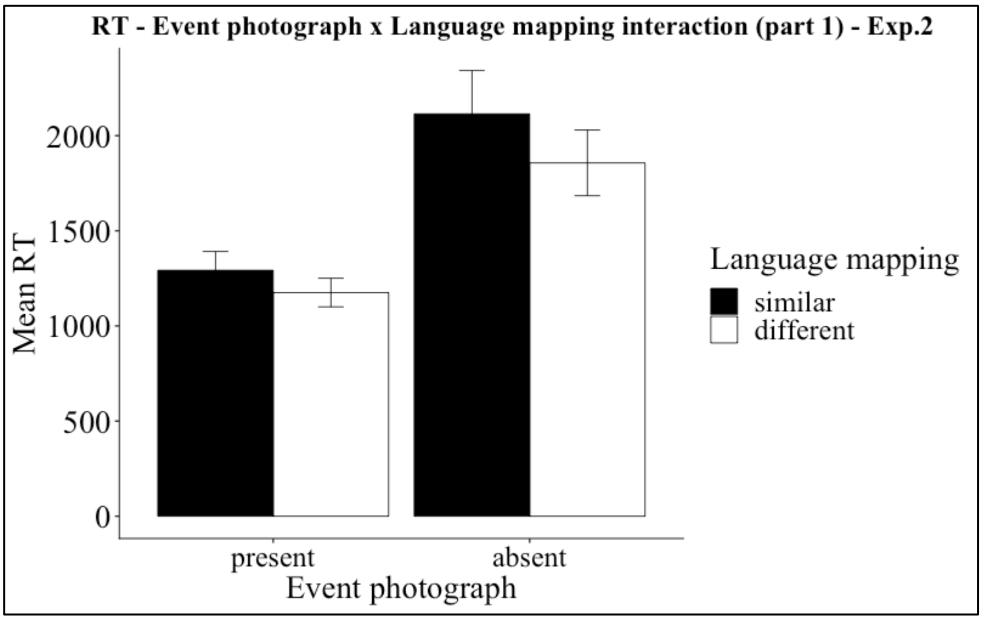


Figure 19a: The interaction between event photograph and language mapping on the reaction time in Part 1/immediate testing of Experiment 2 (error bars represent 95% CI)

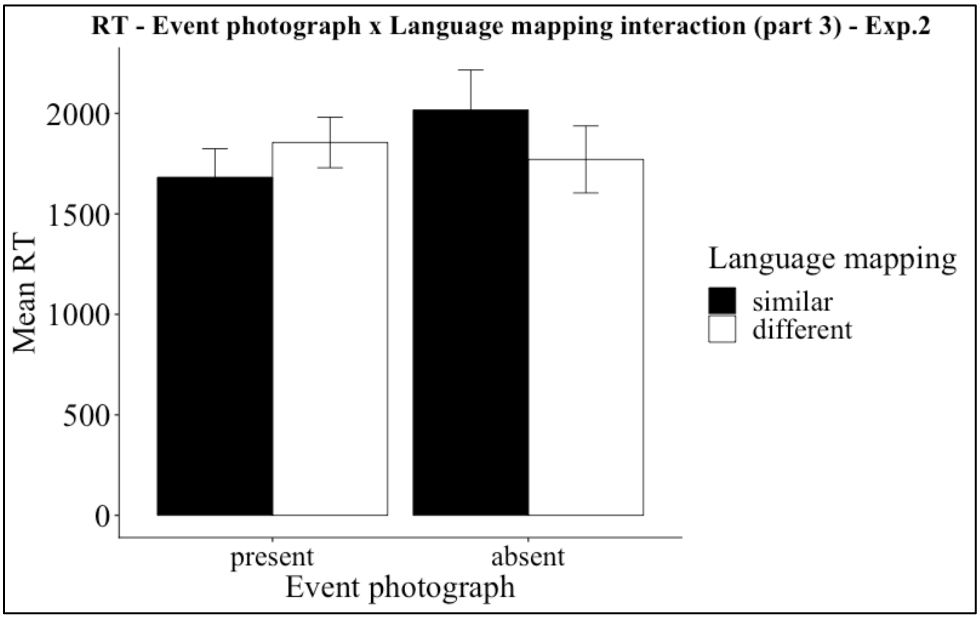


Figure 19b: The interaction between event photograph and language mapping on the reaction time in Part 3/delayed testing of Experiment 2 (error bars represent 95% CI)

Without testing part in the analysis, we obtained a significant effect of event presence (see Figure 20) on the reaction time (SE = 0.01432, t(576) = -10.585, p < 0.01) with faster responses for the event-present items than event-absent items.

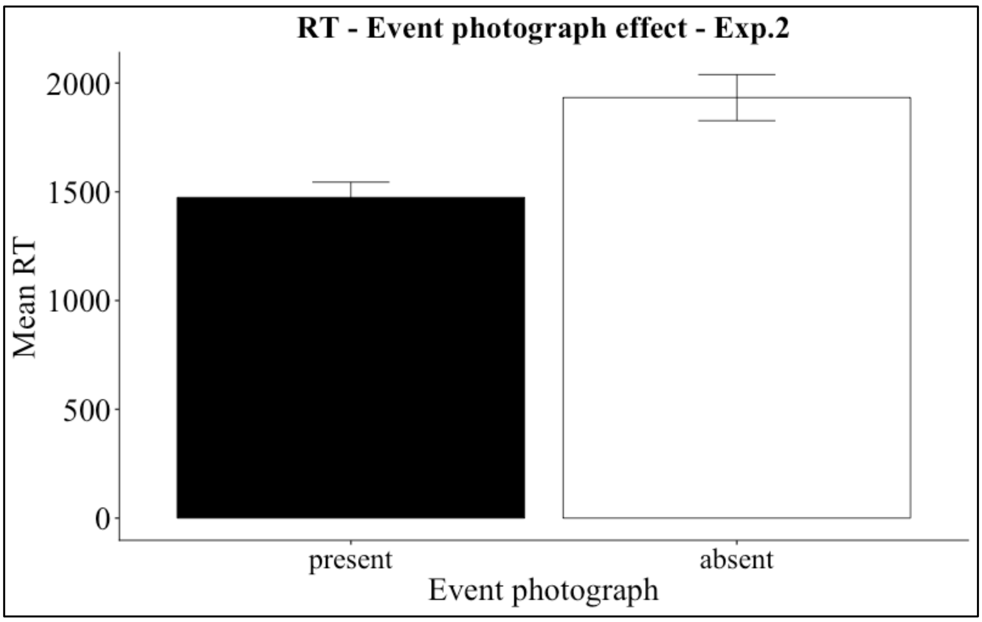


Figure 20: The effect of event photograph on the reaction time in Experiment 2 (error bars represent 95% CI)

The analyses also revealed a significant event presence by language mapping interaction (SE = 0.014, = 2.023, p = 0.043). Figure 21 shows that when events were absent, participants took much longer to respond accurately in similar, rather than in different, language mapping contexts. In contrast, when events were present, participants' responses were a little faster in similar language mapping conditions than in different ones.

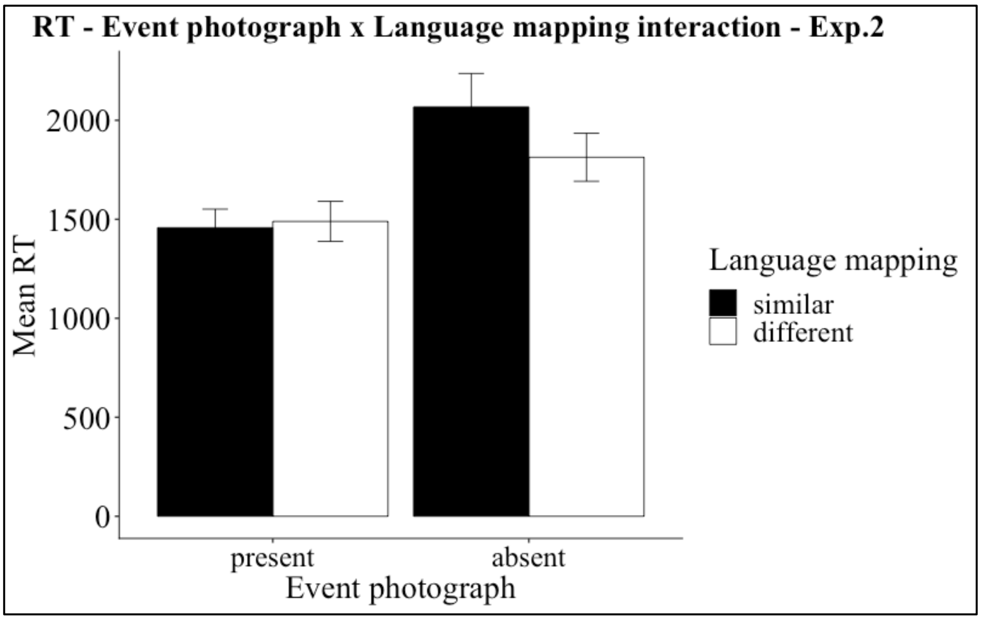


Figure 21: The interaction between event photograph and language mapping on the reaction time in Experiment 2 (error bars represent 95% CI)

Because we analyzed reaction times for correct choices only, we were concerned that imbalances in the number of correct responses among learning conditions might have affected the results. For that reason, we also created data groups for the four learning conditions (present-similar, present-different, absent-similar, absent-different), balanced per condition per part (e.g., 54 responses for each condition in Part 1 based on the smallest number of correct choices in a condition). Results from these post hoc analyses corroborated a significant effect of event presence on reaction time (SE = 0.01727, t(370) = -9.123, $p < 0.001$) and a main effect of testing part on reaction time (SE = 0.02147, t(30.9) = -4.230, $p < 0.0002$). When the testing part was included in the analysis model, the event depiction by language mapping interaction and the language mapping by testing part interaction were statistically significant with the same patterns as the results obtained in the full reaction time data set.

8.2.5.3 Discussion for Experiment 2

We recorded all the results of Experiment 2 in Table 12. With the same learning regime as in the first experiment, we changed the testing in Experiment 2. Instead of choosing one of two object photographs referring to a noun, after hearing a verb sound in Experiment 1, participants in Experiment 2 had to select one of two event photographs after listening to a verb-noun phrase. That meant that participants in Experiment 2 were not tested by matching a verb with a noun to complete a phrase. We used a binary forced-choice decision task to examine whether the participants might have learned and remembered every verb-noun phrase. Also, participants had only one chance to respond, and they received no feedback. Results showed that participants were significantly more accurate when they viewed supportive event photographs during L2 learning compared to when learning happened without event photographs. With the testing design of Experiment 2 (e.g., every full phrase tested), event photographs played a key role in L2 learning.

Table 12: Main results in Experiment 2

Exp.2	**Accuracy**	**Reaction time**
Without Part	event presence effect	event presence effect event presence by language mapping interaction
With Part	event presence effect testing part effect event presence by testing part interaction	event presence effect testing part effect event presence by testing part interaction language mapping by testing part interaction event presence by language mapping interaction

We also observed a significant effect of event photographs on participants' responses in the delayed testing with new photograph materials. Participants could remember the meanings of phrases (specific actions) via depicted event photographs. In the delayed testing, they were able to generalize which one of two new event photographs was correct. By contrast, participants did not have enough information about verb-noun phrases to identify the meaning of phrases when the verb sounds were new, and the events had been absent during training. Therefore, participants' choices in testing were made based on chance. Results in both testing parts showed that participants' accuracy was around 50% for the event-absent conditions.

It is noteworthy that when learning with event photographs, participants responded much faster in the immediate testing compared with the delayed testing. That might be because the photograph materials were new in the delayed testing part. Hence, participants took more time to identify the phrase sound and the event they had learned in the two previous sections.

We found no main effect of language mapping when beginners learned L2 phrases in the L2 learning design. The robustly significant effect of event depictions on L2 learning might cause the influence of the experimental factors—language mapping was not precise. Also, interactions such as language mapping by testing part and event presence by language mapping had opposite directions from the ones predicted by the L1–L2 transfer theory. We do not know whether these were chance interactions or the results of individual learners' differences.

In the next section, we discuss more differences of testing in Experiment 1 and Experiment 2 as well as differences in L2 vocabulary learning success and the effects of experimental factors on L2 learning.

8.3 Second Language Vocabulary Testing and its Effects on Second Language Vocabulary Learning Success in Experiments 1 and 2

In Experiments 1 and 2, we examined how nonlinguistic visual context and language transfer modulate L2 learning in beginners. As predicted, the participants had higher accuracy and faster responses in event-present than in event-absent learning conditions, but only in a suitable learning and testing environment (Experiment 2). In Experiment 2, one of two learned phrases (event pictures, e.g., *read-book* or *clean-table*) was directly tested along with a new event picture in a testing trial (e.g., *choose-book*) related to the noun learned (the object picture, e.g., *book*). Before testing, inspecting the event pictures improved participants' reaction times and accuracy in the binary forced-choice event selection task. Event presence in training enabled participants to infer differences between making tea and drinking tea or reading a book and taking a book. Otherwise (in the event-absent condition), they could only guess which of two event photographs was the referent of the verb-noun phrase. These results are in line with previous findings (Koehne et al., 2015; Smith et al., 2008; Yu et al., 2007) on learning or processing language (e.g., words and phrases) in visual contexts with the effective supplements of object picture-audio pairs and event picture-audio pairs to learn nouns and phrases.

Unlike Experiment 2, the results in Experiment 1 indicated no differences in L2 learning success (measured by accuracy and reaction time) when participants were trained in event-present or event-absent conditions. We think that the divergence of the results is due to the different forced-choice selections in the two experiments. In Experiment 1, learners had to match exactly one of two presented object pictures (e.g., *book* or *table*) with a verb audio sound (e.g., *read*) to complete an exact verb-noun phrase (e.g., *read-book*). The key to L2 phrase learning in Experiment 1 was to extract the verb sounds and remember the associated object pictures. We think that because of the focus of this task (learning associations between verbs and nouns), seeing the event photographs in learning was not advantageous since verb-noun phrase sounds were displayed in all learning conditions. By contrast, in Experiment 2, the choice was between two events (*make-tea* vs. *drink-tea*). Inferring event meaning was only possible in the event-present conditions of Experiment 2.

The originally experimental purpose was to investigate how beginners learned L2 phrases in four different learning conditions. Therefore, (i) testing by choosing one of two referents (*a person reading a book* vs. *a person looking for a book*) pairing with a spoken phrase (e.g., *read-book*) was appropriate. In Experiment 1, (ii) testing by matching a spoken verb (e.g., *read*) with one of two different objects (e.g., *book* vs. *table)* to complete a phrase (e.g., *read-book)* did not fit with the planned L2 learning purpose. Although both (i) and (ii) are binary forced-choice tasks, (i) is a more suitable test than (ii) for the following reasons:

In (ii), participants could still respond accurately in testing without seeing event photographs illustrating phrases before, so the role of event presence could not be examined.
In (i), the task was a recognition task. By hearing a spoken phrase first, the participants had time to recognize which sound they had listened to before. Then, inspecting two event photographs/two actions related to an object, they realize which event (if they had seen it before) pairs with the spoken phrase. If they made correct choices in event-present conditions, they had learned verb-noun phrases via the useful support of event depictions.

9. Influences of Individual Differences on Second Language Vocabulary Learning Success

Following the findings of Experiment 2, we wanted to replicate the experiment for another adult group with the same age range (18–31 years) and other older adult groups (32–65 years). That was why we carried out a set of three experiments (Experiment 2R—"R" means replication, Experiment 3 and Experiment 4) at the same time. These experiments were made to investigate:

- whether L2 vocabulary learning success in a visual context would be replicated with the same effects as in Experiment 2
- whether there was a main effect of language mapping in the L2 learning situation
- how personal differences (i.e., age-related differences and cognitive ability) affected adults' L2 vocabulary learning.

9.1 Age Ranges for Adulthood

Adulthood begins at around 20 years old and has three distinct stages of development of psychology: early adulthood (20–30 years), mature adulthood or middle age (30–65 years), and old age (65+ years). Each stage brings its own set of rewards and challenges.

The stages of adulthood examined in the present study include young adults (ages 18–31), early middle-aged adults (ages 32–45), and late-middle-aged adults (ages 45–65). We also wanted to test older adults (ages 65 and older), but we could not realize this goal because it was challenging to recruit participants.

9.2 Young Adults (Experiment 2R)

9.2.1 Methods

We ran Experiment 2R with the same materials, experimental design, and procedure as in Experiment 2. We tested a further 32 participants (18–31 years old, mean age: 23.6, Male: 14).

Before the main learning experiment, participants took a cognitive test comprising five subtests of the WAIS-IV test in the German version. Table 13 describes the five subsets of the cognitive test. Four specific domains of learners' intelligence (i.e., verbal comprehension, perceptual reasoning, working memory, and processing speed) were examined.

Table 13: A description of the cognitive test - Wechsler Adult Intelligence Scale - (WAIS-IV)

<u>Part A: General cognitive abilities</u> (fluid intelligence—"the capacity to reason and solve novel problems, independent of any knowledge from the past")	**<u>Part B: Linguistic abilities</u>** (crystallized intelligence—"the ability to use the skills, knowledge, and experiences relying on accessing information from the long-term memory")
Subtest 1: Bilder ergänzen (WIE 1)/Picture completion => Perceptual Reasoning Scale Subtest 2: Zahlen-Symbol-Test (WIE 3)/Digit Symbol => Processing Speed Scale Subtest 3: Zahlen nachsprechen (WIE 8) /Digit Span => Working Memory Scale	Subtest 4: Wortgenerierungsaufgabe/Animal and Word naming Subtest 5: Gemeinsamkeiten finden (WIE 4)/Similarities => Verbal Comprehension Scale

Figure 22 shows all the main parts when the learners were in the laboratory participating in the L2 learning experiments (Experiment 2R, 3, and 4). First, they read the information for participants before signing the consent form and filling out the questionnaire. They then took a cognitive test for around 20 minutes before participating in the main L2 vocabulary learning experiment. Last, they answered some post-experimental questions before signing and receiving €11 as a reward for their participation. The experimental procedure was consistent with Experiment 2 (see Figure 15).

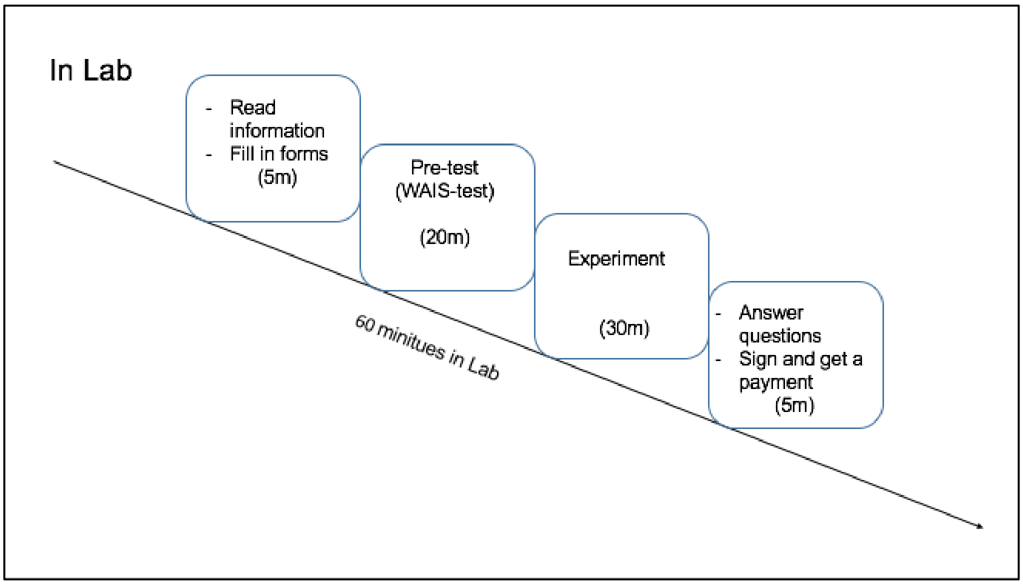


Figure 22: The full procedure in the Lab for experiments 2R, 2N, 3, and 4

9.2.2 Predictions

Predictions for the cognitive test (WAIS scores): Participants with higher scores in the cognitive test would do better in the L2 learning experiment (higher accuracy and shorter reaction time) than participants with comparatively lower scores.

Predictions for main results of L2 vocabulary learning success: Predictions for the results of Experiment 2R were based on the significant main effects of event depiction and testing part on L2 accuracy and reaction time in Experiment 2. Although no main effect of language mapping had emerged in Experiment 2, we still predicted its effect in Experiment 2R. This means that if verb mapping influences language learning success, then participants should respond more slowly in different verb mappings than in the same verb mappings. We also found some significant interactions (event depiction by testing part; event depiction by language mapping; language mapping by testing part) in Experiment 2. Therefore, we used these results to predict Experiment 2R:

(a) If the event effects differ depending on language mapping, then we should see a significant interaction between event photograph depictions and verb mappings in their influence on participants' learning. Participants in the event-present learning condition would take shorter reaction time and get more correct answers in similar verb mapping conditions than in different verb mapping conditions. By contrast, participants in event-absent learning conditions would have approximately equal learning speed and accuracy in both similar and different verb mapping conditions.

(b) If the event effects differ depending on the testing part, then we should see a significant interaction in participants' L2 learning between visual context (event photograph depiction) and testing part (1 and 3). That means that participants who had inspected event photographs performed their tasks better in Part 1 (shorter reaction time, more correct answers) than in Part 3. However, participants who had not inspected event photographs would perform their tasks equally well in Part 1 and Part 3 (approximately 50% of correct answers and similar reaction times).

9.2.3 Results and Discussion

With the same analysis methods and settings as in experiments 1 and 2 (see 8.1.4), we obtained the following results.

9.2.3.1 Accuracy Results

When testing part was included as a factor in the analysis model, three results were replicated (see Table 14), including a main effect of event presence ($p < .001$) in both testing parts, a main effect of testing part ($p < .001$), and a significant event presence when testing part interaction ($p < .001$, see Figure 23).

Table 14: Results from generalized linear mixed model fit by maximum likelihood for accuracy in Experiment 2R

Term	Estimate	SE	z value	p
Intercept	1.206	0.176	6.84	<0.001***
Event presence	1.238	0.110	11.24	<0.001***
Language mapping	-0.076	0.169	-0.45	0.65
Testing Part	0.531	0.103	5.13	<0.001***
Event presence* Testing Part	0.522	0.104	5.04	<0.001***
Language mapping* Testing Part	0.067	0.103	0.65	0.52
Event presence*Language mapping	0.213	0.108	1.98	0.049*

Event presence*Language mapping* Testing Part	0.076	0.103	0.73	0.46

Present event photographs were helpful for beginners in L2 phrase learning: The learners responded accurately at 96.1% in immediate testing and 77.34% in delayed testing, while their correct choices were 50.4% and 50% in immediate and delayed testing respectively, when event photographs were absent in learning phase.

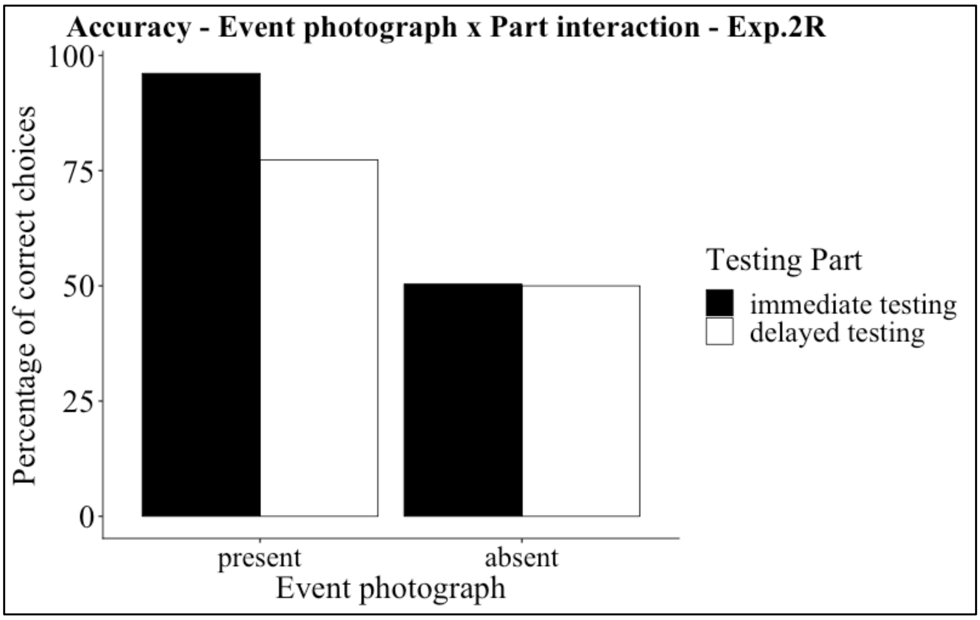


Figure 23: The interaction between event photograph and testing part on the accuracy in Experiment 2R

The only new result (compared with Experiment 2) was a significant interaction between event presence and language mapping (p = 0.049, see Figure 24.). When event photographs appeared in learning, participants, in total, equally successfully learned L2 phrases both in similar language mapping items (86.72% correct responses) and different language mapping items (86.72% correct responses). When event photographs were absent in the learning phase, they were much more accurate for different language mapping items (55.86% correct responses) than for similar language mapping items (44.53% correct choices).

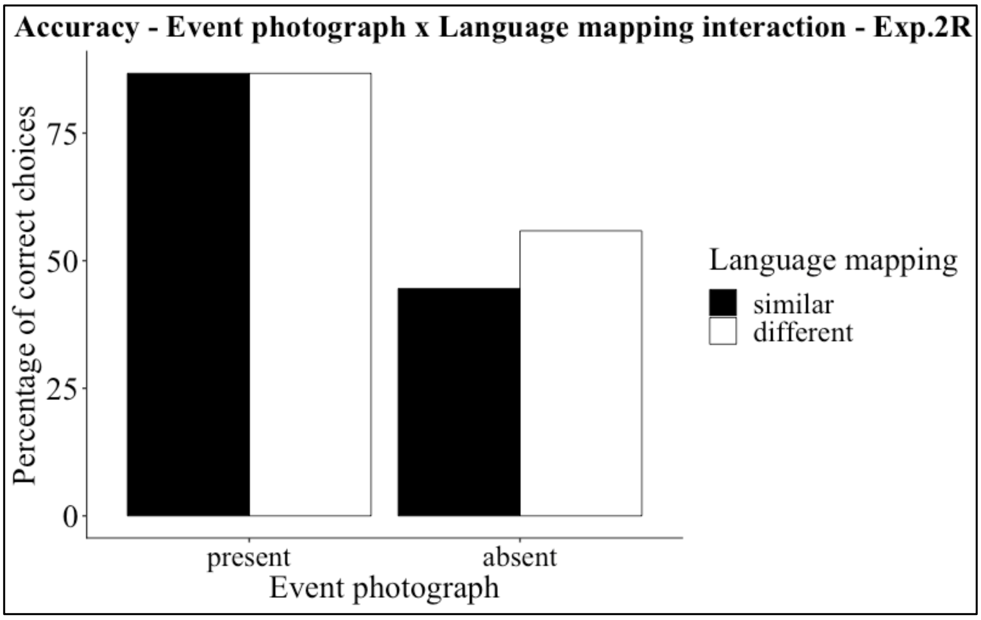


Figure 24: The interaction between event photograph and language mapping on the accuracy in Experiment 2R

Omitting the testing part, we replicated a significant effect of event depictions on participants' accuracy, as in Experiment 2 ($p < 0.001$). The participants made 86.7% correct choices in the event-present learning conditions (512 event-present trials in total). Participants were much less accurate than in the event-absent learning conditions, with 50.2% correct choices for 512 event-absent trials (see Figure 25).

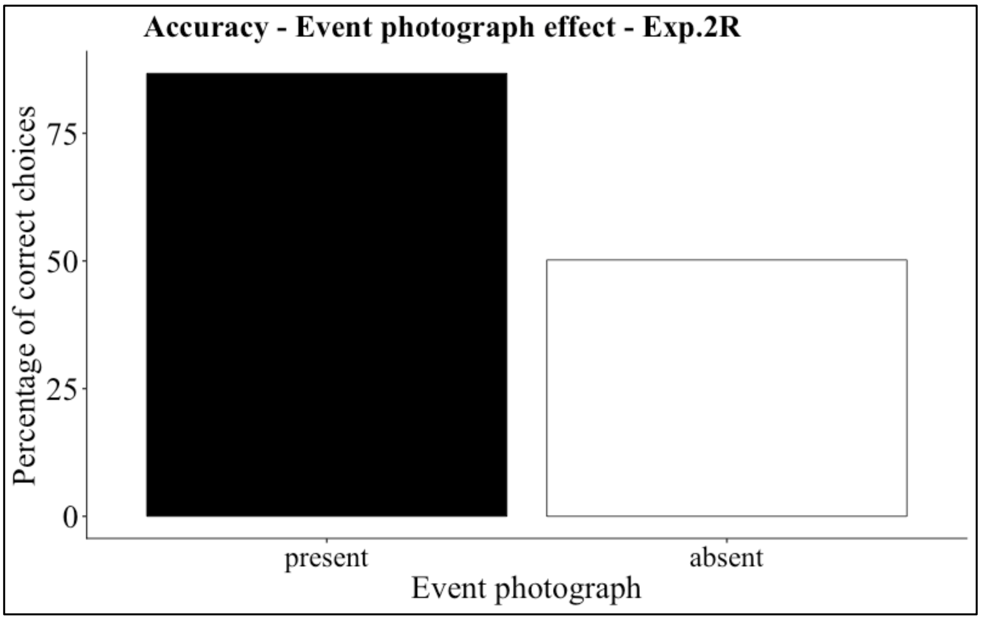


Figure 25.: The effect of event photograph on the accuracy in Experiment 2R

9.2.3.2 Reaction Time Results

Table 15 shows all results investigated from the linear mixed-effects model when testing part as a factor was included in the analysis model. As expected, the model replicated a significant effect of event presence (SE = 0.01, = 15.45, $p < 0.001$) on participants' response speed, which meant that participants responded faster in the event-present than in the event-absent conditions. Also, the analyses corroborated significantly faster reaction times in the immediate testing than in the delayed testing (SE = 0.02, = 5.82, $p < 0.001$).

Table 15: Results from linear mixed effects model for reaction time in Experiment 2R

Term	**Estimate**	**SE**	**df**	**|t|**	**p**
Intercept	7.359	0.038	41	193.96	<0.001***
Event presence	-0.181	0.011	630	15.45	<0.001***
Language mapping	0.007	0.019	30	0.39	0.7
Testing Part	-0.125	0.021	32	5.82	<0.001***
Event presence* Testing Part	-0.104	0.011	611	9.34	<0.001***
Language mapping* Testing Part	0.039	0.011	611	3.52	<0.001***
Event presence*Language mapping	0.002	0.011	631	0.24	0.81
Event presence*Language mapping* Testing Part	0.010	0.011	608	0.97	0.34

The significant event presence by testing part interaction (SE = 0.01, = 9.34, $p < 0.001$, $p < 0.001$, see Figure 26a) indicated that learners were faster in the immediate testing than in the delayed testing if events had been present in the learning phase. Their response times were approximately the same in both testing parts when events had been absent previously.

Moreover, we found an unpredicted significant two-way interaction between language mapping and testing part (see Figure 26b) in the model (SE = 0.01, = 3.52, $p < 0.001$). In the immediate testing, participants responded more slowly in similar language mapping conditions compared to different conditions. By contrast, in the delayed testing, they were slower in responding for different than for similar language mapping conditions. Unlike in Experiment 2, the interaction between event presence and language mapping in Experiment 2R was non-significant.

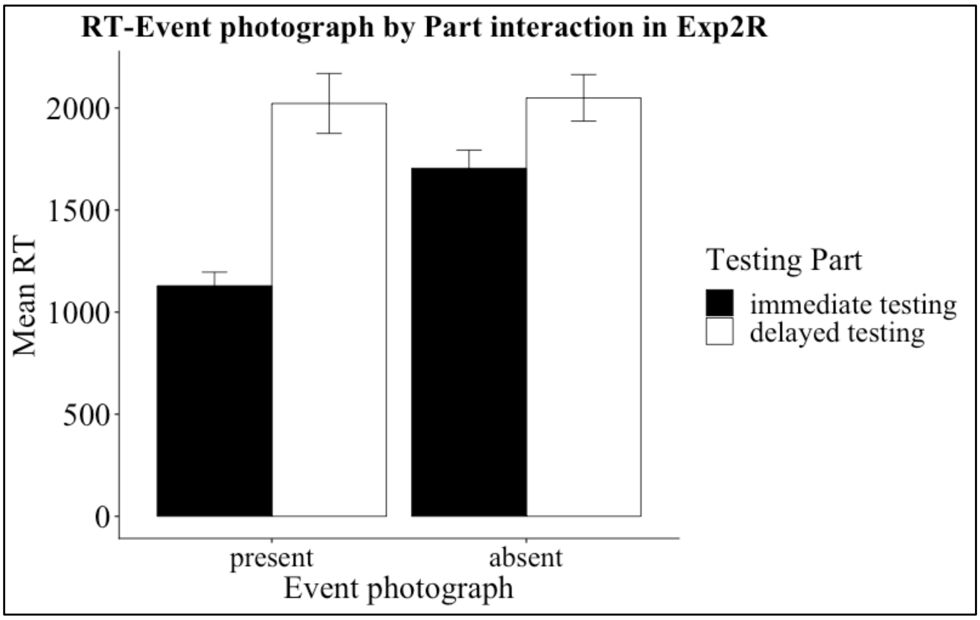


Figure 26a: The interaction between event photograph and part on the reaction time in Experiment 2R (error bars represent 95% CI)

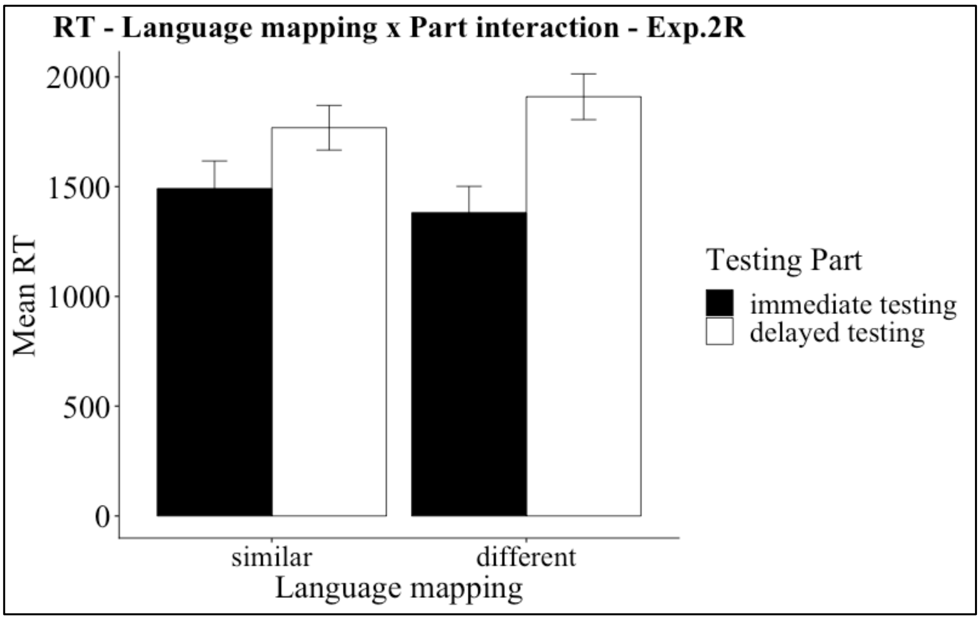


Figure 26b: The interaction between language mapping and part on the reaction time in Experiment 2R (error bars represent 95% CI)

Results without testing part corroborated a significant effect of event presence on reaction time (see Figure 27), as in Experiment 2 ($SE = 0.01$, $= 13.65$, $p < 0.001$). That result seemed to indicate that participants' response time for correct choices was much shorter in the learning phase with supportively present event photographs than without them. No further effects involving the manipulated factors were found.

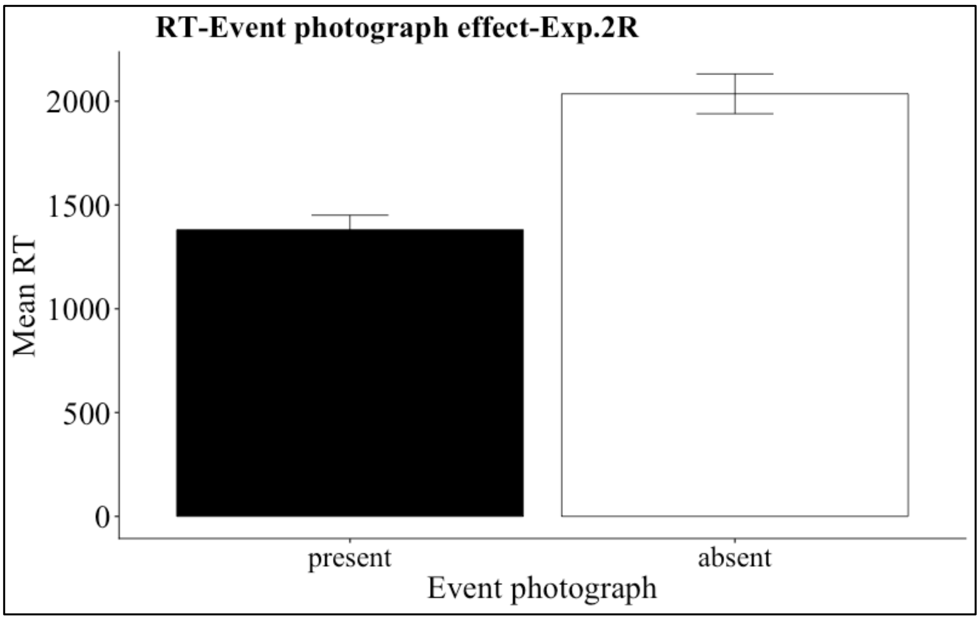


Figure 27: The effect of event photograph on the reaction time in Experiment 2R (error bars represent 95% CI)

9.2.3.3 Cognitive Test Results

Young adults' cognitive test scores are shown in Figure 28a. Participants performed best in the similarities task and worst in the digit span task. The mean for the verbal fluency task was 41.625 words generated.

To test whether accuracy results and the cognitive test scores correlate, we computed a Pearson correlation test because both variables were normally distributed. The result revealed that learners' cognitive tests and their accuracy scores in L2 vocabulary learning were not significantly correlated (a correlation coefficient of -.18, p-value > .05). Figure 28b illustrates no correlation, but it suggests that the lower the cognitive scores, the higher the accuracy scores.

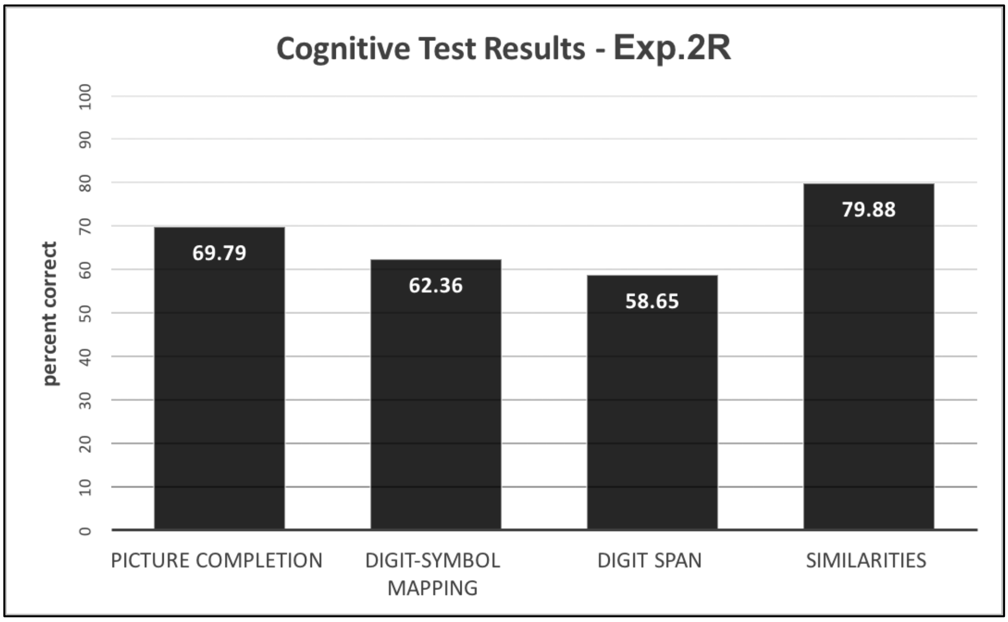


Figure 28a: Cognitive test scores for the WAIS test (exp. 2R). The y-axis displays the percentage of correct answers averaged across participants. The percentages are shown in the center of each bar. Note that the verbal fluency test scores are not depicted. Since the task was free naming, there is no upper limit that can be reached.

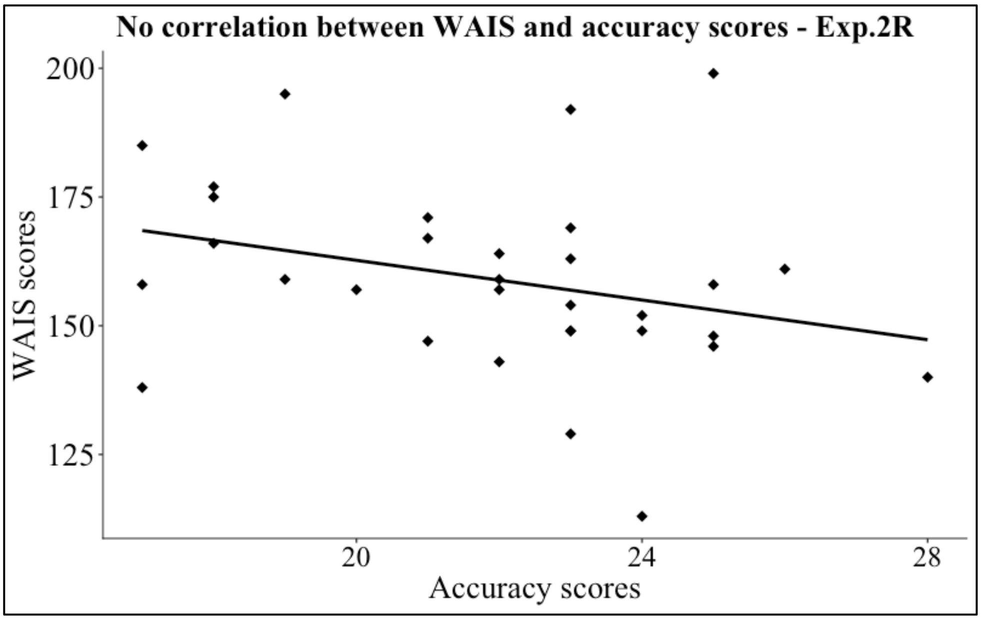


Figure 28b: No correlation between WAIS and accuracy scores (exp. 2R). WAIS scores (including the verbal fluency task scores) are displayed on the y-axis, accuracy scores are displayed on the x-axis. The trend line shows the line of best fit.

The WAIS scores of the participants and their response times in L2 learning are significantly correlated in Experiment 2R, with a correlation coefficient of -.41 and a p-value $< .02$. Figure 28c depicts the correlation, which shows a clear trend that the higher the cognitive scores, the faster the response speed for correct choices.

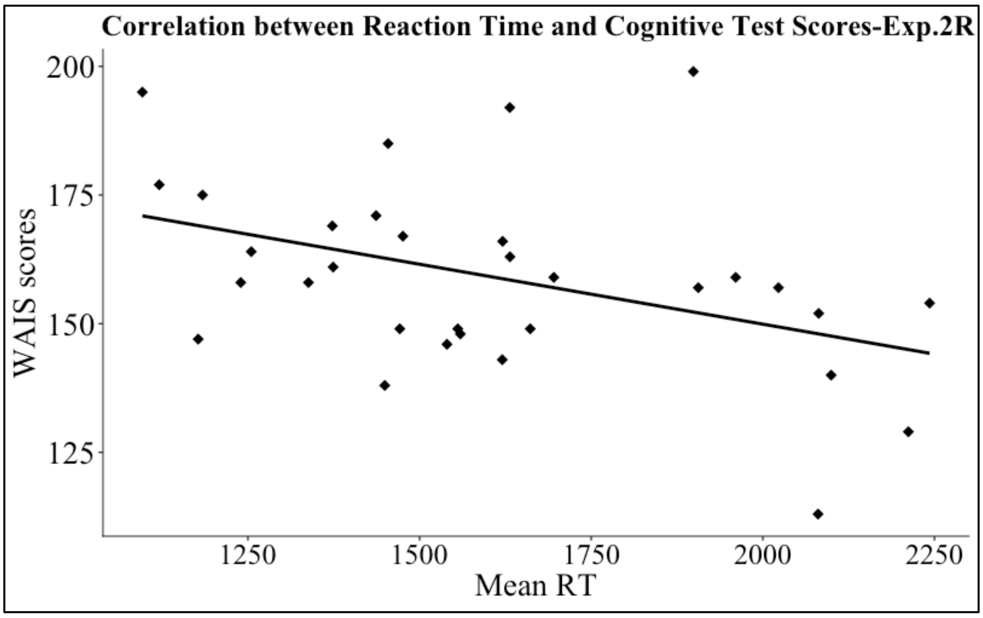


Figure 28c: A significant correlation between WAIS and reaction time (exp. 2R). WAIS scores (including the verbal fluency task scores) are displayed on the y-axis, Reaction time is displayed on the x-axis. The trend line shows the line of best fit.

In general, as expected in 9.2.2, the correlation between participants' cognitive scores and their L2 vocabulary learning success in Experiment 2R was indicated. That meant that participants with higher scores in the cognitive test would do better (higher accuracy and shorter reaction time) than participants with comparatively lower scores in the L2 learning experiment.

9.2.4 Discussion for Experiment 2R

We summarized the main results of Experiment 2R in Table 16. Experiment 2R replicated the significant event presence effect, the testing part effect, and the significant event presence by testing part interaction on L2 learning success (accuracy and reaction time) of Experiment 2. These results confirmed the effectiveness of event-present (compared with event-absent) conditions in both testing parts (immediate vs. delayed) because learners made more correct choices and had shorter reaction times when event photographs were present than when they were not. While the significant event depiction by language mapping interaction in the reaction time analyses for Experiment 2 was not replicated in Experiment 2R, the analyses replicated the significant two-way interaction between language mapping and part. In particular, the results showed, much as in Experiment 2, that in the immediate testing, participants were slower in similar language mapping conditions than in different conditions. However, in the delayed testing, they were faster in similar language mapping conditions than in different conditions, as predicted. We further replicated the absence of positive transfer from L1 to L2 on learning success, for instance, with no main effect of similar (vs. different) verb mapping. The cognitive test scores were not significantly correlated with total accuracy scores and accuracy scores for event-present conditions. However, they were significantly correlated with accuracy scores for event-absent conditions (i.e., the higher the cognitive scores, the lower the accuracy scores).

Table 16: Main findings of Experiment 2R

	Accuracy	**Reaction time**
Without Part	event presence effect	event presence effect
With Part	event presence effect testing part effect event presence by testing part interaction *event presence by language mapping interaction* *	event presence effect testing part effect event presence by testing part interaction language mapping by testing part interaction
Correlations: significant correlation between learners' cognitive test scores and their reaction time		

9.3 Early Middle-aged Adults (Experiment 3)

9.3.1 Participants

Thirty-two adults aged between 32 and 45 (Mean: 36.9, SD: 3.74, Male: 11) participated in Experiment 3. All participants were monolingual native speakers of German. Early middle-aged adults were recruited by LingEx, the Humboldt University mailing list for employment, and by Ebay-kleinanzeigen.de. Testing took place in the reaction time laboratory of the Psycholinguistic Group at Humboldt University. Participants gave informed consent and received €11 for their participation. Testing took approximately 60 minutes in total. The study was approved by an ethics vote (DGfS).

9.3.2 Predictions

We ran Experiment 3 with the same materials, experimental design, and procedure as in Experiment 2R.

Predictions for the L2 vocabulary learning success within Experiment 3 were the same as those for Experiment 2R (see 9.2 for details) because these two experiments were carried out simultaneously (in April and May 2018) after getting results from Experiment 2.

9.3.3 Results and Discussion

The same analysis models were applied in Experiment 3, as in Experiment 2R, generating the following results.

9.3.3.1 Accuracy Results

With testing part as a factor, three outcomes, including the main effect of event presence ($p < .001$) in both testing parts, replicated (see Table 17) the main effect of testing part ($p < .001$), and a significant effect of event presence when testing part interaction ($p < .001$, see Figure 29a).

Table 17: Results from generalized linear mixed model fit by maximum likelihood for accuracy in Experiment 3

Term	Estimate	SE	z value	p
Intercept	1.086	0.210	5.16	<0.001***
Event presence	1.213	0.112	10.82	<0.001***
Language mapping	-0.324	0.192	-1.68	0.0914
Testing Part	0.561	0.104	5.39	<0.001***
Event presence* Testing Part	0.542	0.103	5.22	<0.001***
Language mapping* Testing Part	-0.222	0.103	-2.14	0.0318*

Event presence*Language mapping	0.031	0.109	0.28	0.774
Event presence*Language mapping* Testing Part	-0.183	0.103	-1.77	0.764

L2 adult beginners aged from 32 to 45 also learned L2 phrases much more successfully in present event conditions than absent-event conditions. Overall, they performed testing tasks after the learning phase more accurately in immediate testing than they did in delayed testing. More specifically, they scored around 50% of correct choices in both testing parts for event-absent conditions. In comparison, they responded with higher accuracy for event-present conditions in immediate testing (94.1%) than in delayed testing (72.3%).

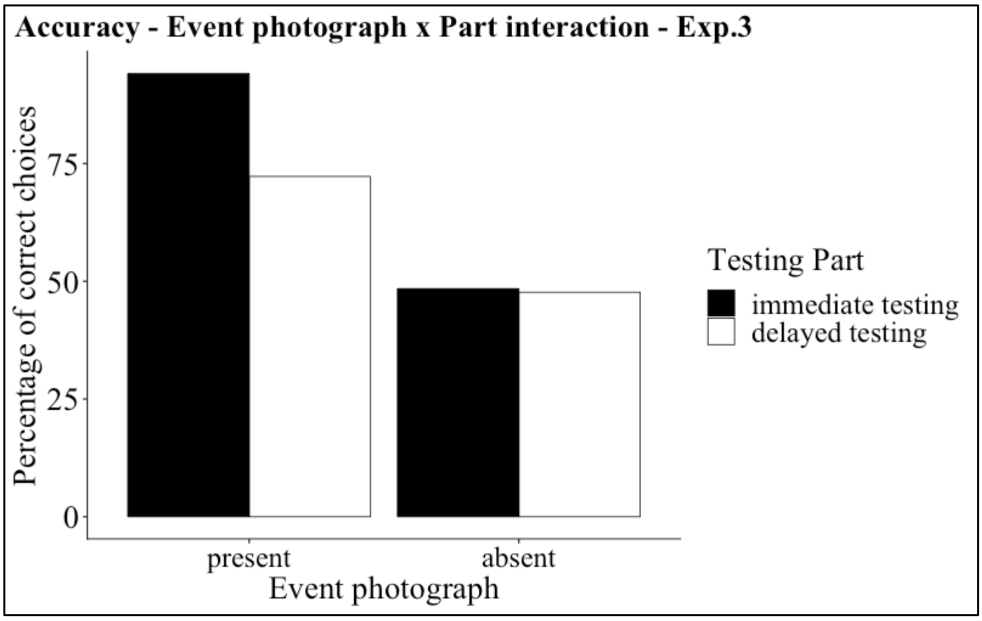


Figure 29a: The significant interaction between event photograph and testing part in the accuracy data in Experiment 3

The only distinct result (when compared with Experiments 2 and 2R) was a significant interaction between language mapping and testing part (p = .0318) in the accuracy data (see Figure 29b). Participants were more accurate in both conditions of language mapping in the immediate testing (similar: 66% and different: 67.6%). Their accuracy percentage in delayed testing reduced to 62.1% for different mapping conditions and declined to 57.8% for similar mapping conditions. Very surprisingly, participants had many more correct choices in total for different language mapping items than for similar ones without the main effect of language mapping.

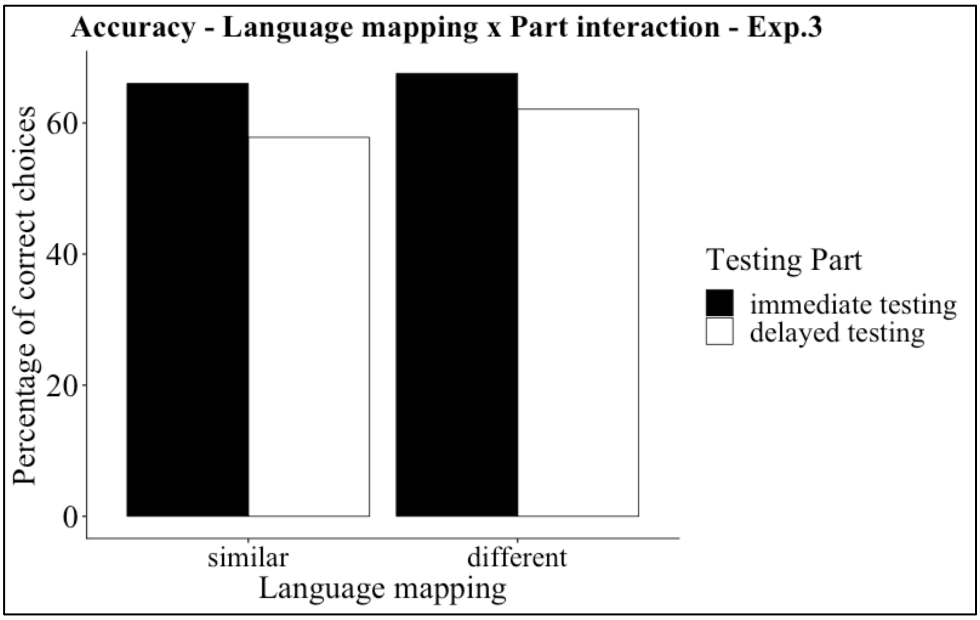


Figure 29b: The significant interaction between language mapping and testing part in the accuracy data in Experiment 3

While not testing part, we replicated a significant effect of event depictions on participants' accuracy as in Experiments 2 and 2R ($p < 0.001$). 83.2% correct choices were made In the event-present learning conditions (in total 512 event-present trials), 83.2% correct choices were made. Participants were much more accurate than in the event-absent learning conditions with 48.04% correct choices for a total of 512 event-absent trials (see Figure 30).

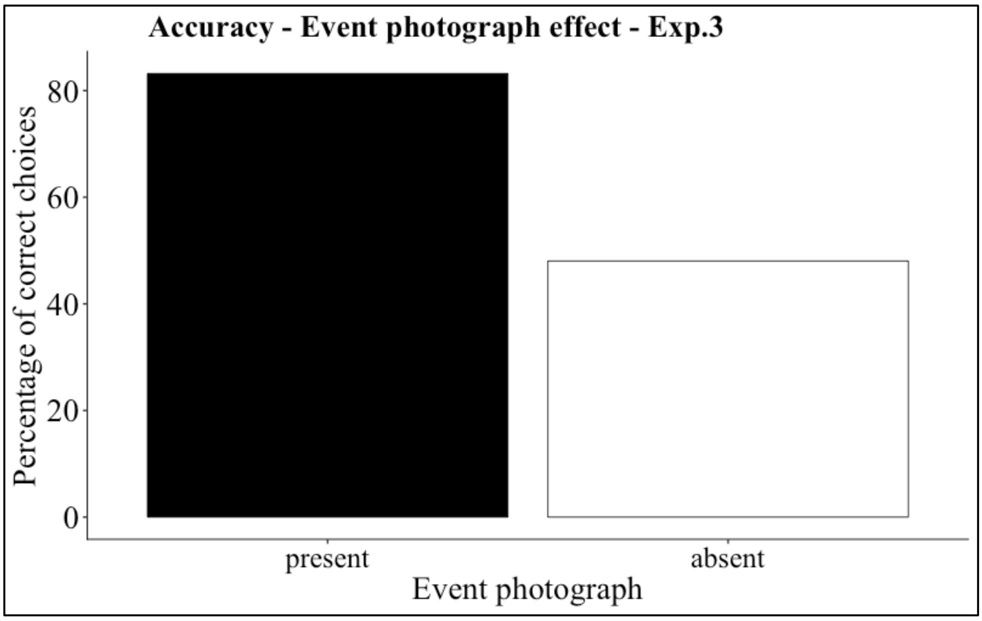


Figure 30: The effect of event photograph on the accuracy in Experiment 3

9.3.3.2 Reaction Time Results

All results from the linear mixed-effects model, including the testing of part as a factor, can be seen in Table 18. As predicted, the analysis model continued to replicate a significant effect of event presence (SE = 0.01, = 12.42, p < 0.001) on participants' reaction times. That means that learners were faster in event-present conditions than in event-absent conditions when they correctly responded in testing. Moreover, a significant effect of the testing of part was also found (SE = 0.02, = 3.97, p < 0.001) when participants decided to make the correct choices more quickly in immediate testing than in delayed testing.

Table 18: Results from linear mixed effects model for reaction time in Experiment 3

Term	**Estimate**	**SE**	**df**	**|t|**	**p**
Intercept	7.460	0.038	42	192.47	<0.001***
Event presence	-0.151	0.012	602	12.42	<0.001***
Language mapping	-0-010	0.020	31	0.52	0.6
Testing Part	-0.080	0.020	33	3.97	<0.001***
Event presence* Testing Part	-0.102	0.011	588	8.78	<0.001***
Language mapping* Testing Part	0.018	0.011	583	1.62	0.103
Event presence*Language mapping	0.007	0.021	600	0.63	0.52
Event presence*Language mapping* Testing Part	0.016	0.011	580	1.40	0.16

A significant event presence by testing part interaction (SE = 0.01, = 8.87, p < 0.001, see Figure 31) was observed. It also indicated that learners were slower for accurately responding to event-present conditions in delayed testing than in immediate testing. They, for event-absent conditions, were faster in delayed testing than in immediate testing. Other interactions (event presence by language mapping, language mapping by part, or event presence by language mapping by part) were non-significant.

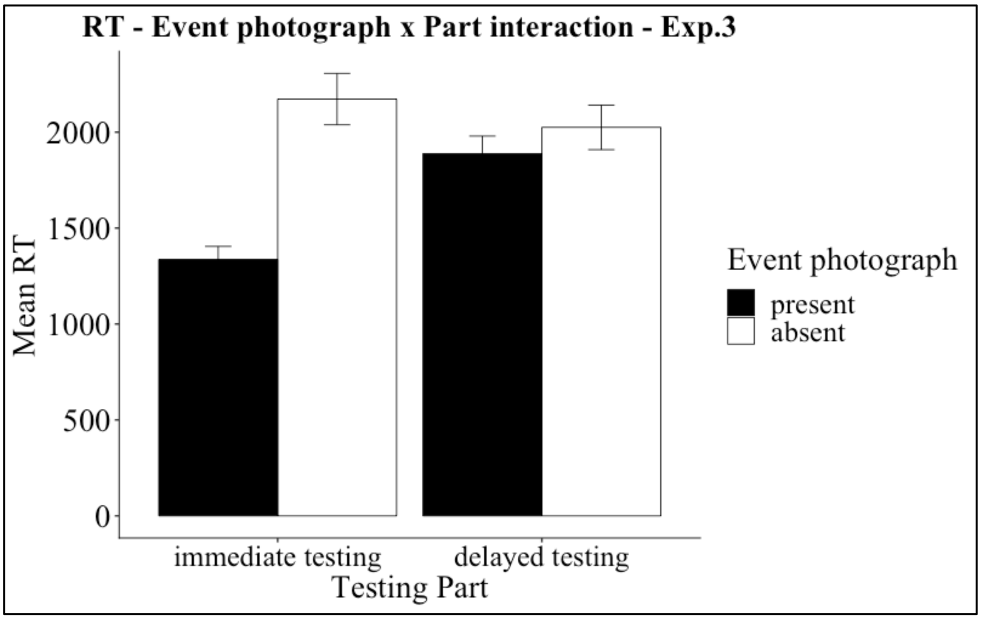


Figure 31 : The interaction between event photograph and testing part in the reaction time data in Experiment 3 (error bars represent 95% CI)

Reaction time results without the testing of part established a significant effect of event presence on reaction time (see Figure 32), as in Experiment 2 (SE = 0.014, = 11.45, $p < 0.001$). The result showed that participants accurately responded faster in event-present conditions than they did in event-absent conditions. We found no further effects involving the manipulated factors.

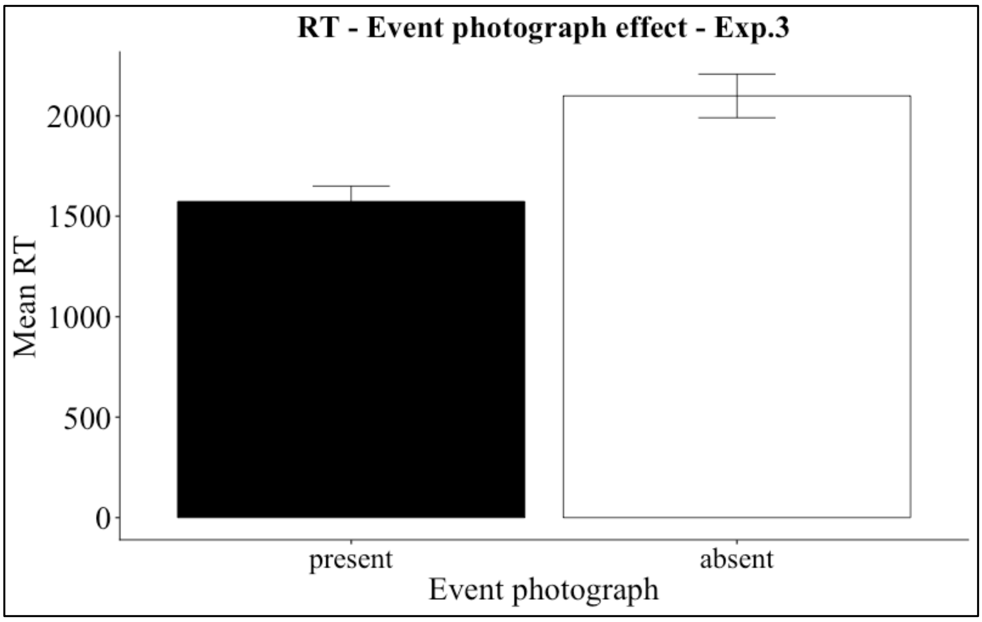


Figure 32 : The effect of event photograph on the reaction time in Experiment 3 (error bars represent 95% CI)

9.3.4 Cognitive Test Results

Early middle-aged adults' cognitive test scores are shown in Figure 33a. We can see from the chart that this group performed best in the similarities task and worst in the picture completion task. The mean for the verbal fluency task was 43.87 words generated.

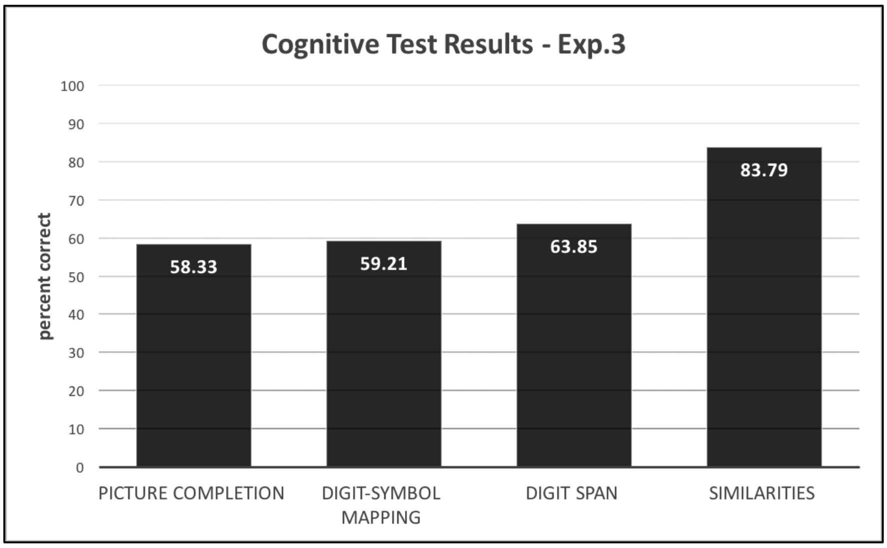


Figure 33a: Cognitive test scores for the WAIS test (exp.3). The y-axis displays the percentage of correct answers averaged across participants. The percentages are shown in the center of each bar. Note that the verbal fluency test scores are not depicted. Since the task was free naming, there is no upper limit that can be reached.

We also computed a Pearson correlation test with both variables that were normally distributed to test whether the accuracy results and the cognitive test scores correlate. The result showed that learners' cognitive test scores and their accuracy scores in L2 vocabulary learning were not significantly correlated (a correlation coefficient of .023, p-value = .9). Figure 33b does not indicate a correlation, but it does depict a gradual increase, meaning that the higher the cognitive scores, the higher the accuracy scores.

Also, the WAIS scores and the response times of early middle-aged adults in L2 learning were not significantly correlated with a correlation coefficient of -.025 and a p-value = 0.89. Figure 33c shows no correlation, just a slight decrease in response time when participants' cognitive scores were higher.

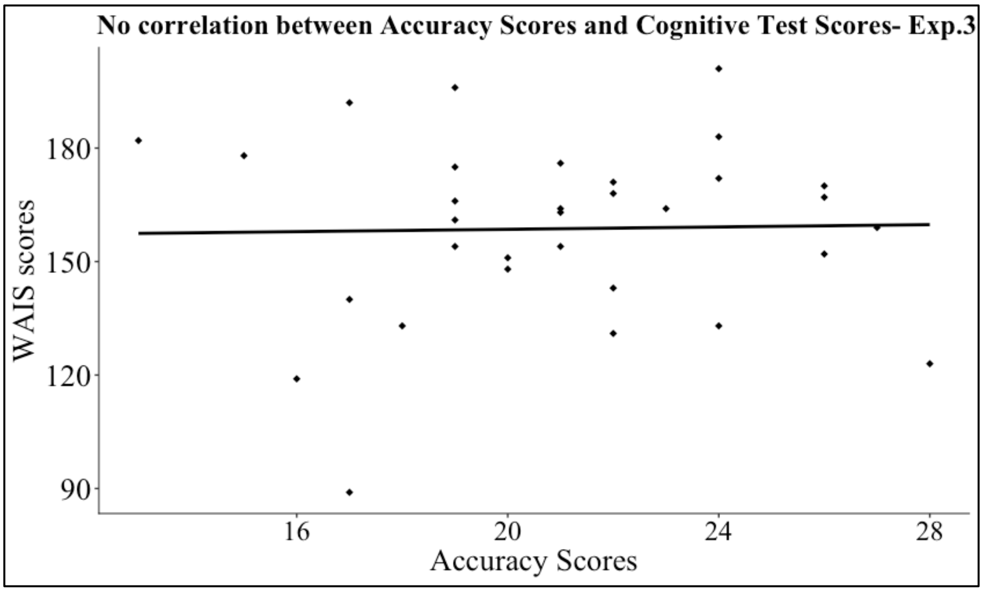


Figure 33b: No correlation between WAIS and accuracy scores (exp. 3). WAIS scores (including the verbal fluency task scores) are displayed on the y-axis, accuracy scores are displayed on the x-axis. The trend line shows the line of best fit

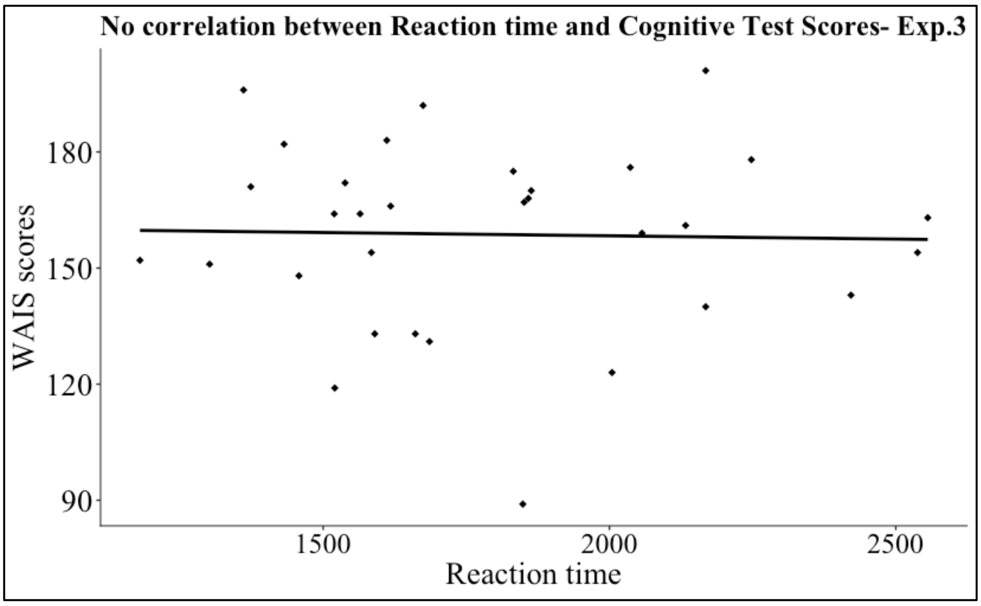


Figure 33c: No correlation between Reaction time and WAIS scores (exp. 3). WAIS scores (including the verbal fluency task scores) are displayed on the y-axis, accuracy scores are displayed on the x-axis. The trend line shows the line of best fit

There is no correlation between accuracy and reaction time in L2 phrase learning for early middle-aged adults' accuracy because of a correlation coefficient of .058 and p-value = 0.75. Figure 33d displays no correlation, but the line depicts a slight increase in accuracy when the reaction time is longer.

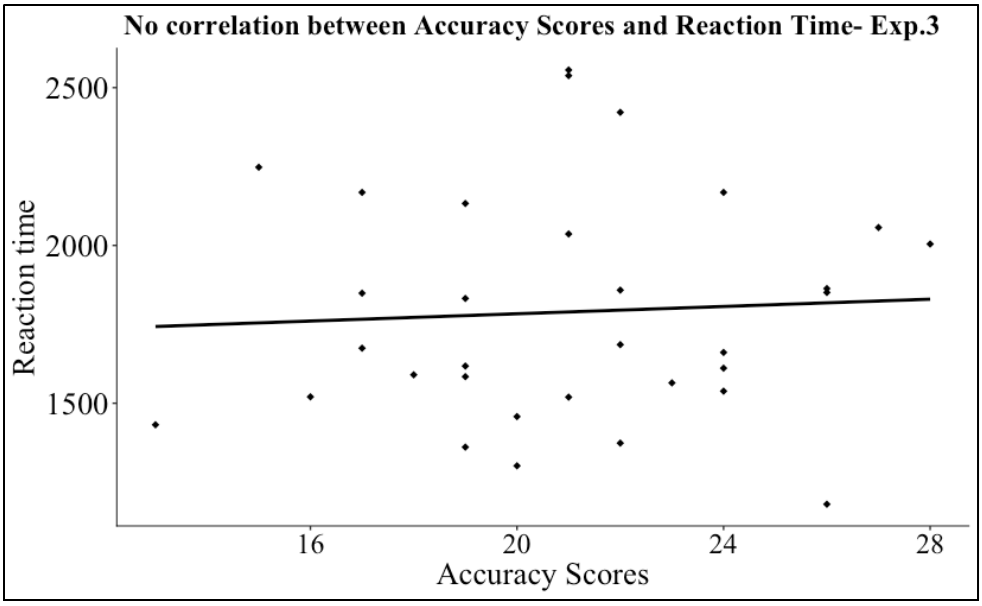


Figure 33d: No correlation between Accuracy Scores and Reaction time (exp. 3). WAIS scores (including the verbal fluency task scores) are displayed on the y-axis, accuracy scores are displayed on the x-axis. The trend line shows the line of best fit

9.3.5 Discussion for Experiment 3

The main results of Experiment 3 are shown in Table 19. Results of Experiment 3 replicated the significant event presence effect, the testing part effect as well as the significant event presence by testing part interaction on L2 learning success (accuracy and reaction time) from Experiment 2 and 2R. These results confirmed the effectiveness of event-present (compared with event-absent) conditions in both testing parts (immediate vs. delayed) because learners made more correct choices and had shorter reaction times when event photographs were present than absent. The analyses replicated no significant two-way interactions from Experiments 2 and 2R, but there was a new significant two-way interaction between language mapping and testing part on learners' accuracy. We further replicated the absence of the main effect of language mapping, as in the preceding two experiments, 2 and 2R. The cognitive test scores were neither significantly correlated with accuracy scores nor with the response time of adult learners in the group.

Table 19: Main findings of Experiment 3

	Accuracy	**Reaction time**
Without Part	event presence effect	event presence effect

With Part	event presence effect testing part effect event presence by testing part interaction *language mapping by part interaction*	event presence effect testing part effect event presence by testing part interaction
Correlations: No significant correlations: between learners' cognitive test scores and their reaction time; between learners' cognitive test scores and their accuracy; between learners' accuracy and their reaction time		

9.4 Late Middle-Aged Adults (Experiment 4)

9.4.1 Participants

Thirty-two adults aged between 46 and 65 (Mean: 54.25, SD: 5.96, Male: 10) participated in the experiment. All participants were monolingual native speakers of German. Later middle-aged adults were recruited via the Humboldt University mailing list and Social Networks such as Facebook, Ebay-kleinanzeigen.de. Testing took place in the reaction time laboratory of the Psycholinguistic Group at Humboldt University. Participants gave informed consent and received €11 for their participation, and testing took approximately 60 minutes in total. The study was approved by an ethics vote (DGfS).

9.4.2 Predictions

We conducted Experiment 4 with the same materials, experimental design, and procedure as in Experiments 2R and 3. Predictions for the L2 vocabulary learning success in Experiment 4 were changed compared to Experiments 2R and 3 because we had the results of Experiments 2R and 3 before making the predictions for Experiment 4.

Predictions for the cognitive test (WAIS scores): If participants had higher scores in the cognitive test, then they would do better in the L2 learning experiment (higher accuracy and shorter reaction time) than others.

Predictions for the main results of L2 vocabulary learning success:

If events are useful in learning, then participants could have more accurate and faster responses (by pressing one of two keyboard keys, Q or P) in the "event-present" condition compared with the "event-absent" one (as in Experiments 2 and 3). That means that participants would get many more correct choices in the whole experiment and in each testing part when they were presented with event photographs as compared to no event photographs. They could also take a shorter time to respond correctly in the event-present condition than the event-absent one.
If there is a strong effect of part, then participants could give higher correct choices and faster responses in Part 1 compared with Part 3 because participants have to access new photographs referring to the same verb-noun phrases as in Part 1 (as in Experiments 2 and 3).
If the event effects differ depending on the testing part, then we should see a significant interaction between visual context (event photograph depiction) and parts and 3 on participants' phrase learning (as in Experiments 2 and 3). That means that participants performed their tasks better in Part 1 (shorter reaction time, more correct answers) than in Part 3 if they had seen event photographs. However, participants performed their tasks equally well in Part 1 and Part 3 (approximately 50% of answers correct, and similar reaction time) if they had not seen event photographs.
No main language mapping effect and interactions related to the "language mapping" was factor were predicted because we found no significant effect of language mapping in three previous experiments (2, 2R, and 3). Also, some interactions were not identical among experiments.

9.4.3 Results and Discussion

With the same analysis methods and settings, as in previous studies, we obtained the results below.

9.4.3.1 Accuracy Results

With testing part included as a factor, the results replicated (see Table 20) were the main effect of event presence ($p < .001$) in both testing parts, the main effect of testing part ($p < .001$), and a significant event presence by testing part interaction ($p < .001$). The only new result compared with Experiments 2 and 3 (but the same as in Experiment 2R) was a just-significant interaction between event presence and language mapping ($p = 0.0485$).

Table 20: Results from generalized linear mixed model fit by maximum likelihood for accuracy in Experiment 4

Term	Estimate	SE	z value	p
Intercept	1.206	0.176	6.83	<0.001***
Event presence	1.238	0.110	11.23	<0.001***
Language mapping	-0.076	0.169	-0.45	0.65
Testing Part	0.531	0.103	5.13	<0.001***
Event presence* Testing Part	0.522	0.103	5.04	<0.001***
Language mapping* Testing Part	0.067	0.103	0.65	0.52
Event presence*Language mapping	0.213	0.108	1.98	0.0485*
Event presence*Language mapping* Testing Part	0.076	0.103	0.73	0.46

Late middle-aged adults learned L2 phrases more successfully in event-present conditions than event-absent conditions. In event-absent conditions, late middle-aged adults made around 38% correct choices for both testing parts. In comparison, they responded with much higher accuracy for event-present conditions in immediate testing (84.4%) than in delayed testing (67.6%). Figure 34a illustrates the significant event depictions when testing part interaction in Experiment 4.

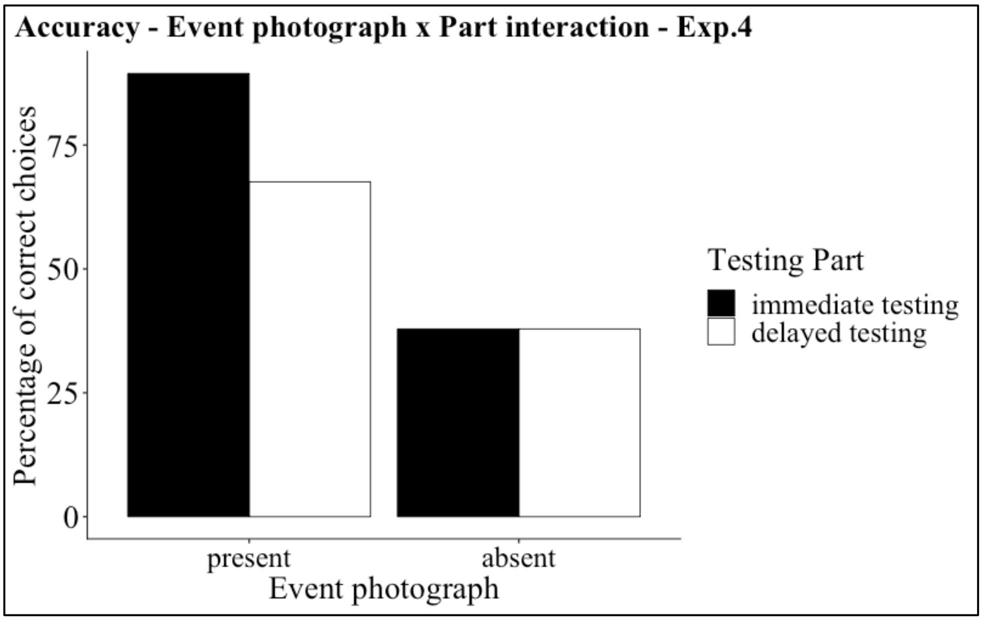


Figure 34a: The interaction between event photograph and testing part on the accuracy in Experiment 4

The significant language mapping by event photograph interaction (see Figure 34b) indicated that adults in the group performed better in different verb mapping conditions than they did in similar verb mapping conditions. When event photographs were present, they made 82.03% correct choices for different mapping and 75% for similar mapping. When event photographs were absent, they made 42.19% correct choices for different mapping and 33.59% for similar mapping. We did not expect the direction of the interaction, and we could not explain it.

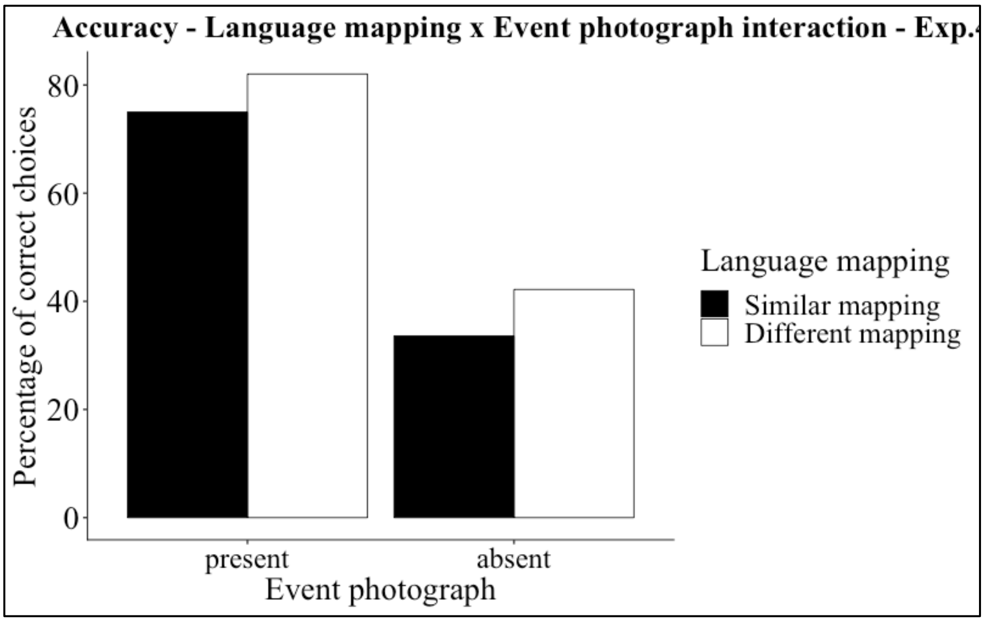


Figure 34b: The interaction between mapping and event photograph on the accuracy in Experiment 4

Omitting the testing part, we replicated a significant effect of event depictions on participants' accuracy, just as we did in three previous studies ($p < 0.001$). In the event-present learning conditions (in total 512 event-present trials), 78.51% correct choices were made. Participants were much more accurate than in the event-absent learning conditions, with 37.89% correct choices for a total of 512 event-absent trials (see Figure 35).

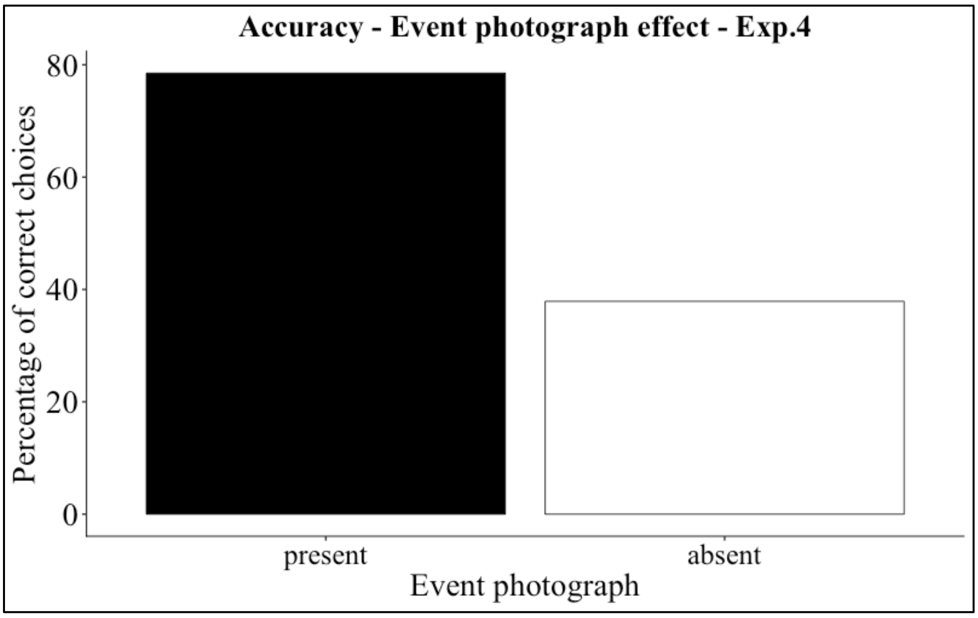


Figure 35. The effect of event photograph on the accuracy in Experiment 4

9.4.3.2 Reaction Time Results

All results from the linear mixed-effects model, including the testing part as a factor, are presented in Table 21. As predicted, the analysis model continued to replicate a significant effect of event presence (SE = 0.016, = 8.99, $p < 0.001$) on participants' reaction time. That means that learners were faster in present event conditions than in absent-event conditions when they responded correctly in testing. Furthermore, a significant effect of the testing part was also found (SE = 0.021, = 6.07, $p < 0.001$) when participants made their correct choices more quickly in immediate testing than in delayed testing.

Table 21: Results from linear mixed-effects model for reaction time in Experiment 4

Term	**Estimate**	**SE**	**df**	**|t|**	**p**
Intercept	7.479	0.036	40	202.69	<0.001***
Event presence	-0.145	0.017	554	8.36	<0.001***
Language mapping	-0.001	0.023	36	0.053	0.957
Testing Part	-0.154	0.016	551	9.17	<0.001***
Event presence* Testing Part	-0.063	0.016	544	3.76	= 0.001***
Language mapping* Testing Part	0.038	0.016	548	2.301	0.0217*
Event presence*Language mapping	0.012	0.017	553	0.711	0.4774
Event presence*Language mapping* Testing Part	0.014	0.016	542	0.853	0.3937

A significant event presence by testing part interaction (SE = 0.019, = 8.092, $p < 0.001$, see Figure 36a) was found. It showed that learners were much faster in immediate testing than in delayed testing for both event-present conditions and event-absent conditions. More noteworthy, the mean of response time for event-present items increased much more from immediate testing to delayed testing than for event-absent items.

Another interaction between language mapping and testing part was replicated from Experiment 2 (see Figure 36b). While participants, in immediate testing, were slower to respond in similar language mapping conditions than in different language mapping conditions, they were faster for similar ones in delayed testing than for different ones.

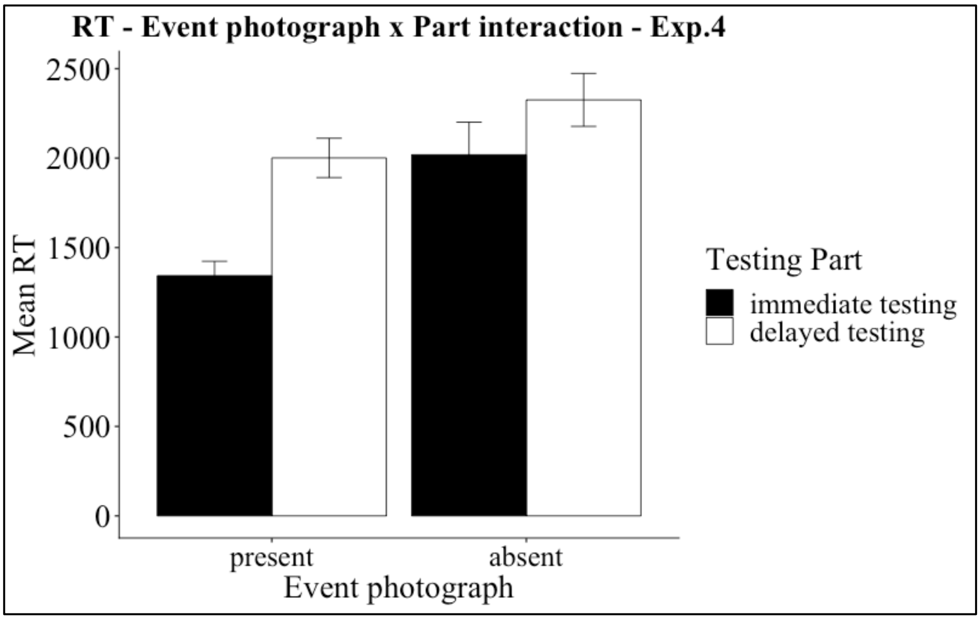


Figure 36a: The interaction between event photograph and part on the reaction time in Experiment 4 (error bars represent 95% CI)

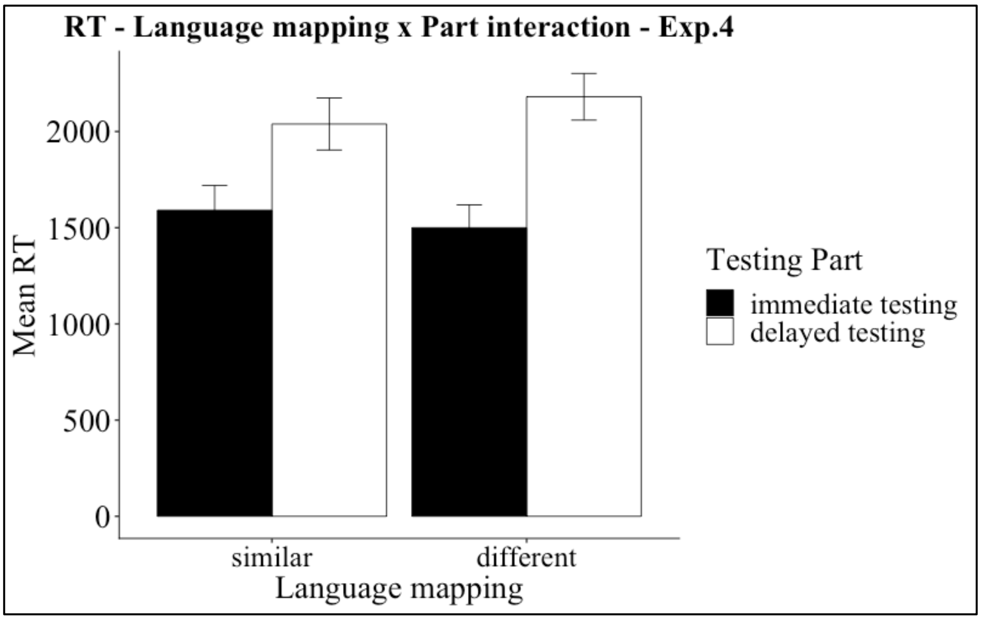


Figure 36b: The interaction between language mapping and testing part on the reaction time in Experiment 4 (error bars represent 95% CI)

Reaction time results without the testing part showed a significant effect of event depictions on reaction times (see Figure 37), as was the case in three previous experiments (SE = 0.019, = 8.09, $p < 0.001$). We found no further effects related to the manipulated factors.

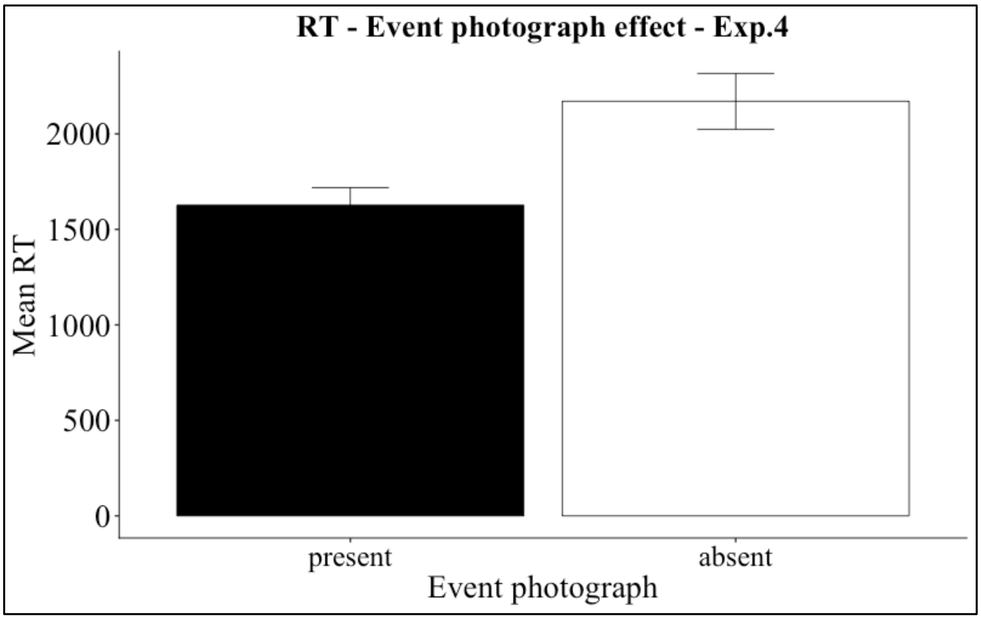


Figure 37: The effect of event photograph on the reaction time in Experiment 4 (error bars represent 95% CI)

9.4.3.3 Cognitive Test Results

The cognitive test scores of late middle-aged adults are depicted in Figure 38. As we can see from the chart, they performed best in the similarities task and worst in the digit symbol mapping task. The mean for the verbal fluency task was 41.46 words generated.

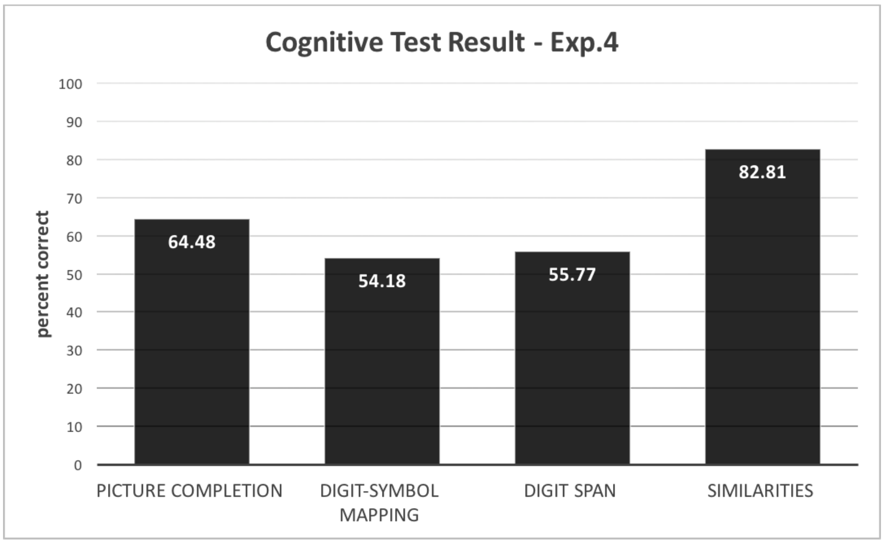


Figure 38: Cognitive test scores for the WAIS test (exp.4). The y-axis displays the percentage of correct answers averaged across participants. The percentages are shown in the center of each bar. Note that the verbal fluency test scores are not depicted. Since the task was free naming, there is no upper limit that can be reached.

A Pearson correlation test was computed with the two variables that were typically distributed to examine whether the accuracy results and the cognitive test scores correlate significantly. The result showed that learners' cognitive test scores and their accuracy scores in L2 vocabulary learning were not significantly correlated (a correlation coefficient of .05, p- = .78). Figure 38b indicates no correlation, but it describes a gradual increase, which means the higher the cognitive scores, the higher the accuracy scores.

We found no significant correlation between WAIS scores and reaction times in the L2 learning of late middle-aged adults. We found a correlation coefficient of -.09 and a p-value = 0.62. Figure 38c indicates no correlation, but an apparent decrease in response time when participants' cognitive scores are higher.

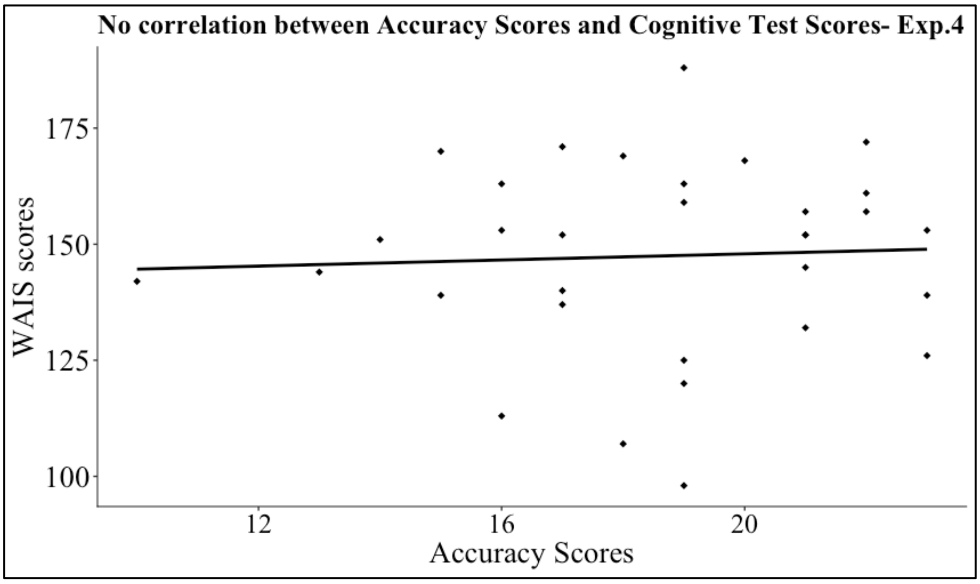


Figure 38b: No correlation between WAIS and accuracy scores (exp. 4). WAIS scores (including the verbal fluency task scores) are displayed on the y-axis, accuracy scores are displayed on the x-axis. The trend line shows the line of best fit

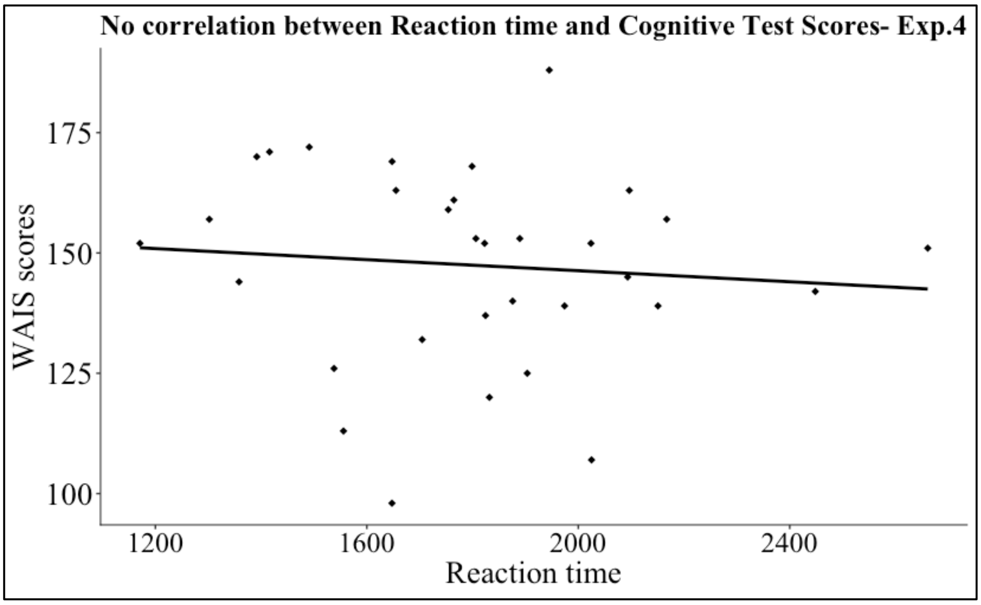


Figure 38c: No correlation between Reaction time and WAIS scores (exp. 4). WAIS scores (including the verbal fluency task scores) are displayed on the y-axis, accuracy scores are displayed on the x-axis. The trend line shows the line of best fit

No correlation between accuracy and reaction time in L2 phrase learning for later middle-aged adults was found. We found a correlation coefficient of -.23 and a p-value = 0.18. Figure 38d indicates no correlation, but the line depicts an apparent increase in accuracy scores when the reaction time is shorter.

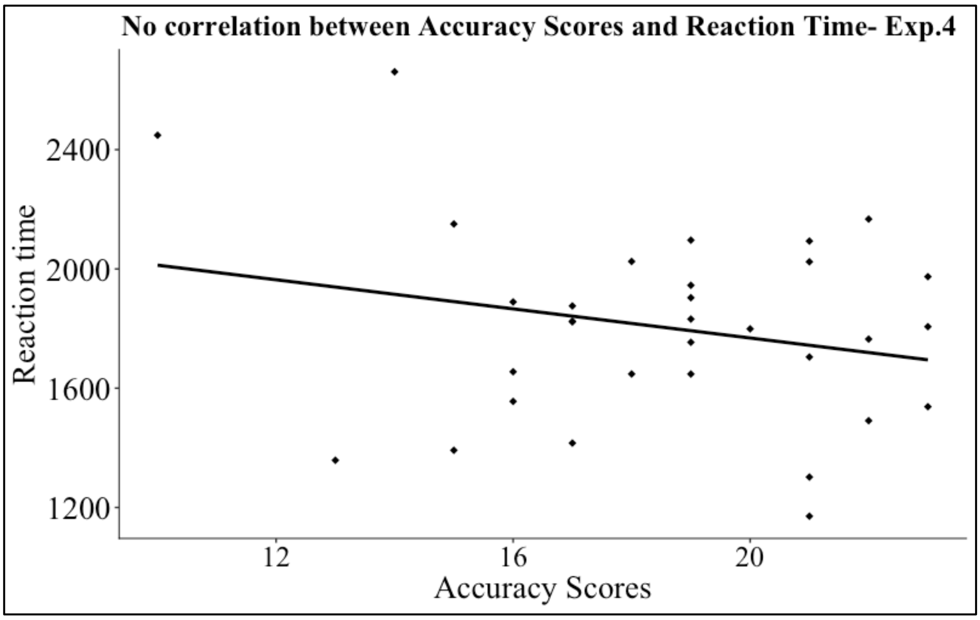


Figure 38d: No correlation between Accuracy scores and Reaction time (exp. 4). The means of reaction time are displayed on the y-axis, accuracy scores are displayed on the x-axis. The trend line shows the line of best fit.

9.4.4 Discussion for Experiment 4

The main results of Experiment 4 are summarized in Table 22. Experiment 4 replicated the significant event presence effect, the testing part effect, and the significant event presence by testing part interaction on L2 learning achievement (accuracy and reaction time) from experiments 2, 2R, and 3. These results confirmed the effectiveness of event-present (compared with event-absent) conditions in the two testing parts (immediate vs. delayed). Learners made more right choices and had quicker reaction times when event photographs were present than when they were absent. The analyses replicated a significant two-way language mapping by testing part interaction in reaction time data from Experiments 2 and 2R. We further confirmed no main effect of language mapping on L2 learning success as in the three previous studies 2, 2R, and 3. The cognitive test scores were not significantly correlated with the accuracy scores and the response times of adult learners in Experiment 4.

Table 22: Main findings of Experiment 4

	Accuracy	**Reaction time**
Without Part	event presence effect	event presence effect
With Part	event presence effect testing part effect event presence by testing part interaction	event presence effect testing part effect event presence by testing part interaction *language mapping by part interaction*

Correlations: No significant correlations: between learners' cognitive test scores and their reaction time; between learners' cognitive test scores and their accuracy; between learners' accuracy and their reaction time

9.5 Individual Differences of Adults aged from 18 to 65 and Second Language Vocabulary Learning Success

First, we want to summarize the similarities and differences between the main effects of the three L2 vocabulary learning experiments. The collective findings and various significant interactions of the three studies are listed in Table 23. All three adult groups learned L2 vocabulary more successfully (higher accuracy and shorter reaction time) in depicted event photograph conditions than in undepicted events. Also, they all performed much better and faster in L2 vocabulary immediate testing than in delayed testing. Unexpectedly, L1–L2 similar language mapping conditions did not benefit participants' L2 vocabulary learning compared with different conditions in our specific L2 learning studies. Some significant interaction between language mapping and other experimental factors (i.e., event presence or testing part) varied in different experiments and measured dependent variables (i.e., accuracy or reaction time). The one significant interaction between language mapping and testing part on reaction time in Experiments 2R and 4 had the same pattern. However, they were a little different compared to that interaction in Experiment 2.

Table 23: Comparisons of main findings in Experiments 2R, 3 and 4

Analysis model: glmer (Accuracy) and lmer (Reaction time)	
Accuracy	*** Similarities:** a significant effect of visual context (event depiction) a significant effect of testing part a significant visual context by testing part interaction
	*** Differences:** Experiment 2R: a significant event presence by language mapping interaction Experiment 3: a significant interaction between language mapping and testing part
Reaction time	*** Similarities:** a significant effect of visual context (event depiction) a significant effect of testing part a significant visual context by part interaction
	*** Differences:** Experiment 2R: a significant language mapping by testing part interaction; a significant correlation between learners' cognitive test scores and their reaction times Experiment 4: a significant interaction between language context and testing part

9.5.1 Comparisons of Cognitive Test Scores for Experiments 2R, 3, and 4

We predicted the relationship between cognitive scores and participants' L2 vocabulary learning success per experiment, as described in previous sections (9.2; 9.3; 9.4). The only significant correlation between learners' cognitive test scores and their reaction times (i.e., young adult learners) was found in Experiment 2R.

We also wanted to see whether there were significant correlations between learners' cognitive test scores and their accuracy or reaction times across three different age groups even though we did not predict them.

Two more Pearson correlation tests were computed with the two normally distributed variables to examine whether accuracy and reaction time results and the cognitive test scores have a significant correlation. Results indicated that adult learners' cognitive test scores and their accuracy scores in L2 vocabulary learning were not significantly correlated (correlation coefficient of .05, p- = 0.59). However, we got a marginally significant correlation between cognitive test scores and reaction time in the L2 vocabulary learning for all three adult groups with a correlation coefficient of -.18 and p- = 0.071 (see Figures 39a and 39b).

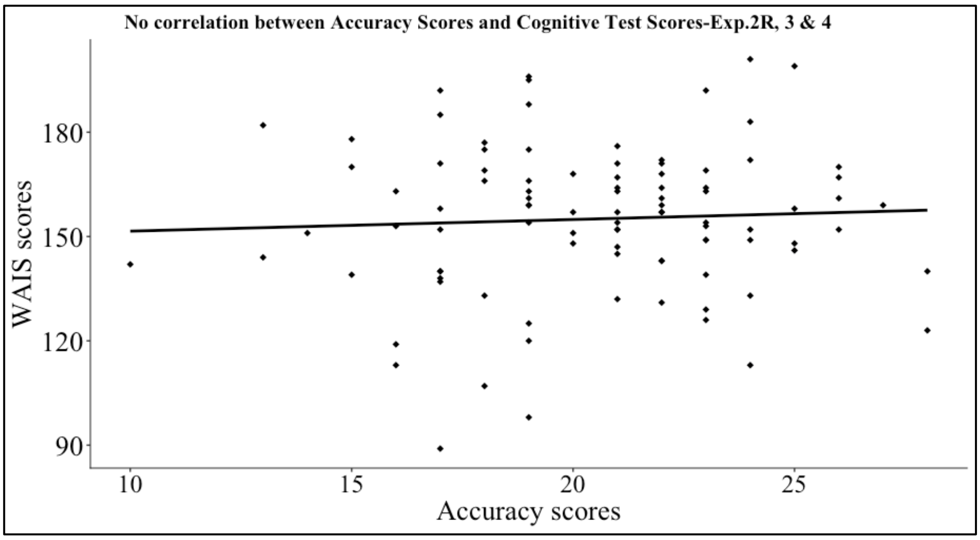


Figure 39a: No correlation between Cognitive Test Scores and Accuracy scores (96 participants). The means of cognitive test scores are displayed on the y-axis, accuracy scores are displayed on the x-axis. The trend line shows the line of best fit.

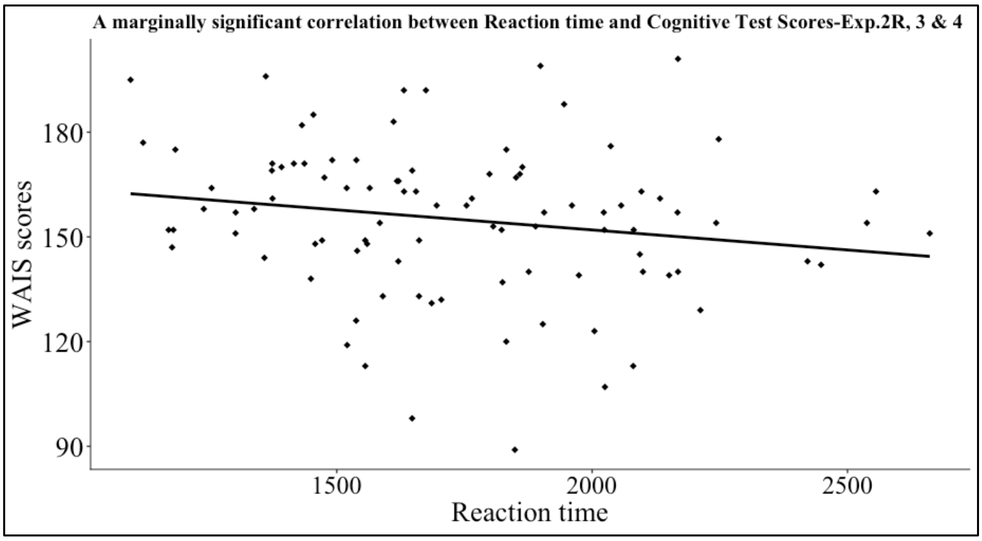


Figure 39b: A marginally significant correlation between Cognitive Test Scores and Reaction Time (96 participants). The means of cognitive test scores are displayed on the y-axis, the means of reaction time are displayed on the x-axis. The trend line shows the line of best fit.

9.5.2 Age Comparison for Experiments 2R, 3 and 4

In the section, we made a comparison to see how age-related differences can influence adult learners' cognitive ability and their L2 vocabulary learning success.

9.5.2.1 Predictions and Analysis Methods

Predictions for differences of WAIS scores: The youngest age group will achieve the highest scores for general cognitive abilities, and the oldest age group will get the highest scores for linguistic abilities.

Predictions for Accuracy Scores and Reaction Times: Age-related differences can also strongly affect learners' L2 vocabulary learning success.

- For accuracy scores: We had the accuracy results for Experiments 2 and 3 before making the following predictions. We predicted no substantially significant differences in the number of correct choices for the three age groups. This means that participants in all age groups would get an approximate number of correct answers in this experimental design for L2 vocabulary learning.

- For reaction times: We already had results for experiments 2 and 3 showing no significant difference in the mean scores for reaction times between young adults and early middle-aged adults. However, if learner age plays an essential role in L2 learning success, then there would be significant differences in response times (i.e., means for reaction times) among the learner groups. In other words, young adults and early middle-aged adults would respond much faster for correct choices compared with late middle-aged adults.

9.5.2.2 Analyses

To compare age-related differences in L2 vocabulary learning success in more detail, we conducted repeated measure ANOVAs with age as a between-subjects factor for the subject analysis and as a within-subjects factor for the item analyses for the cognitive test scores, accuracy scores, and reaction times. The analysis included all three age groups (younger adults, early middle-aged adults, and late middle-aged adults).

9.5.2.3 Results

9.5.2.3.1 Age and Cognitive Test Results

All the data on the cognitive tests of three adult groups can be seen in Table 24.

Table 24: Cognitive test results and demographic characteristics among adult learner groups

Characteristic	Experiment 2R	Experiment 3	Experiment 4
Age range	18–31	32–45	46–65
Mean age in years	23.62 (3.42)	36.9 (3.74)	54.25
Picture completion[a]	4.18 (0.93)	3.5 (1.16)	3.62 (0.94)
Digit Symbol[a]	82.94 (13.53)	78.75 (17.99)	72.06 (13.94)
Digit Span[a]	17.59 (2.77)	19.15 (4.77)	17.03 (3.09)
Similarities[a]	12.78 (1.86)	13.4 (1.95)	13.25 (2.15)
Vocabulary (Animal naming[b] and Word naming[c])	41.62 (7.29)	43.87 (8.76)	41.46 (8.28)
Total scores for WAIS test	159.12 (18.43)	158.69 (23.93)	147.44 (20.39)
BMIS	7.62 (0.83)	7.4 (1.16)	7.56 (2)
Male/female (n)	14/18	11/21	10/22
Note.			

[a] The latest test German version of the Wechsler Adult Intelligence Scale (WAIS), [79-80].
[b] Task: Name as many animals as possible, time allowed: 1min.
[c] Task: Name as many words as possible starting with the letter "l," time allowed: 1 m

As we predicted, the young adults had the highest scores for *general cognitive abilities (fluid intelligence)* in subtests such as picture completion, digit symbol, and digit span (mean scores: 104.72) compared to two other adult groups (mean scores:101.41 and 92.72). There was a clear gap in performance between young adults (aged 18–31) and the two older groups for picture completion and digit symbol—the two older groups, by contrast, are relatively homogenous (see more Table 24).

For *linguistic abilities (crystalized intelligence)* comprising the two subtests of similarities and vocabulary, we found that early middle-aged adults had the highest scores (mean: 57.28). However, the two other adult groups had approximately the same scores (54.41 for young adults and 54.72 for late middle-aged adults). Our prediction was wrong when we thought late middle-aged adults would outperform the two other groups. Figure 40 contains two lines showing the two main testing parts of the cognitive tests across three experiments.

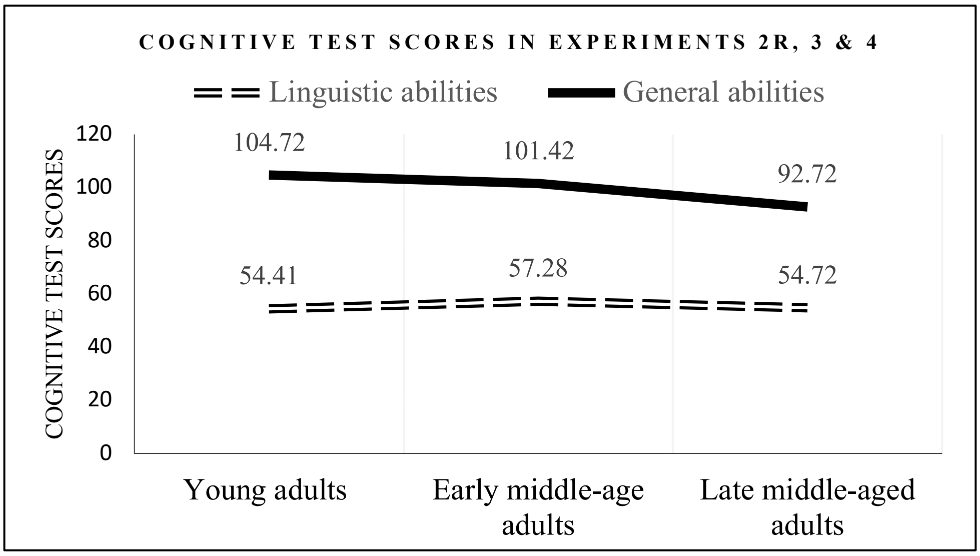


Figure 40: Cognitive test scores of three adult learner groups

We computed a Pearson correlation test to examine whether learners' ages and cognitive test scores were significantly correlated. A significant correlation between the adult age range (18–65) and their WAIS scores was investigated. There was a correlation coefficient of -.25 and p-value = .011. Figure 41 shows a significant correlation indicating that the higher the cognitive test scores, the younger the L2 learners.

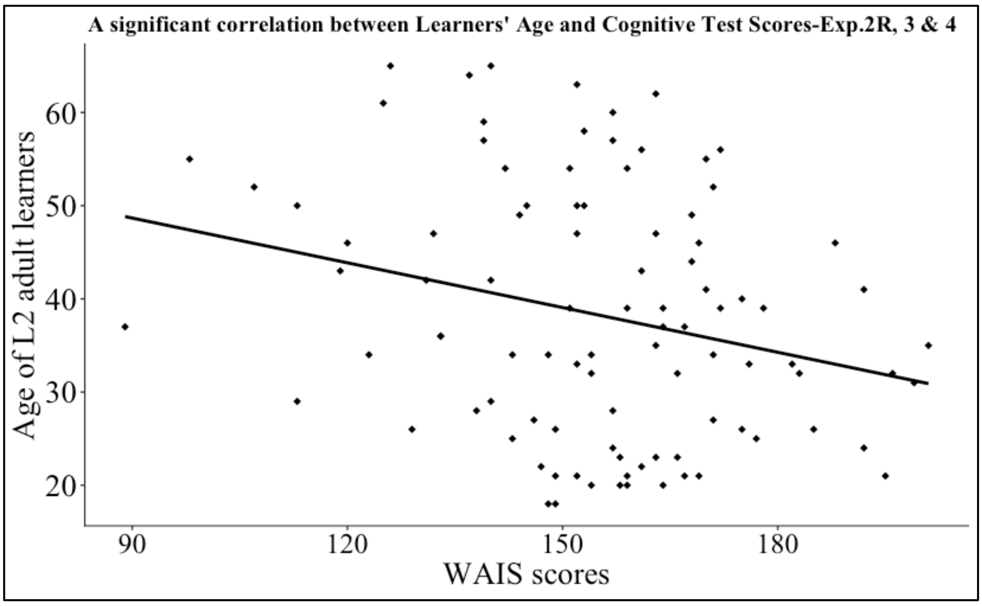


Figure 41: A significant correlation between Age and WAIS scores (exp. 2R, 3 & 4). The means of learners' age are displayed on the y-axis, WAIS scores are displayed on the x-axis. The trend line shows the line of best fit.

We also wanted to know whether there were significant differences in cognitive test scores on the whole among adult learner groups. We computed one-way between-groups ANOVA in R. The result showed a just-significant difference in the mean for cognitive test scores [$F(2, 93) = 3.171$, $p = 0.0465$] between learner groups. Post hoc comparisons using the Tukey test were carried out. There was a marginally significant difference between late middle-aged adults and young adults ($p = 0.07$) with cognitive test scores of 11.69, on average, less than those for young adults. There was also a marginally significant difference between late middle-aged adults and early middle-aged adults ($p = 0.08$) with cognitive test scores of 11.25, on average, less than those for early middle-aged adults.

9.5.2.3.2 Age and Accuracy in Second Language Vocabulary Learning

Adult learners in all three groups benefited from event-present conditions (compared to event-absent conditions) in L2 vocabulary learning. They also performed the task much more accurately in immediate testing than in delayed testing. We found the same patterns of accuracy scores for the event presence effect and the testing part effect in three studies. Table 25 lists these significant effects.

Table 25: Accuracy in L2 vocabulary learning experiments 2R, 3 & 4

Experiment	Event photographs	Testing Part	Correct choices (%)
Exp.2R (18-31 years)	Present	Immediate	96.1
		Delayed	77.34
	Absent	Immediate	50.4
		Delayed	50
Exp.3 (32-45 years)	Present	Immediate	94.1
		Delayed	72.3
	Absent	Immediate	48.4
		Delayed	47.7
Exp.4 (46-65 years)	Present	Immediate	84.4
		Delayed	67.6
	Absent	Immediate	37.9
		Delayed	37.9

A Pearson correlation test indicated that learners' age and accuracy scores were significantly correlated with a correlation coefficient of -.366 and p-value = .00024. Figure 42 illustrates the significant correlation indicating that the higher the accuracy scores in L2 vocabulary learning, the younger the learners. The pattern was the same for the significant correlation between learners' age and their cognitive test scores.

A one-way between-group ANOVA showed that there was a significant difference in the means for accuracy scores [$F(2, 93) = 9.354$, $p < 0.001$] between learner groups. Post hoc comparisons using the Tukey test were computed. There were statistically significant differences: between late middle-aged adults and young adults ($p < 0.001$); between late middle-aged adults and early middle-aged adults ($p = 0.0081$). Both young adults and early middle-aged adults had higher accuracy (means of accuracy scores overall) than late middle-aged adults. Figure 42b shows the differences in the means of the accuracy scores among adult learner groups. Figure 42b shows that participants' average scores for accuracy decrease when they are older.

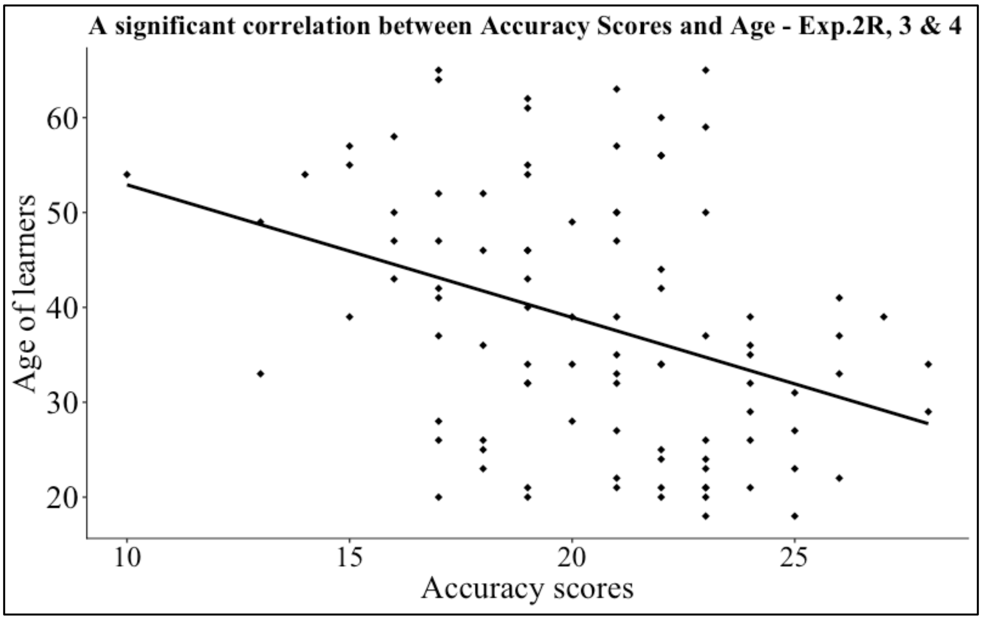


Figure 42a: A significant correlation between Learners' Age and Accuracy Scores (exps. 2R, 3 & 4). The means of learners' age are displayed on the y-axis, Accuracy scores are displayed on the x-axis. The trend line shows the line of best fit.

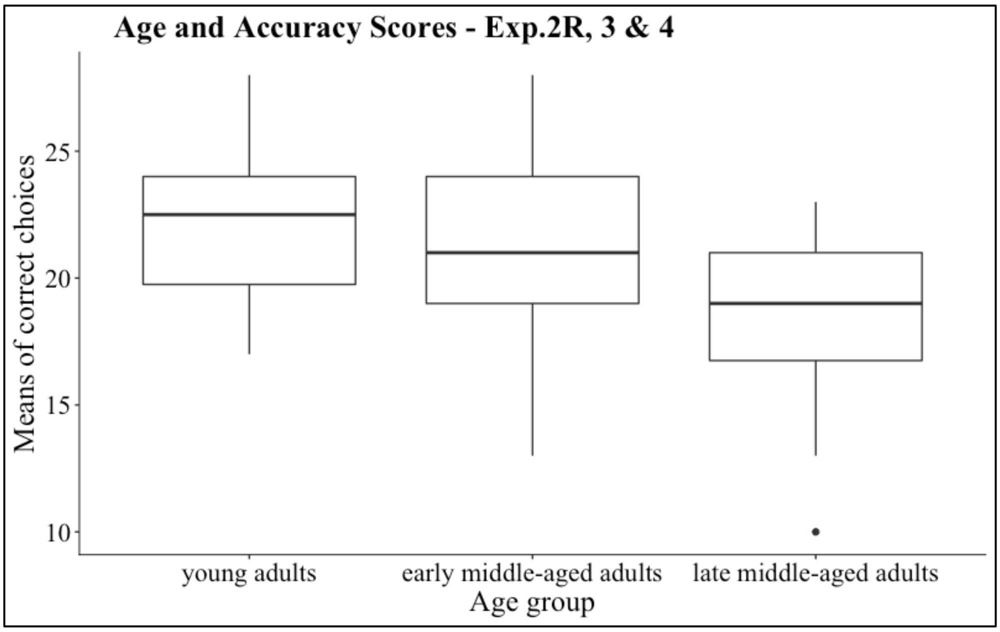


Figure 42b: Mean of Accuracy Scores among three adult groups

The differences in accuracy scores can be seen more clearly when the experimental factor, event depiction, was included in the graph (see Figure 42c). That is because all adult learners benefited from event photographs when they learned L2 Vietnamese phrases in the L2 vocabulary learning design. While young adults and early middle-aged adults had approximately the same mean accuracy scores for event-absent conditions, late middle-aged adults achieved low accuracy scores when they responded in the learning conditions.

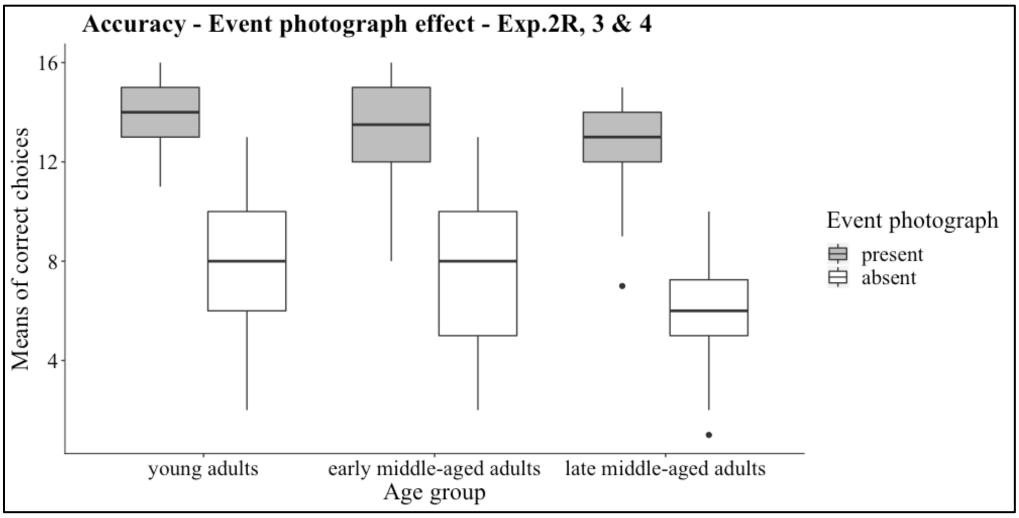


Figure 42c: The main effect of event photographs on L2 accuracy of three adult learner groups

9.5.2.3.3 Age and Reaction Time in Second Language Vocabulary Learning

First, we estimated the means for reaction times per event presence factor (present vs. absent) per experiment. A similarity among the three adult groups was the strongly significant effect of event depictions (i.e., event-present conditions) on their speed of accurately learning L2 phrases.

Figure 48 illustrates much faster responses in event-present conditions than in event-absent conditions in all learner groups. Young adults had the shortest reaction time with/without event photograph learning conditions, and the reaction time pattern of middle-aged adults (early and late) seemed to be the same. One-way between-group ANOVA showed a significant difference in the mean for learners' speed [$F(2, 93) = 3.338$, $p = 0.0398$] between learner groups. Post hoc comparisons using the Tukey test were also computed. There was a statistically significant difference between late middle-aged adults and young adults ($p = 0.0578$). We found a marginally significant difference between young adults and early middle-aged adults ($p = 0.0866$). There was no significant difference in reaction times between the two groups of middle-aged adults. Figure 43a shows the differences related to means for reaction times in adult groups.

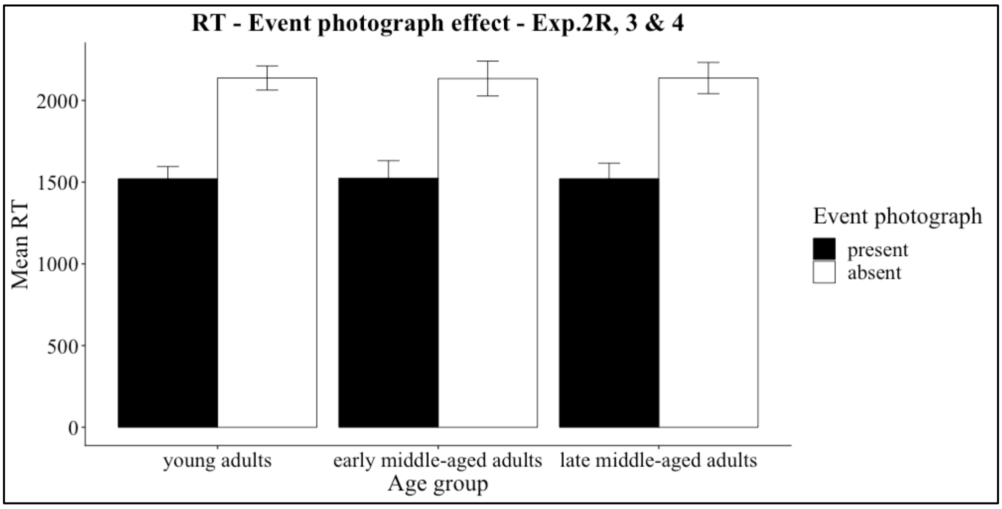


Figure 43a: The main effect of event photographs on speed in L2 vocabulary learning in three adult learner groups

The differences in participants' reaction times were strongly influenced by event depictions. In Figure 43b, we can see clearly that the late-middle-aged adults had the longest mean reaction times, both when events were present and when they were absent. By contrast, young adults had the shortest mean reaction times in the two event learning conditions. The early middle-aged adults had longer mean reaction times in comparison with young adults, but they had shorter mean reaction times compared to late middle-aged adults.

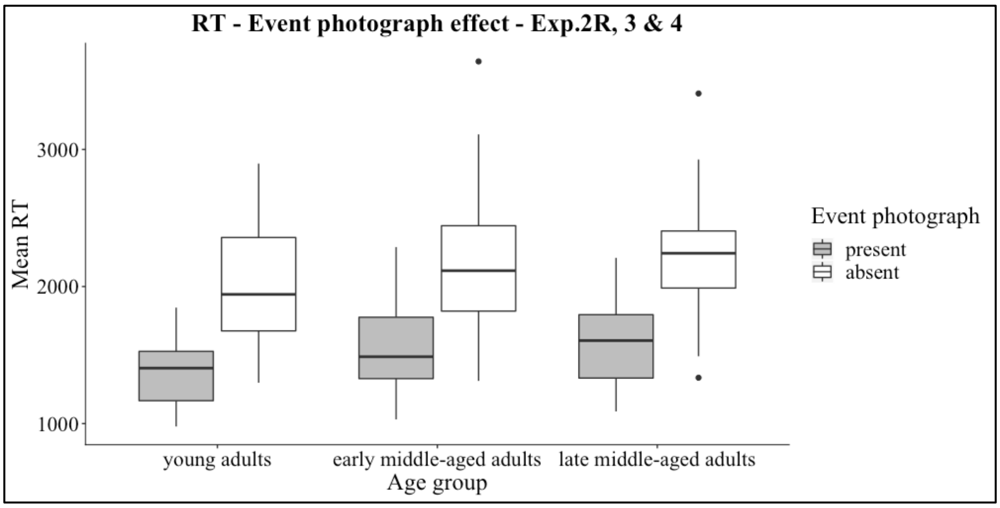


Figure 43b: Mean of reaction time per event photograph condition in three adult groups

We performed a Pearson correlation test to investigate whether learners' age and their speed of responding were significantly correlated. The output showed a correlation coefficient of .26 and p- = .01. Figure 43c depicts the significant correlation indicating that the older the learners, the longer their reaction times.

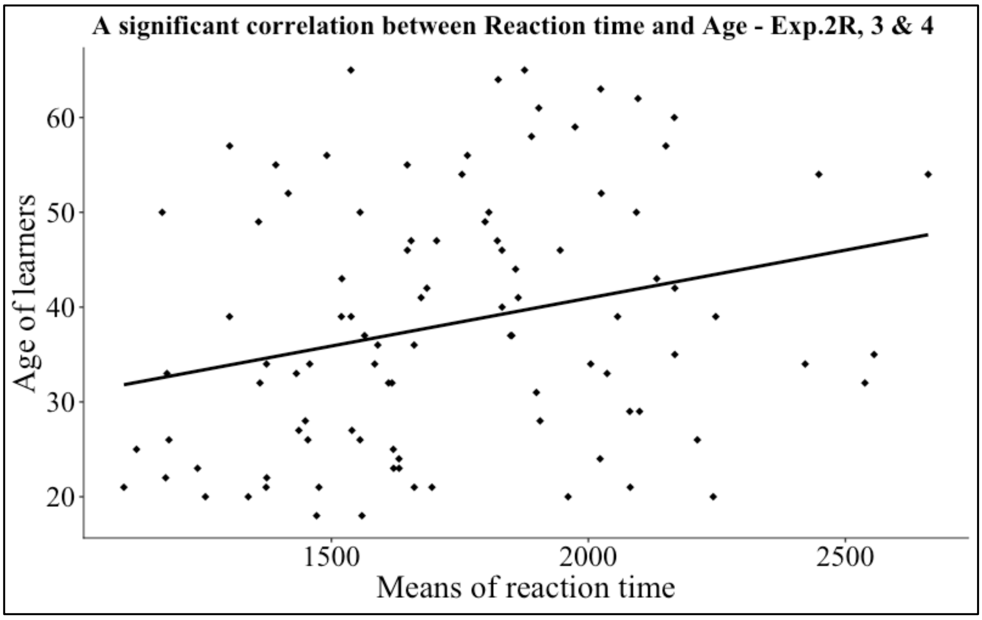


Figure 43c: A significant correlation between Age and Reaction time (exp. 2R, 3 & 4). The means of learners' age are displayed on the y-axis, means of reaction time are displayed on the x-axis. The trend line shows the line of best fit.

9.5.3 Discussion

To conclude, all three age groups could successfully use event-present photographs (compared to event-absent ones) to determine "what happened to the object learned before an action," and then learn L2 phrases. Also, L2 beginners in all age groups did not benefit from L1–L2 similar language mapping (compared to L1–L2 different language mapping) for their L2 vocabulary learning success. Learners' cognitive ability (measured by cognitive test scores) marginally, yet significantly, influenced how quickly they could relate a spoken phrase to an event depiction. Learners' age, as another factor in L2 learning, can powerfully affect L2 learning results. Overall, young adults showed the highest accuracy rates and the shortest reaction times. In contrast, late middle-aged adults had the lowest accuracy, and there were no significant differences in reaction times between late middle-aged adults and early middle-aged adults.

After conducting four experiments (Experiments 2, 2R, 3, and 4), we found very robust significant effects of event photograph presence in short-term L2 vocabulary learning. These effects were investigated not only for young adults but also for early middle-aged adults and even for late middle-aged adults. We were further motivated to investigate whether these effects of event depictions would be replicated in more difficult L2 vocabulary learning and testing situations. Therefore, we conducted one more experiment, Experiment 2N, in which we removed the learning repetition part and added one more testing part as compared to the previous experimental design in Experiments 2, 2R, 3, and 4. In Section 10, we present the new L2 learning design and findings.

An additional motivation for carrying Experiment 2N came from studies of single-shot learning in linguistic contexts. Borovsky, Kutas and Elman (2010) examined the impact of initial learning context on the perception of new word usage. They used event related brain potentials and tested twenty-six English native English speakers (18-25 years) in four different linguistic conditions. Participants saw a word (unknown vs. known) in one of constraining sentence contexts (high vs. low) before giving their plausibility ratings of the word as the object of transitive verbs. For instance, an unknown word (e.g., *MARF*) was seen in the high sentence context (e.g., *He tried to put the pieces of the broken plate back together with MARF*) or in the low one (e.g., *She walked across the room to Mike's messy desk to return his MARF*). Then, they provided their plausibility (plausible or implausible) ratings via two test sentences (e.g., *He needed the MARF; He greeted the MARF* or *She used the MARF; She drove the MARF*). Results showed that plausibility effects were only found in a highly constraining context. The finding also suggested that the contextual constraint could modulate rapid word learning with single-shot learning. The question was whether L2 beginners would successfully learn L2 vocabulary with single-shot learning in non-linguistic context.

10. Second Language Vocabulary Learning: Learning with Single Exposure and Learning with vs. without Repetition

10.1 Experiment 2N

10.1.1 Participants

Thirty-two additional adults aged between 18 and 31 years (Mean: 23.78, SD:3.44) participated in Experiment 2N. All participants were monolingual native speakers of German. Young adult participants were recruited from LingEx, the Humboldt University mailing list for Bachelor students. Testing took place in the reaction time laboratory of the Psycholinguistic Group at Humboldt University. Participants received €11 for their participation, and testing took approximately 60 minutes in total.

10.1.2 Materials and Procedure

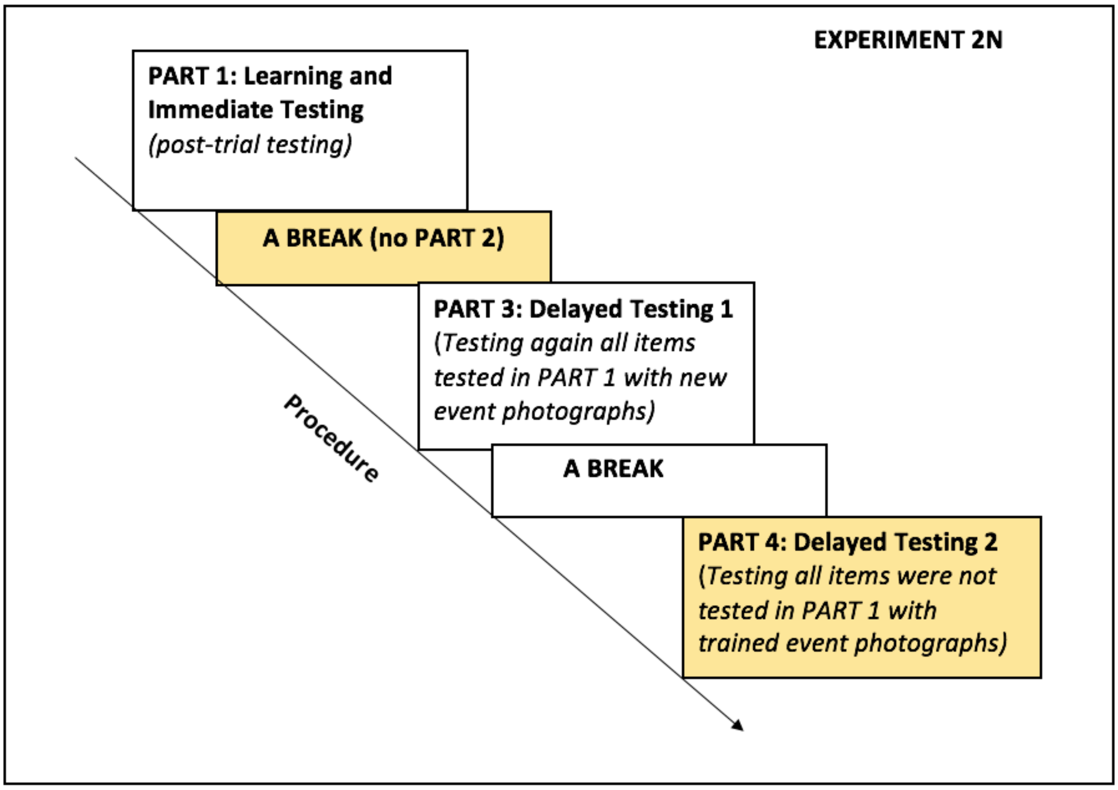


Figure 44: The procedure for Experiment 2N

We conducted Experiment 2N with the same materials as in Experiment 2R. The procedure of Experiment 2R was not identical to that of previous experiments (2, 2R, 3 and 4). Compared to Experiment 2 (see more at 8.2.3), the main learning experiment had some changes. Part 1 and Part 3 remained as in previous experiments, but there was no Part 2 in Experiment 2N (i.e., we kept Part 2 as a short break of around 2 minutes in Experiment 2N). An additional part (Part 4) was included as a second delayed test. Figure 44 shows the main procedure of Experiment 2N. All changes are marked with an orange background color.

10.1.3 Predictions

Because we also tested young adult participants in Experiment 2N, we based our predictions on the results of Experiment 2 and Experiment 2R:

Predictions for the cognitive test (WAIS scores): Participants with higher scores in the cognitive test would do better in the L2 learning experiment (higher accuracy and shorter reaction times) than others.

Predictions for results of L2 vocabulary learning success:
If events are useful in learning, then participants could give more accurate and faster responses in total/each testing part via a button press in the “event-present” condition compared with the “event-absent” condition (as in Experiment 2 and Experiment 2R).

If there is a strong effect of testing part, then participants could give more accurate and faster responses via a button press in Part 1 compared with Part 3 because participants have to process new photographs referring to the same verb-noun phrases as in Part 1 (as in Experiment 2 and Experiment 2R).

Experiment 2N was conducted to investigate how event presences (present vs. absent) can work with participants’ short-term memory in difficult testing situations and whether they can maintain effects on L2 learning success. Therefore, we set up a delayed testing part (Part 3) without another repeated training part (as in Part 2 in Experiment 2).

If the event effects differ depending on testing parts, then we should see a significant interaction between visual context (event photograph depictions: present vs. absent) and testing part (1 and 3) on participants’ phrase learning (as the results for Experiment 2 and Experiment 2R). That means participants who inspected event photographs performed their tasks better in Part 1 (i.e., shorter reaction times, more correct answers) than in Part 3. However, participants who did not inspect event photographs performed their tasks equally well in Part 1 and Part 3 (approximately 50% correct answers and similar reaction times).

If the event depictions can extend their effects to L2 learning with single exposures, then in Part 4, participants could still give more accurate and faster responses in event-present conditions compared with the event-absent ones when we tested half of the items (16) which were not tested in the previous two parts (the new prediction).

We did not predict language mapping effects and significant interactions related to the language mapping factor because results from previous studies (Experiments 2 and 2R) showed no main effect. Before analyzing data from Experiment 2N (in February 2020), we added further predictions:

For Experiment 2N, if the order of testing part in L2 vocabulary learning plays a vital role in participants' L2 learning success, then learners are more accurate and faster in Part 1, immediate testing, and Part 3—the first delayed testing with the support of attention and memory than in Part 4—the second delayed testing.

If L2 vocabulary learning with repetition can help participants to retain or improve their L2 learning results, then learners are more accurate and faster in Part 3 of Experiment 2R (with repetition of learning all items before delayed testing) than in Part 3 of Experiment 2N (without any re-learning before delayed testing).

10.1.4 Results

With the same analysis methods and settings as in previous experiments, we obtained the following results.

10.1.4.1 Accuracy Results

10.1.4.1.1. With Testing Part as a Factor in the Analysis Model

With testing part included as a factor, we analyzed Experiment 2N with three different data sets in the same analysis model:

- Data set 1: Part 1 and Part 3 - The data structure was the same as in previous experiments, meaning that we can compare results to know which effects were replicated.
- Data set 2: Part 1 and Part 4 - We wanted to investigate the event depiction effect with more prolonged delayed testing.
- Data set 3: Part 3 and Part 4 - We wanted to understand how delayed testing would affect the learning outcome (i.e., comparing a single learning exposure and learning with a repetition).

a. Data set 1

Results replicated (see Table 26) a few main effects from Experiment 2R. They were the main effect of event presence ($p < .001$) in both testing parts, the main effect of the testing part ($p < .001$), and a significant event presence by testing part interaction ($p < .001$).

Table 26: Results from generalized linear mixed model fit by maximum likelihood for accuracy in Experiment 2N (Part 1 and Part 3)

Term	**Estimate**	**SE**	**z value**	**p**
Intercept	1.183	0.193	6.1	<0.001***
Event presence	1.267	0.109	11.54	<0.001***
Language mapping	-0.068	0.184	-0.37	0.7115
Testing Part	0.53	0.099	5.264	<0.001***
Event presence* Testing Part	0.475	0.098	4.805	<0.001***
Event presence* Language mapping	0.23	0.107	2.155	0.0312*
Event presence*Language mapping*Part	0.13	0.098	1.359	0.1743

A significant interaction between event presence and language mapping ($p = 0.0312$) on learners' accuracy was also found (see Figure 45). It indicated that learners made approximately the same number of correct choices when event photographs were depicted, both in similar language mapping conditions and different ones. However, when events were not depicted in L2 learning, participants responded much more correctly in different language mapping conditions than in similar ones (this was unexpected).

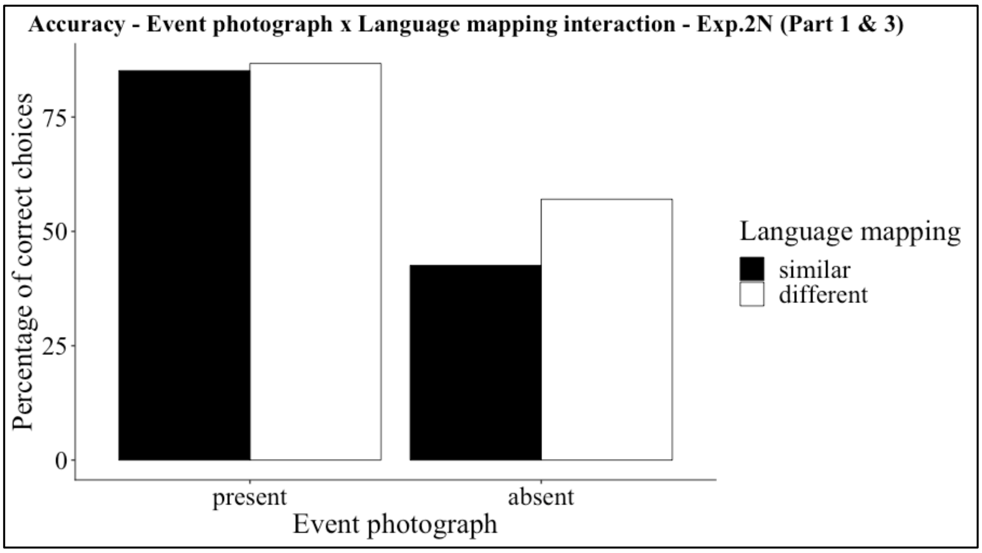


Figure 45: The interaction between event photograph and language mapping in the accuracy data in Experiment 2N

b. Data set 2

The results of this data set are presented in Table 27. We again investigated the main effect of event presence ($p < .001$) in both testing parts, the main effect of the testing part ($p < .001$), and a significant event presence by testing part interaction ($p < .001$).

Table 27: Results from generalized linear mixed model fit by maximum likelihood for accuracy in Experiment 2N (Part 1 and Part 4)

Term	Estimate	SE	z value	p
Intercept	1.335	0.144	9.24	<0.001***
Event presence	1.339	0.104	12.79	<0.001***
Language mapping	-0.067	0.139	-0.48	0.6307
Testing Part	0.25	0.101	2.534	0.0113*

Event presence* Testing Part	0.246	0.104	2.369	0.0179*
Event presence* Language mapping	0.23	0.102	2.293	0.0219*
Event presence*Language mapping*Part	0.043	0.103	0.423	0.6720

A significant interaction between event presence and language mapping (p = 0.0219) was replicated with the same pattern as in the data set 1 (see Figure 46).

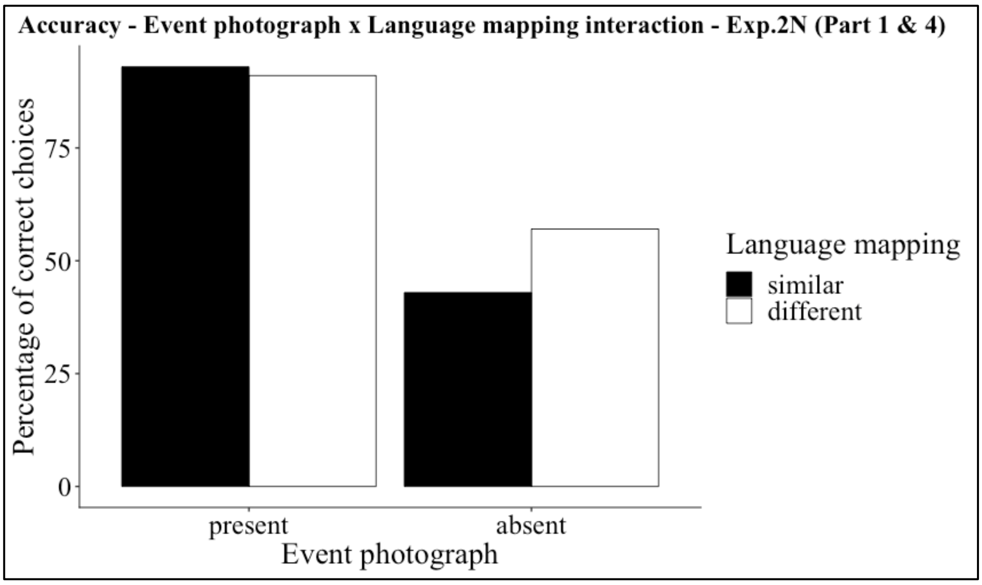


Figure 46: The interaction between event photograph and language mapping in the accuracy data in Experiment 2N

c. Data set 3

The accuracy results of the data set 3 are listed in Table 28. A main effect of event presence (p < .001) in both testing parts, a main effect of testing part (p < .001), and a significant event presence by testing part interaction (p < .001) were investigated in the data set. No significant interaction between event presence and language mapping was found (p = .1).

Table 28: Results from generalized linear mixed model fit by maximum likelihood for accuracy in Experiment 2N (Part 3 and Part 4)

Term	Estimate	SE	z value	p
Intercept	0.863	0.149	5.79	<0.001***
Event presence	0.921	0.081	11.26	<0.001***

Language mapping	-0.077	0.146	-0.52	0.5987
Testing Part	-0.23	0.079	-2.995	0.0027*
Event presence* Testing Part	-0.18	0.081	-2.209	0.0271*
Event presence* Language mapping	0.13	0.080	1.618	0.1056
Event presence*Language mapping*Part	-0.073	0.081	-0.896	0.3701

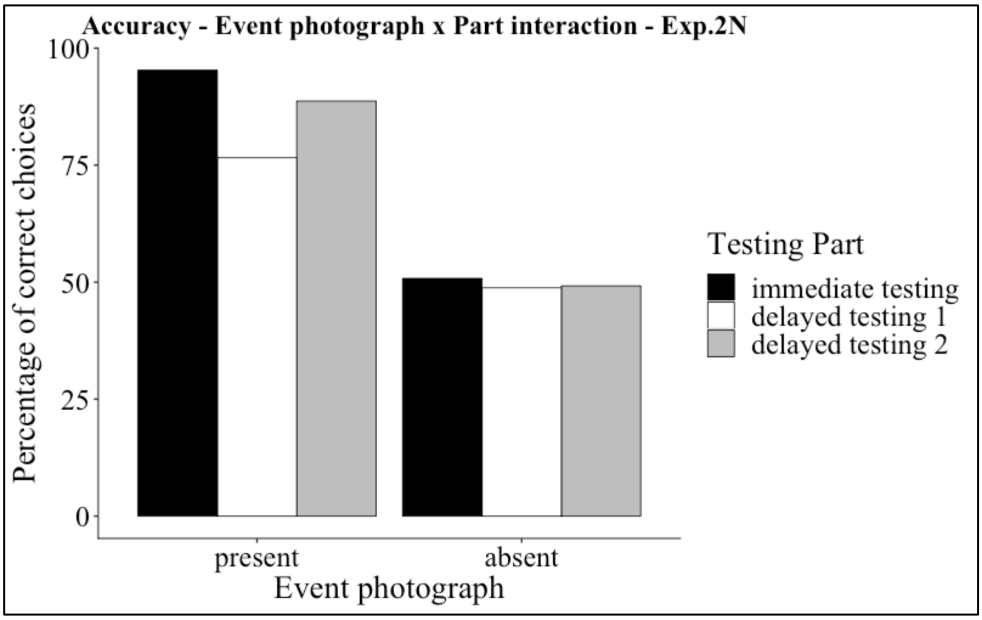


Figure 47: The interaction between event photograph and testing part in the accuracy data in Experiment 2N

From Figure 47, we can visualize a main effect of present event photographs in all three different testing parts and the significant event depiction by testing part interactions. For event-present learning conditions, learners had higher accuracy in immediate testing (95.3% correct choices) than in delayed testing 1 (76.6% correct choices) and in delayed testing 2 (88.7% correct choices) as expected. Unexpectedly, learners responded more accurately in delayed testing 2 than in delayed testing 1. We thought that participants would be more accurate in delayed testing 1 than in delayed testing 2. Many learned phrases were tested for the second time in delayed testing 1, and other phrases were examined for the first time in delayed testing 2. However, the learners performed better in delayed testing 2 than in delayed testing 1. That might be because the same event photographs viewed in Part 1 were tested in delayed testing 2, while new event photographs referred to learned phrases in delayed testing 1. For event-absent learning conditions, there were no significant differences in how many correct choices were made among the three testing parts. Participants made around 50% correct choices for each testing part. Without event photographs, learners already know about an object. However, they had no idea about the action that happened to the object as presented via the audio sound. Therefore, their responses could be guesses.

10.1.4.1.2. Without Testing Part in the Analysis Model

Without the testing part, we replicated a significant effect of event depictions on participants' accuracy, as in Experiment 2R ($p < 0.001$). In the event-present learning conditions (in total 768 event-present trials), 85.9% correct choices were made. Participants were much more accurate than in the event-absent learning conditions, with 49.8.2% correct choices for 768 event-absent trials (see Figure 48).

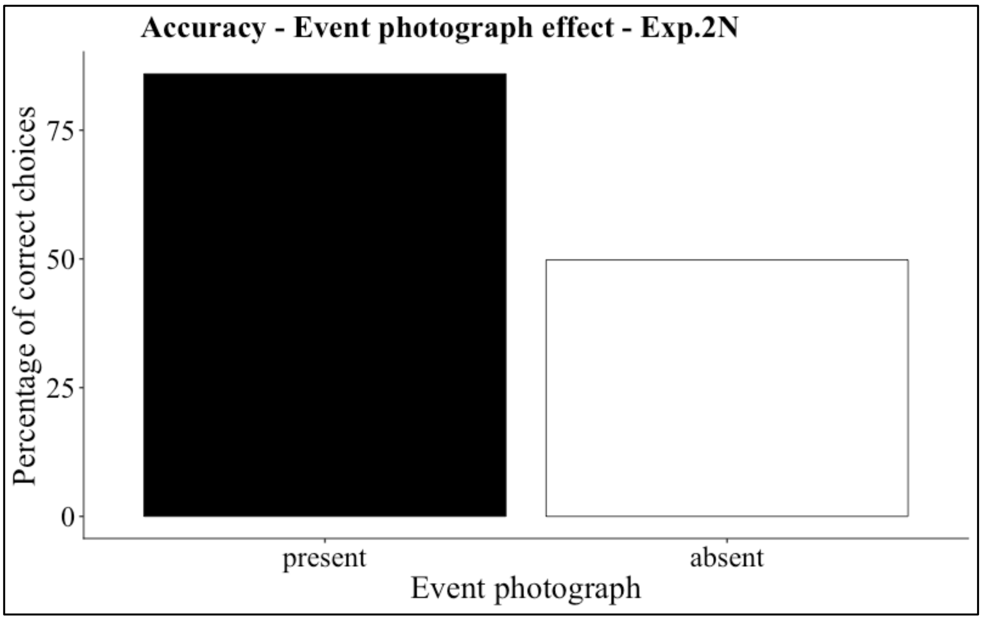


Figure 48: The effect of event photograph on the accuracy in Experiment 2N

10.1.4.2 Reaction Time Results

10.1.4.2.1. With Testing Part as a Factor in the Analysis Model

With the testing part included as a factor, we analyzed the data from Experiment 2N in three different data sets as in the accuracy analysis of the same linear mixed model.

a. Data set 1

All findings from the linear mixed-effects model, including the testing part as a factor, are shown in Table 29. As predicted, the analysis model replicated a significant event presence effect (SE = 0.01, = 13.46, p < 0.001) on participants' reaction time. That establishes that learners were faster in event-present conditions than in event-absent conditions when they responded precisely in testing. Furthermore, a significant effect of the testing part was discovered (SE = 0.01, = 8.07, p < 0.001), when participants made correct choices more quickly in immediate testing than in delayed testing. A significant event presence by testing part interaction (SE = 0.01, = 10.26, p < 0.001) was investigated to show that learners took longer in delayed testing than in immediate testing when responding to event-present conditions. For event-absent conditions, they needed a similar reaction time in both testing parts. Other interactions (event presence by language mapping, language mapping by part, or event presence by language mapping by part) were non-significant.

Table 29: Results from linear mixed effects model for reaction time in Experiment 2N (Part 1 & Part 3)

Term	**Estimate**	**SE**	**df**	**|t|**	**p**
Intercept	7.460	0.038	42	192.47	<0.001***
Event presence	-0.151	0.012	602	12.42	<0.001***
Language mapping	-0-010	0.020	31	0.52	0.6
Testing Part	-0.080	0.020	33	3.97	<0.001***
Event presence* Testing Part	-0.102	0.011	588	8.78	<0.001***
Language mapping* Testing Part	0.018	0.011	583	1.62	0.103
Event presence*Language mapping	0.007	0.021	600	0.63	0.52
Event presence*Language mapping* Testing Part	0.016	0.011	580	1.40	0.16

When testing part was not included in the analysis model, a significant effect only for the event presence factor (SE = 0.01, = 11.63, p < 0.001) on participants' speed was observed. That result confirmed that learners performed much faster in depicted event conditions than in undepicted ones.

b. Data set 2

The analysis with testing part as an experimental factor showed a significant effect of event presence (SE = 0.016, = 12.76, $p < 0.001$) on participants' reaction times. Participants responded faster for event-present items in both testing parts than for event-absent items. Also, a significant effect of testing part (SE = 0.02, = 7.76, $p < 0.001$) was observed in which participants were much slower to respond in Part 4 (delayed testing 2) than in Part 1 (immediate testing) for the two event depiction conditions. Last, we found a significant event presence by testing part interaction (SE = 0.011, = 5.49, $p < 0.001$). That result indicated that L2 learners in both testing parts gave faster responses for event-present items than for event-absent items. In both testing parts, participants were tested with the same event photographs that they had viewed before. However, the mean reaction time for both event presence conditions increased significantly in Part 4 (delayed testing 2) in relation to Part 1 (immediate testing). The results of data set 2 are included in Table 30.

Table 30: Results from linear mixed effects model for reaction time in Experiment 2N (Part 1 & Part 4)

Term	Estimate	SE	df	|t|	p
Intercept	7.356	0.039	37	187	<0.001***
Event presence	-0.216	0.016	25	12.76	<0.001***
Language mapping	0.024	0.017	28	1.43	0.1635
Testing Part	-0.162	0.02	31	7.76	<0.001***
Event presence* Testing Part	-0.063	0.011	629	5.498	<0.001***
Language mapping* Testing Part	0.004	0.011	608	0.377	0.706
Event presence*Language mapping	0.009	0.012	28	1.748	0.4607
Event presence*Language mapping* Testing Part	0.02	0.011	635	1.845	0.0655

Another analysis not including testing part as an experimental factor resulted in a main effect of event depictions. Overall, participants' responses were much slower when events were not depicted than when they were depicted.

c. Data set 3

With testing part included in the analysis model for the data set, we found a significant main effect of event depiction and a significant interaction between event depiction and testing part. However, no significant effect of testing part was investigated as in the two other data sets. All findings from the linear mixed-effects model, including testing part as a factor, are shown in Table 31. For the significant effect of event presence (SE = 0.01, = 8.38, $p < 0.001$) on participants' reaction time, learners responded accurately and more quickly in event-present conditions than in event-absent conditions. A significant event presence by testing part interaction (SE = 0.01, = 4.85, $p < 0.001$) was also found. That result indicated that learners, for event-present items, were much faster in Part 4 (the second delayed test with the same event pictures as in the learning part) than in Part 3 (the first delayed test with different event pictures compared to the event pictures in the learning part). However, for event-absent items, they were slower in Part 4 than in Part 3.

Table 31: Results from linear mixed effects model for reaction time in Experiment 2N (Part 3 & Part 4)

Term	Estimate	SE	df	|t|	p
Intercept	7.498	0.041	38.3	182.05	<0.001***
Event presence	-0.092	0.01	596	8.38	<0.001***
Language mapping	0.016	0.017	29	0.927	0.361
Testing Part	-0.014	0.128	35	1.15	0.257
Event presence* Testing Part	0.05	0.01	601	4.885	<0.001***
Language mapping* Testing Part	-0.003	0.01	581	0.28	0.779
Event presence*Language mapping	-0.01	0.01	600	0.992	0.321
Event presence*Language mapping* Testing Part	-0.003	0.011	604	0.347	0.729

Without including testing part as a factor in the analysis model, we perceived a significant effect of the event presence factor (SE = 0.01, = 8.27, $p < 0.001$) on participants' reaction time. That replicated that learners were much faster in event-present conditions compared to absent-event conditions.

Taking three data sets together, Figure 49 illustrates the main effect of event-present conditions per testing part and the significant interactions between event depiction and testing part in data sets 1 (immediate testing and delayed testing 1) and 2 (immediate testing and delayed testing 2).

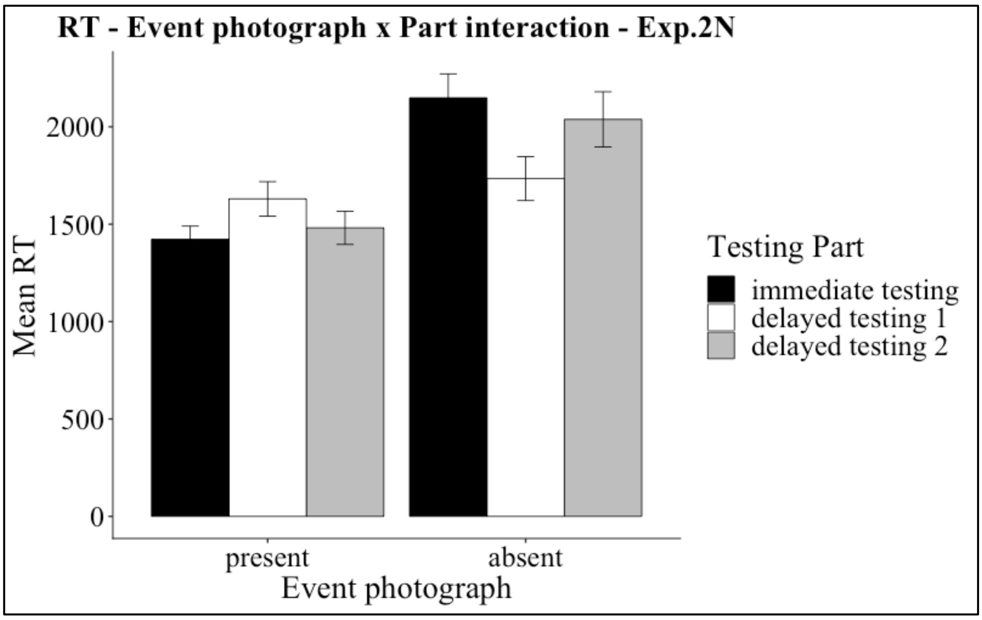


Figure 49: The interaction between event photograph and part in the reaction time data of Experiment 2N (error bars represent 95% CI)

10.1.4.2.2. With Testing Part as a Factor in the Analysis Model

We also analyzed the reaction times in all data in the three testing parts, but the testing part was not set up as an experimental factor. The result corroborated only one significant effect of event presence on reaction time (see Figure 50) as in the previous experiments (SE = 0.014, = 13.85, $p < 0.001$). We found no additional effects involving the manipulated factors.

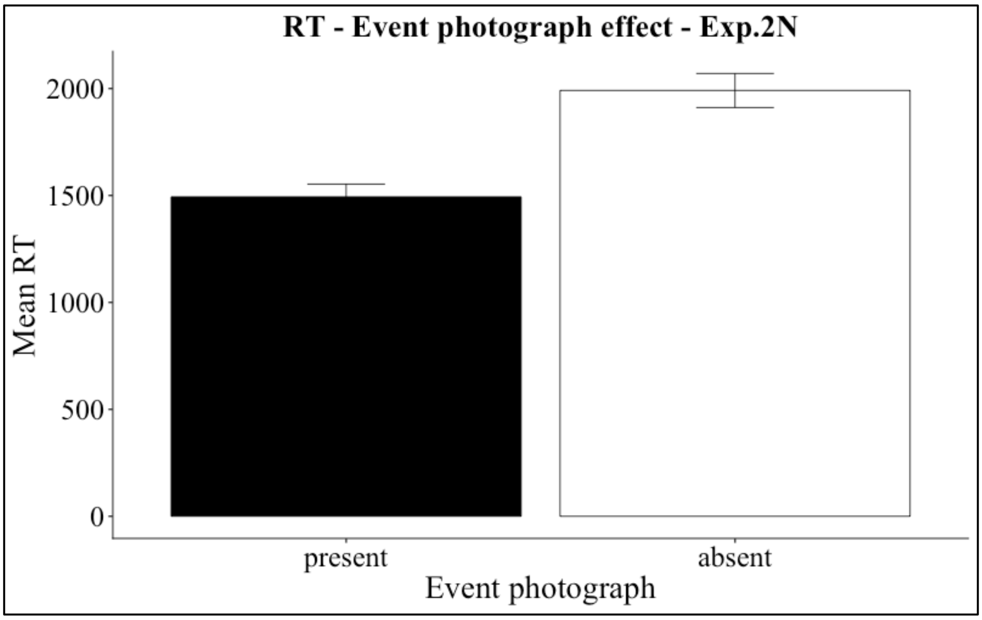


Figure 50: The effect of event photograph on the reaction time in Experiment 2N (error bars represent 95% CI)

10.1.4.3 Cognitive Test Results

Young adults' cognitive test scores in Experiment 2N are shown in Figure 51. As shown in this graph, they performed best in the similarities task and worst in the digit span task (the results are replicated from Experiment 2N). The mean for the verbal fluency task was 41.625 (SE = 8.09) words generated (approximately the same as in Experiment 2).

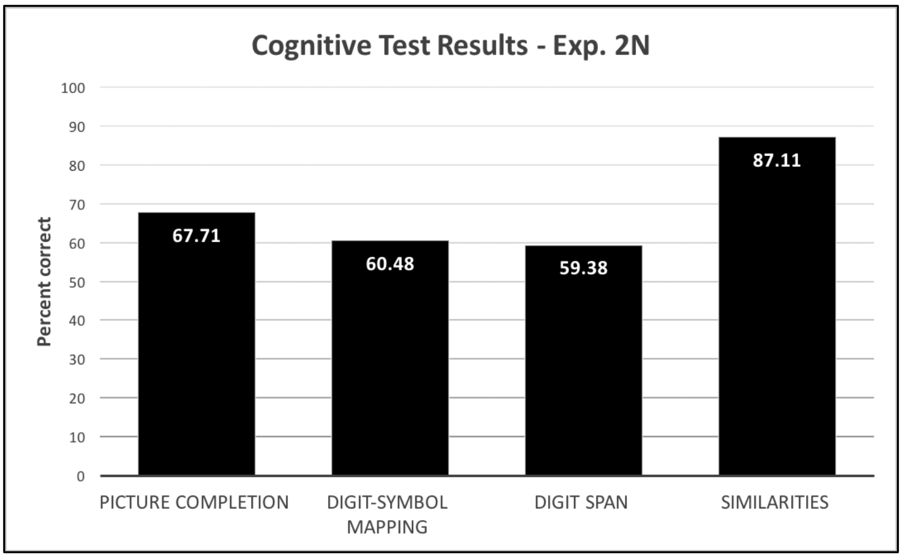


Figure 51: Cognitive test scores for the WAIS test (exp. 2N). The y-axis displays the percentage of correct answers averaged across participants. The percentages are shown in the center of each bar. Note that the verbal fluency test scores are not depicted. Since the task was free naming, there is no upper limit that can be reached.

To examine whether accuracy results and the cognitive test scores correlate, we computed a Pearson correlation test because the two variables were normally distributed. The result indicated that learners' cognitive test scores and their accuracy scores in L2 vocabulary learning were not significantly correlated (a correlation coefficient of -.08, p-value = .656). Figure 52a illustrates no correlation, but it suggests the same tendency as in Experiment 2R that the lower the cognitive scores, the higher the accuracy scores.

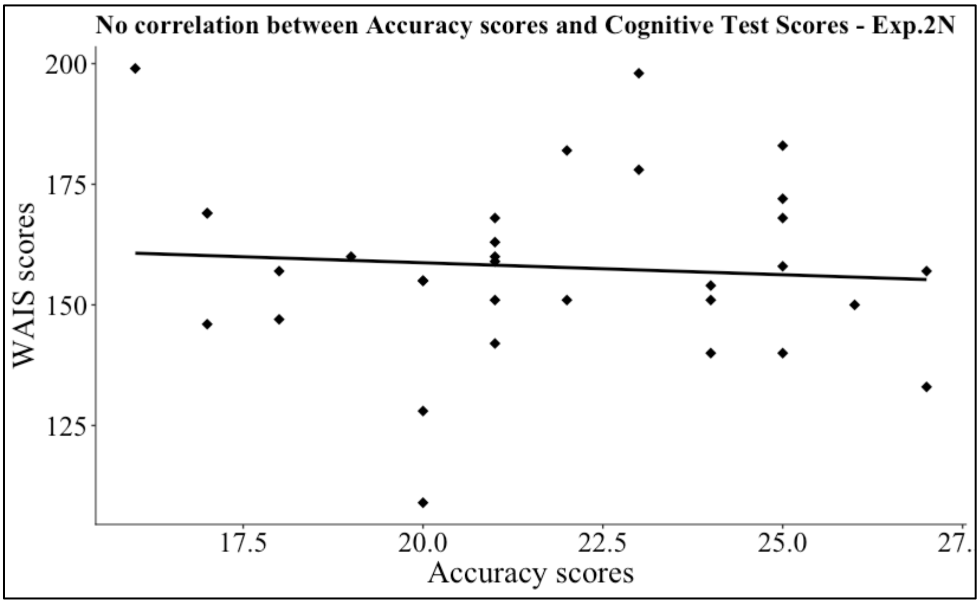


Figure 52a: No correlation between WAIS and accuracy scores (exp. 2N). WAIS scores (including the verbal fluency task scores) are displayed on the y-axis, accuracy scores are displayed on the x-axis. The trend line shows the line of best fit.

Learners' WAIS scores and their response times in L2 learning were significantly correlated in Experiment 2N, with a correlation coefficient of -.45 and p-value < .008. Figure 52b shows the significant correlation with a clear trend that the higher the cognitive scores, the faster the response speed for correct choices (as in Experiment 2R).

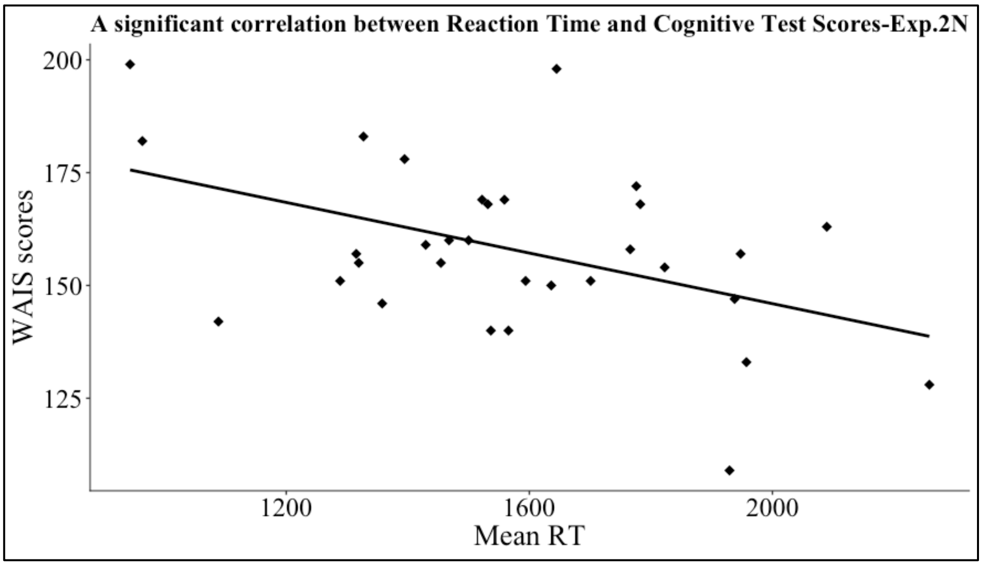


Figure 52b: A significant correlation between Reaction Time and Cognitive Test Scores (exp. 2N). WAIS scores (including the verbal fluency task scores) are displayed on the y-axis, accuracy scores are displayed on the x-axis. The trend line shows the line of best fit.

10.1.5 Discussion for Experiment 2N

Experiment 2N was the last experiment in our set of six experiments. The results of Experiment 2N with two first testing parts (Part 1 and Part 3) replicated all significant main effects on L2 vocabulary success (i.e., accuracy scores and means of reaction times) of previous studies. They comprised a main effect of the event depiction factor, a main effect of the testing part (when the testing part was included in analysis models), and a significant event depiction when testing part interaction. Also, no main effect of language mapping was found. These findings confirmed the robustly significant effect of event depictions and testing parts on beginners' L2 vocabulary learning success.

Part 4 (delayed testing 2) was an additional part of Experiment 2N as compared to previous experiments. The results of this part confirmed the robust significant effect of event depictions in L2 short-term learning. L2 beginners responded more accurately and more quickly for event-present items (vs. event-absent items) that they had learned with single exposures in the first part, and on which they had not been tested in two previous testing parts (i.e., Part 1 and Part 3).

Another investigation was also related to event pictures. L2 beginners responded more accurately to same event pictures (e.g., A is reading a book) in testing Part 1 and testing Part 4 (e.g., B is cleaning a table), which they had inspected and learned before than they responded for new event photographs (e.g., C is a reading book) in Part 3.

The cognitive test scores of young adults in Experiment 2N only correlated significantly with participants' speed (as in Experiment 2R).

We made a further comparison between Experiment 2N and Experiment 2R in the next section (see 10.2). We wanted to know whether young adults as L2 beginners, in L2 vocabulary learning, were more accurate and faster in Part 3 (delayed testing) of Experiment 2R (with learning repetition) than in Part 3 (delayed testing 1) of Experiment 2N (with single learning exposure).

10.2 A Comparison between Experiments 2N and 2R

Before comparing the L2 learning success of young adults between two experiments, we estimated the cognitive test scores of these groups. Table 33 lists the means and standard deviation of all subtests in the cognitive test for participants in Experiment 2R and 2N. Overall, no significant difference in the cognitive test results was found between the two young adult groups. In other words, the cognitive test results of Experiment 2N replicated all results of Experiment 2R.

Table 33: Cognitive test results and demographic characteristics of young adults

Characteristic	**Experiment 2R**	**Experiment 2N**
Age range	18–31	18–31
Mean age in years	23.62 (3.42)	23.78 (3.44)
Picture completion[a]	4.18 (0.93)	4.06 (0.76)
Digit Symbol[a]	82.94 (13.53)	80.44 (14.3)
Digit Span[a]	17.59 (2.77)	17.81 (3.44)
Similarities[a]	12.78 (1.86)	13.94 (1.41)
Vocabulary (Animal naming[b] and Word naming[c])	41.62 (7.29)	41.62 (8.09)
Total scores of the cognitive test	159.12 (18.43)	157.87 (18.74)
BMIS	7.62 (0.83)	7.26 (1.18)
Male/female (n)		10/22

Note.

[a]latest German version of the Wechsler Adult Intelligence Scale (WAIS), [79-80].

[b]Task: Name as many animals as possible, time allowed: 1 min.

[c]Task: Name as many words as possible starting with the letter "l", time allowed: 1 min.

Experiment 2N replicated all the effects of event photographs in previous experiments. We investigated the differences in L2 learning success between learning with and without repetition. There were no significant differences in accuracy scores for event-present items in both the immediate and delayed testing of Experiments 2R and 2N (see Table 34). Without learning repetition, young adults in Experiment 2N still made 76.6% correct choices for event-present items in delayed testing. We expected young adults in Experiment 2N to have much higher accuracy because they had one more learning chance before delayed testing. However, they had 77.34% correct responses, not much more than the participants in Experiment 2N. Furthermore, for the accuracy, Experiment 2N replicated all results from Experiment 2R (i.e., approximately the same percentage of correct choices per learning condition per testing part).

Table 34: Second language vocabulary learning with vs. without a repetition in Experiments 2R and 2N

Experiment	**Event photograph**	**Testing part**	**Correct choices (%)**
2R (with a repetition learning before delayed testing)	Event present	Immediate	96.1
	Event absent		50.4
	Event present	Delayed	*77.34*
	Event absent		50
2N (without a repetition learning before delayed testing)	Event present	Immediate	95.3
	Event absent		50.8
	Event present	Delayed	*76.6*
	Event absent		48.8

We also wanted to investigate whether learners' overall accuracy scores were correlated with their cognitive test scores. We performed a Pearson correlation test with the data from 64 young adults in Experiments 2R and 2N. The results indicated that learners' cognitive test scores and their accuracy scores in L2 vocabulary learning were not significantly correlated (a correlation coefficient of -.18, p-value = 0.1446). Figure 53a represents no correlation, but it indicates a tendency that the lower the cognitive scores, the higher the accuracy scores.

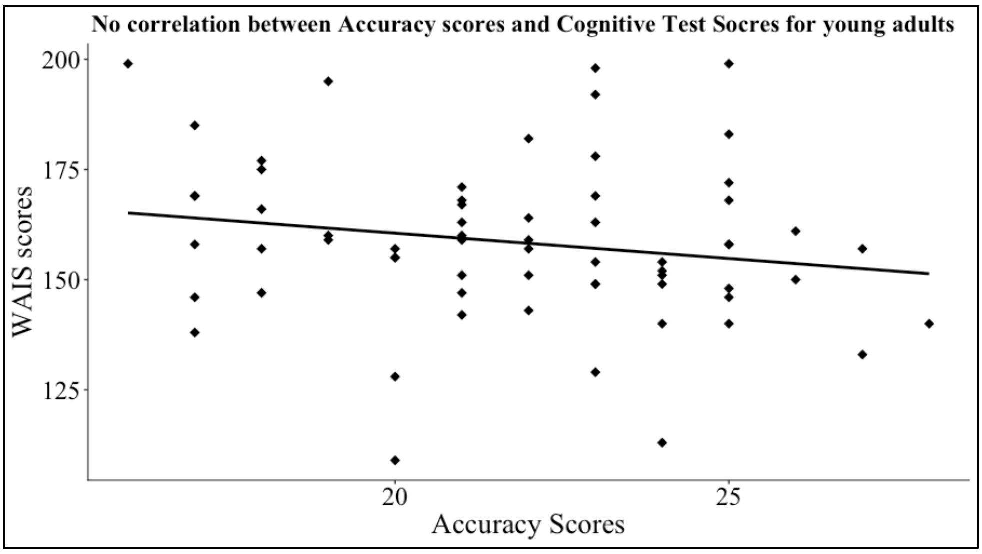


Figure 53a: No correlation between cognitive test scores and accuracy scores of young adults (exp. 2R and 2N with 64 participants). WAIS scores (including the verbal fluency task scores) are displayed on the y-axis, accuracy scores are displayed on the x-axis. The trend line shows the line of best fit.

The response speed of young adult learners' (data from 64 participants) in L2 vocabulary learning was significantly correlated with their cognitive test scores (a correlation coefficient of -.43, p-value = 0.0003). Figure 53b shows a significant correlation indicating that participants with higher cognitive test scores responded much faster than others.

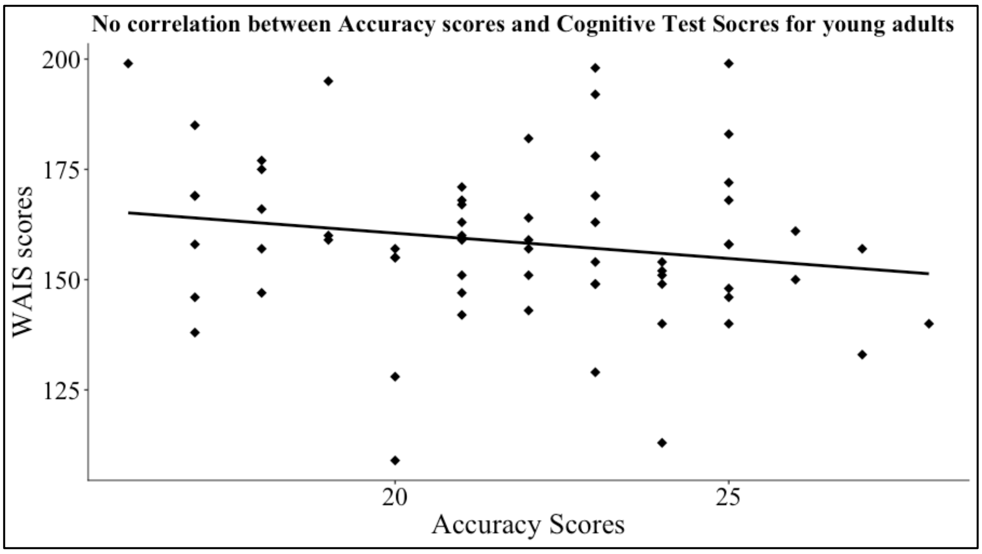


Figure 53b: A significant correlation between cognitive test scores and reaction time of young adults (exp. 2R and 2N, 64 participants). WAIS scores (including the verbal fluency task scores) are displayed on the y-axis, means of reaction time are displayed on the x-axis. The trend line shows the line of best fit

The main new finding in Experiment 2N in comparison with Experiment 2R was that a learning repetition for L2 beginners in a short-term L2 vocabulary learning did not help them to perform more successfully than a single learning exposure did. That might be because of the effects of human short-term working memory.

11. General Discussion and Conclusion

Six reaction time L2 learning studies investigated the extent to which L2 adult beginners can use depicted event photographs and similarity in L1–L2 mappings to successfully learn L2 phrasal vocabulary (i.e., verb-noun phrases). In addition, we examined whether age (young adults vs. early middle-aged adults vs. late middle-aged adults) and participants' cognitive ability can be predictors of L2 phrasal learning success in a visual context. In this section, we discuss the main findings of our experiments.

11.1 Effects of Event Photographs in Second Language Phrasal Vocabulary Learning

In previous studies of L2 language learning, L2 beginners usually started learning L2 single nouns via a word-picture pair or a word-referent pair. In Section 2.2.1, we reviewed studies showing that L2 learners could successfully recall or recognize an L2 noun (i.e., name or object) after learning it via a combination of a spoken form and a corresponding referent (i.e., picture). Our research results provided more evidence for referent effects (i.e., depicted object and event photographs) on L2 vocabulary learning. Moreover, we replicated and extended findings by Nassaji (2012) and Havas et al. (2017). They found an advantage of depicted referents (i.e., auditory-visual learning conditions) over no referents (i.e., auditory only) when adults learned L2 single words.

Across all six experiments in our research project, participants used depicted events (vs. non-depicted events) extensively for L2 vocabulary learning. Furthermore, this effect was anticipatory, meaning that all L2 beginners, including young adults, early middle-aged adults, and late middle-aged adults in the immediate testing responded much more accurately and quickly when they had inspected event photographs in learning trials (vs. they had not inspected event photographs before). Even though the event photographs were different in the delayed testing compared to the learning phase, all three learner groups retained their learning success (i.e., higher accuracy and shorter reaction times) in event-present conditions than in event-absent conditions.

For event-present conditions, the correct responses of young adult learners from 18 to 31 years old (Experiments 2, 2R, and 2N) were around 95% in the immediate testing and around 75%

in the delayed testing. Young adults had the highest accuracy for both testing parts when they had seen event photographs before. Early middle-aged adults aged from 32 to 45 learned L2 phrasal vocabulary successfully with the same approximate percentage of correct choices as young adults. The correct responses of late middle-aged adults aged from 46 to 65 were 10% lower than the two younger learner groups (i.e., 84.4% in immediate testing; 67.6% in delayed testing), but that still means that event photographs were beneficial in learning.

For event-absent conditions, young adults and early middle-aged adults made around 50% correct choices in both testing parts, as we predicted. Late middle-aged adults seemed confused in the learning conditions because they responded accurately to about 38% in both testing parts.

For reaction times, we found the same patterns across experiments for both testing parts. In immediate testing, learners responded much faster for present event conditions than in the absence of event conditions. However, participants took much longer to respond accurately in both learning conditions in delayed testing. The distance in the means for reaction time between event-present conditions and event-absent ones was narrow. To summarize our research findings, L2 adult beginners could learn successfully L2 verb-noun phrases in a visual context with supportive event photographs.

In our review of the research literature, we did not find any study of L2 phrasal vocabulary learning for L2 beginners, especially adult beginners. The studies of Wolter and Gyllstad (2011) and Yamashita and Jiang (2010) tested adults on their L2 collocational knowledge (i.e., verb-noun phrases with high frequency) when they had a high L2 proficiency. Hence, from our study findings, a higher level of L2 vocabulary learning (i.e., phrasal vocabulary such as verb-noun phrases) could be suggested to apply to L2 adult beginners in a visual context.

Our research has potentially a more specific implication. For instance, adult L2 beginners can learn Vietnamese verb-noun phrases successfully. They will be able to make and to use simple Vietnamese sentences in daily life. That is for the following reasons:

- The original forms of Vietnamese verbs and nouns do not change when they appear in sentences with any subject or tense (see Table 35).
- The most general structure in Vietnamese sentences is SVO. That means the order of a verb and a noun is identical when they stand in an infinitive verb-noun phrase and a sentence (see Table 35).

Table 35: The constant form of a Vietnamese verb-noun phrase in sentences

	Vietnamese	German	English
A phrase	***uống*** *nước*	*Wasser* ***trinken***	to ***drink*** water
SVO sentences	▪ Tôi ***uống*** *nước.* ▪ Cô ấy ***uống*** *nước.* ▪ Tôi vừa ***uống*** *nước.* ▪ Tôi đã ***uống*** *nước.* ▪ Bạn ***uống*** *nước* à?	▪ Ich ***trinke*** *Wasser.* ▪ Sie ***trinkt*** *Wasser.* ▪ Ich ***habe*** *Wasser* ***getrunken***. ▪ Ich ***trank*** *Wasser.* ▪ ***Trinkst*** du *Wasser*?	▪ I ***drink*** *water.* ▪ She ***drinks*** *water.* ▪ I ***have drunk*** *water.* ▪ I ***drank*** *water.* ▪ ***Do*** you ***drink*** *water*?
Comparison	- The forms of the verb and the noun are constant in every language context. - The order of the verb and the noun does not change.	- The forms of the verb are not constant. They depend on pronouns and tenses. - The order of the verb and the noun changes according to tenses.	

In the current study, L1 German adults learned L2 Vietnamese verb-noun phrasal vocabulary successfully with supportive depicted event photographs. We could continuously train them to make simple SVO sentences with learned L2 verb-noun phrases as the next learning step. Then, they could easily apply these L2 sentences in real life for various purposes of learning and using L2.

The initial research question relating to the effects of event depiction in L2 learning was motivated by some earlier studies of language processing (Münster, 2016; Zhang & Knoeferle, 2012). These studies examined the effects of depicted action events compared to non-depicted ones in eye-tracking experiments (Section 2.2.2). Zhang and Knoeferle (2012) observed online and offline effects of depicted over non-depicted action events for children, and online but no offline effects of event depiction for adults when participants processed their L1 OVS sentences together with seeing depicted actions. Münster (2016) found that young and older adults and children benefited from depicted event actions for online sentence processing. Language comprehenders processed OVS sentences very well in real time when they saw the agent depicted performing (vs. not performing) the action mentioned in the spoken sentence. All three age groups tested, especially children, answered the post-sentence comprehension question, "who does what to whom?" more correctly for depicted action events than they did for non-depicted action events.

In our L2 vocabulary learning study, we also found significant effects of directly depicted action events (vs. no action events) on L2 phrasal vocabulary learning. L1 German adults learned L2 Vietnamese verb-noun phrases with high accuracy and short reaction times in learning conditions containing present event photographs. The L2 vocabulary learning success of participants was maintained in delayed testing. Therefore, we would say that L1 comprehenders and L2 learners benefit from depicted action events in real-time L1 processing and L2 vocabulary learning.

Table 36: Second language phrasal learning with non-depicted event photographs

a. Absent action event photographs *(in our current study)*	b. Non-depicted action event photographs *(suggested for future studies)*

After conducting L2 learning experiments, we realized a limitation of our experimental design. In L2 learning conditions of absent action events, although L2 learners heard the spoken L2 verb-noun phrase (e.g., Sound 2 in Table 36a), they did not see anything. The absence of event made L2 beginners extremely confused in the learning phase. It would have been more contextual if we had replaced no events with photos that showed a person and an object without depicting the action. In the learning conditions of events with non-depicted actions, participants looked at the object they had learned again with a person, but the person does not act on the object (e.g., Picture 2 and Sound 2 in Table 36b). L2 learners still could not indicate the meaning of the spoken verb (e.g., *đọc/read*) in the combination of the spoken verb-noun phrase (e.g., *đọc-sách/read-book*). However, by seeing a person and an object depicted, they could guess that the person will do something with the object.

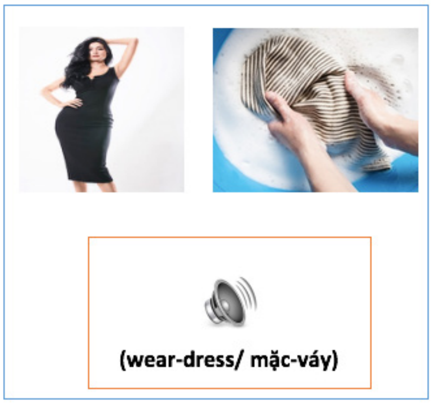
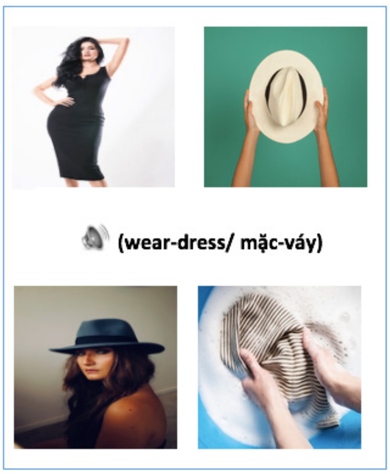
Another idea for future study is that we could set up three levels for the independent experiment factor—event depiction such as absent events, events with depicted actions, and events without depicted actions. We could in this way differences in L2 phrasal vocabulary learning outcomes when L2 beginners learned them in three visual learning inputs.

11.2 Second Language Transfer in the Specific Second Language Learning Situation

In six reaction time studies of L2 vocabulary learning, we found no main effects of L1–L2 language mapping (similar vs. different), which indicated that L2 adult beginners learned L2 verb-noun phrases successfully in the two verb mapping conditions. We investigated several verb mapping effects by testing part interaction and verb mapping by event depiction interaction across experiments.

The results did not show an advantage of similar verb mapping (i.e., L1–L2 correspondence) compared to different verb mapping (i.e., L1–L2 difference) as reviewed in previous studies in Section 4.2. Many researchers have agreed that L1 lexical knowledge could strongly influence the L2 vocabulary learning process and results at a very early stage of L2 learning. However, our studies did not show any negative or positive transfer in two specific language mapping conditions. Why was this the case? The first main reason we found after conducting six studies was the inappropriate combination of learning and testing design. In the original experimental design (Experiment 1), we created contrasts in learning and testing (see Experiment 1; Table 37a and 37b). Participants learned the verb "to wear" in Vietnamese with two different forms and collocations (i.e., *wear dress* vs. *wear-hat* ⇔ *mặc-váy* vs. *đội-mũ*), then they were tested by choosing one of two objects after hearing a verb sound meaning "to wear" (e.g., *mặc).* Even though the design of learning and testing in Experiment 1 seemed to be useful for examining the language mapping factor, we found no significant effect of language mapping and no main effect of event photographs. Events were not tested in the task by matching.

Table 37: Learning and testing Second Language phrasal vocabulary in different verb mapping conditions.

	Experiment 1	**Experiments 2, 2R, 3, 4, and 2N**
a. Learning	*wear dress* vs. *wear-hat* 	=> learning in a contrast: two different L2 verbs (e.g. *mặc, đội*) corresponding to one L1 verb (e.g. *tragen/ wear*)
b. Testing	*dress* vs. *hat*  => testing in a contrast, but event photographs were not tested	*wear dress* vs. *wash dress*  => event photographs were tested, but testing has no contrast for different verb mapping because the phrase "*wear-dress*" was tested with a new phrase "*wash-dress*".
c. Suggested testing		=> event photographs were tested in a contrast for different verb mapping conditions with four events/phrases (e.g., *wear-dress, wear-hat, wash-dress,* and *carry-hat*)

In Experiments 2 to 2R, 3, 4, and 2N, we changed to testing events after learning; then, we investigated the critical role of present events in learning compared to absent events. However, we still found no primary influence of language mapping. We initially did not see that we set up no contrast in testing for the level of different language mapping. Looking at Table 37a and 37b, we can see participants learned *wear dress* vs. *wear-hat (mặc-váy* vs. *đội-mũ)* in a learning trial, but they were not tested in the same testing trial. We tested two different actions related to an object per testing trial (e.g., *wear dress* vs. *wash dress*). No specific contrast in testing might not result in a difference in participants' L2 learning success between similar verb mapping and different verb mapping conditions. Hence, we thought about constructing a testing trial that enables event photographs to be tested, with a contrast of different verb mapping conditions. Table 37c presents our suggestion for a possible new testing design for future studies.

The second leading cause for the absence of a main effect of language mapping in our studies might be the robustly significant effects of event depictions. In event-present conditions, L2 learners could easily recognize events after learning based on reliable visual and auditory cues. L2 learners may have concentrated on the visual context (i.e., event depictions) as explicit information (i.e., event photographs were absent or present) during their L2 vocabulary process. However, they might have ignored the language context (i.e., L1–L2 similar and different verb mappings) because it was implicit information (i.e., they heard two different Vietnamese verbs meaning "to wear" while they inspecting "a person wearing dress" and "a person wearing hat") in the L2 learning situation.

Furthermore, as mentioned in Section 4, many researchers (Bardel & Falk, 2007; Falk & Lindqvist, 2014; Navés, Miralpeix, & Celaya, 2005; Pfenninger & Singleton, 2016; Williams & Hammarberg, 1998) have confirmed that form-based transfer is dominant when L2 learners are at lower levels of L2 proficiency, while meaning-based transfer expands when learners have high L2 proficiency. In our L2 learning study, we could explain the absence of L1–L2 meaning-based transfer caused by the very low L2 proficiency of L2 beginners. Also, short-term laboratory learning experiments with a limited amount of L2 might not be enough to measure negative or positive L1–L2 lexical transfer.

11.3 Age Differences in Second Language Vocabulary Learning Success

All three adult age groups could use the direct cue in all studies, i.e., present event photographs to respond correctly in testing. However, these event effects differed by the age of the learners.

Young adults (18–31) performed best in the L2 learning task (the highest accuracy and the fastest response speeds) in event-present conditions among the three groups. There were no significant differences in L2 vocabulary learning success between young adults (18–31) and early middle-aged adults (32–45). However, we still found a slight decrease in the percentage of correct choices from 96.1% to 94.2% and 77.3% to 72.3% in immediate testing and delayed testing, respectively, in older adults. Significant differences were investigated between the two younger adult groups and the oldest adult groups (i.e., late middle-aged adults aged from 46 to 65). The number of correct responses was approximately 10% to 12% less in immediate testing and around 5% to 10% less in delayed testing. Also, the data analysis showed a significant correlation between learners' age and reaction times in which the older the learners, the longer the reaction time (see Section 9.5.3).

Our study examined L2 learners aged from 18 to 65, learning the same amount of L2 vocabulary. However, the results did not replicate previous findings because of testing a different age range. In Section 5, we focused on reviewing the importance of controlling for the amount of L2 exposure (Muñoz, 2008). In the age range of L2 learners from 8 to 38 (i.e., children and young adults), the main finding of this research line indicated that late L2 starters such as young adults (18–38 years) learned or processed L2 more successfully than early starters such as children (7 to under 18). In our L2 vocabulary learning experiments, the data on participants' accuracy and response speeds indicated that when L2 adults get older, especially when aged from 46 to 65, they are less accurate in and slower at responding to the L2 verb-nouns learned. However, young adults (18–31) were as accurate as early middle-aged adults (32–45), and the early middle-aged adults (32–45 years) were as slow as the late-middle-aged adults (46–65). Hence, in our study, we could say that young adults (18-31) were the best L2 vocabulary learners (i.e., the highest accuracy and the fastest speeds). Early middle-aged adults (32–45) could maintain the same high accuracy as young adults (18–31 years). However, they took much longer to respond. The disadvantage of age was clearly shown for late middle-aged adults (46–65 years) since they responded most slowly and most inaccurately. In summary, the findings of the reaction time experiments indicated that beginners' L2 phrasal vocabulary learning success in the early L2 learning stage was modulated by age.

Our current study supplies more evidence for the age range from 32 to 65 for the research line—different age at the time of testing, the same amount of L2 exposure. When we combine our results with other previous findings in Section 5, we can conclude that young adults are the best group of L2 learners compared to children (7 to under 18) and older adults (32–65 years) when all of them have the same amount of L2 exposure to both long-term and short-term L2 learning.

11.4 Learners' Cognitive Ability and their Second Language Vocabulary Learning Success

The results of cognitive tests in our study replicated all of the findings of Carminati and Knoeferle (2013). Fistly, younger adults (18–30) performed significantly better than older adults (32–45 and 46–65) in the picture completion and the digit symbol tasks of the WAIS test. Next, in other tests (the digit span, similarities, and vocabulary tests), the three groups did not differ significantly. Another finding of our WAIS test indicated that participants' cognitive ability was modulated by age. Marginally significant differences in cognitive test scores were shown between young adults and late middle-aged adults and between early middle-aged adults and late middle-aged adults.

Münster (2016) examined whether cognitive test scores of participants were linked to their accuracy for comprehension questions in each age group. She found a reliably significant correlation between accuracy scores and cognitive scores for older adults (60–80), but not for children.

In our L2 learning study, no significant correlation between adult learners' accuracy scores and their cognitive test scores was found in any of three age groups. However, differences in cognitive test scores could explain the differences between participants' high and low accuracy among groups tested. Young adult learners (18–31) were the best performers for the picture completion and the digit symbol subtests of WAIS. Hence, they could perceive visual cues in L2 learning experiments much more effectively than other older adults (32–65). As a result, young adult learners had the highest accuracy among groups after they efficiently inspected depicted event photographs in L2 vocabulary learning.

We also checked the link between participants' cognitive test scores and means of reaction time in L2 vocabulary learning experiments. We found only one significant correlation in the young adult group, which indicated that the higher the cognitive test scores, the faster the participants' correct responses.

11.5 Conclusion

To conclude, we argue for integrating visual and linguistic factors into accounts of beginners' L2 phrasal vocabulary learning success. Our studies have demonstrated that visual information (event photographs) can effectively facilitate adults' L2 verb-noun phrase learning. Indeed, more research (i.e., L2 vocabulary learning and then processing) is needed to investigate how cross-linguistic influences are when beginners learn L2 phrasal-level vocabulary in the case of L1-L2 partial conceptual equivalence. The learning process of this conceptual equivalence relationship (Pavlenko, 2009) is a combination of restructuring known L1 concepts and creating new links between L2 words and other concepts that do not exist in L1. We also want to examine L2 beginners with different L1 learning L2 phrasal vocabulary in the same learning design.

Additionally, our studies highlight the importance of taking learner characteristics such as age (i.e., physical characteristics) into account when investigating the accuracy and speed of L2 vocabulary learning. However, we found no correlation between adults' (18-65 years) cognitive characteristics (scores of WAIS-IV test, German version) and their L2 vocabulary learning success in the thesis project. In comparison, Münster (2016) found a reliable correlation between listeners' cognitive scores and their accuracy scores in language processing studies for older adults (60-80 years), but not for children. In future studies, investigating how children and older adults interact in this L2 learning context is wanted.

Lastly, the thesis study has a high implication of teaching L2-Vietnamese phrasal vocabulary for L1-German speakers or speakers of other transfiguration languages. Vietnamese is not a transfiguration language, then learning a Vietnamese verb-noun phrase (e.g., *uống-nước/drink-water*), language beginners can easily make a sentence (e.g., *Lina uống nước./ Lina drinks water.)* in speaking and writing. L2 beginners for specific L2 can start learning phrasal vocabulary items in which they learn nouns, verbs, and verb-noun phrases in relevance.

References

Asher, J., & Price, B. (1967). The Learning Strategy of the Total Physical Response: Some Age Differences. *Child Development*, *38*(4), 1219. https://doi.org/10.2307/1127119

Borovsky, A., Kutas, M., & Elman, J. (2010). Learning to use words: Event-related potentials index single-shot contextual word learning. *Cognition*, *116*(2), 289-296. https://doi.org/10.1016/j.cognition.2010.05.004

Carminati, M., & Knoeferle, P. (2013). Effects of Speaker Emotional Facial Expression and Listener Age on Incremental Sentence Processing. *Plos ONE*, *8*(9), e72559. https://doi.org/10.1371/journal.pone.0072559

Comesaña, M., Soares, A., Sánchez-Casas, R., & Lima, C. (2011). Lexical and semantic representations in the acquisition of L2 cognate and non-cognate words: Evidence from two learning methods in children. *British Journal Of Psychology*, *103*(3), 378-392. https://doi.org/10.1111/j.2044-8295.2011.02080.x

Dagut, M., & Laufer, B. (1985). Avoidance of Phrasal Verbs—A Case for Contrastive Analysis. *Studies In Second Language Acquisition*, *7*(1), 73-79. https://doi.org/10.1017/s0272263100005167

Ellis, R. (1994). A theory of instructed second language acquisition. In N. Ellis (Ed.), Implicit and explicit learning of languages. *Academic Press*, 79-114. Google Scholar

Emirmustafaoğlu, A., & Gökmen, D. (2015). The Effects of Picture vs. Translation Mediated Instruction on L2 Vocabulary Learning. *Procedia - Social And Behavioral Sciences*, *199*, 357-362. https://doi.org/10.1016/j.sbspro.2015.07.559

Fisher, C. (2002). Structural limits on verb mapping: the role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science*, *5*(1), 55-64. https://doi.org/10.1111/1467-7687.00209

Gass, S. M. & Selinker, L. 1992. Language transfer in language learning. Amsterdam: John Benjamins Pub.

Gass, S. M. & Selinker, L. (2001). Second Language Acquisition, an Introductory Course (2nd ed.). *Mahwah, NJ: Lawrence Erlbaum Associates.* Google Scholar

Gleitman, L. (1990). The Structural Sources of Verb Meanings. *Language Acquisition*, *1*(1), 3-55. https://doi.org/10.1207/s15327817la0101_2

Havas, V., Taylor, J., Vaquero, L., de Diego-Balaguer, R., Rodríguez-Fornells, A., & Davis, M. (2018). Semantic and phonological schema influence spoken word learning and overnight consolidation. *Quarterly Journal Of Experimental Psychology*, *71*(6), 1469-1481. https://doi.org/10.1080/17470218.2017.1329325

Hulstijn, J., & Marchena, E. (1989). Avoidance. *Studies In Second Language Acquisition*, *11*(3), 241-255. https://doi.org/10.1017/s0272263100008123

Jarvis, S. (2008). *Crosslinguistic Influence in Language and Cognition*. Routledge. https://doi.org/10.4324/9780203935927

Jarvis, S. (2000). Methodological Rigor in the Study of Transfer: Identifying L1 Influence in them Interlanguage Lexicon. *Language Learning*, *50*(2), 245-309. https://doi.org/10.1111/0023-8333.00118

Jarvis, S., & Odlin, T. (2000). Morphological type, spatial reference, and language transfer. *Studies In Second Language Acquisition,* 22(4), 535-556. https://doi.org/10.1017/s0272263100004034.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics*, *21*(1), 47-77. https://doi.org/10.1093/applin/21.1.47

Jiang, N. (2002). Form-meaning mapping in vocabulary acquisition in an second language. *Studies In Second Language Acquisition*, 24(4), 617-637. https://doi.org/10.1017/s0272263102004047.

Jiang, N. (2004). Semantic Transfer and Its Implications for Vocabulary Teaching in a Second Language. *The Modern Language Journal*, *88*(3), 416-432. https://doi.org/10.1111/j.0026-7902.2004.00238.x

Kaushanskaya, M., Yoo, J., & Van Hecke, S. (2013). Word Learning in Adults With Second-Language Experience: Effects of Phonological and Referent Familiarity. *Journal Of Speech, Language, And Hearing Research*, *56*(2), 667-678. https://doi.org/10.1044/1092-4388(2012/11-0084)

Knoeferle, P., & Guerra, E. (2012). What's non-linguistic visual context? *Linguistik Aktuell/Linguistics Today*, 129-150. https://doi.org/10.1075/la.196.09kno

Knoeferle, P., & Guerra, E. (2016). Visually Situated Language Comprehension. *Language And Linguistics Compass*, *10*(2), 66-82. https://doi.org/10.1111/lnc3.12177

Knoeferle, P., Crocker, M., Scheepers, C., & Pickering, M. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. *Cognition*, *95*(1), 95-127. https://doi.org/10.1016/j.cognition.2004.03.002

Koehne, J., & Crocker, M. (2014). The Interplay of Cross-Situational Word Learning and Sentence-Level Constraints. *Cognitive Science*, *39*(5), 849-889. https://doi.org/10.1111/cogs.12178

Krashen, S. (1985). *Inquiries & insights*. Alemany Press.

Laufer, B., & Eliasson, S. (1993). What Causes Avoidance in L2 Learning. *Studies In Second Language Acquisition*, *15*(1), 35-48. https://doi.org/10.1017/s0272263100011657

MacDonald, K., Yurovsky, D., & Frank, M. (2017). Social cues modulate the representations underlying cross-situational learning. *Cognitive Psychology*, *94*, 67-84. https://doi.org/10.1016/j.cogpsych.2017.02.003

Masny, D., Gass, S., & Selinker, L. (1985). Language Transfer in Language Learning. *The Modern Language Journal*, *69*(2), 166. https://doi.org/10.2307/326512

Meriläinen, L. 2010. Language Transfer in the Written English of Finnish Students. Publications of the University of Eastern Finland. Dissertations in Education, Humanities, and Theology, No 9. Joensuu: University of Eastern Finland.

Mitchell, R. & Myles, F. (2004). Second language learning theories (2nd ed.). *London: Hodder Arnold*

Miralpeix, I. (2006). Chapter 4. Age and Vocabulary Acquisition in English as a Foreign Language (EFL). *Age And The Rate Of Foreign Language Learning*, 89-106. https://doi.org/10.21832/9781853598937-006

Miralpeix, I. (2007). Lexical knowledge in instructed language learning: The effects of age and exposure. *International Journal of English Studies*, 7(2), 6183.

Morvay, G. (2015). The role of L1 reading ability, L2 proficiency and non-verbal intelligence in L2 reading comprehension. *Argentinian Journal of Applied Linguistics*, 3(1), 26-47.

Münster, K. (2016). Effects of emotional facial expressions and depicted actions on situated language processing across the lifespan. *Pub.Uni-Bielefeld.De*. https://pub.uni-bielefeld.de/record/2906648

Nassaji, H., & Hu, M. (2012). The relationship between task-induced involvement load and learning words from context. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 50, 69-86.

I S P Nation. (2001). *Learning Vocabulary in Another Language*. Cambridge University Press.

Odlin, T. (1989). Language Transfer: Cross-Linguistic Influence in Language Learning. In *Cambridge University Press*. Cambridge University Press.

Odlin, T. (1990). Word Order Transfer, Metalinguistic Awareness, and Constraints on Foreign Language Learning. In W. Patten, & J. Lee (Eds.), Second Language Acquisition/Foreign Language Learning (pp. 95-117). Clevedon: Multilingual Matters.

Odlin, T. 2003. "Cross-Linguistic Influence". In The Handbook of Second Language Acquisition, ed. Doughty, C., & Long, M. H., 436–486. Malden (Ma.): Blackwell.

Ong, J. H., & Chan, A. H. D. (2019). The influence of referent type and familiarity on word-referent mapping. *PLOS ONE*, *14*(7), e0219552. https://doi.org/10.1371/journal.pone.0219552

Olson, L. L., & Jay Samuels, S. (1973). The Relationship between Age and Accuracy of Foreign Language Pronunciation. *The Journal of Educational Research*, *66*(6), 263–268. https://doi.org/10.1080/00220671.1973.10884476

Pavlenko, A. (2009) Conceptual representation in the bilingual lexicon and second language vocabulary learning. In Pavlenko, A. (ed.) The bilingual mental lexicon: Interdisciplinary approaches. Clevedon, UK: Multilingual Matters, pp. 125-160.

Rouhi, A. & Mohebbi, H. (2013). Glosses, spatial intelligence, and L2 vocabulary learning in multimedia context. *3L: The Southeast Asian Journal of English Language Studies*, 19 (2), 75 – 87.

Ringbom, H. (1983). Borrowing and lexical transfer. In Applied Linguistics 4, 207–212.

Ringborn, H. (1987). The role of the first language in foreign language learning. Clevedon, UK: Multilingual Matters.

Ringbom, H. (1992). On L1 transfer and L2 comprehension and L2 production. In Language Learning 42, 85–112.

Ringbom, H. (2007). Cross-Linguistic Similarity in Foreign Language Learning. Clevedon: Multilingual matters.

Sajavaara, K. (1987). Second language speech production: Factors affecting fluency. In Dechert, H.W. & Raupach, M. (Eds.), *Psycholinguistic models of production* (pp. 45–65). Norwood, NJ: Ablex.

Shakouri, N., Sheikhy Behdani, R., & Teimourtash, M. (2017). On the relationship between linguistic intelligence and recalling lexical items in SLA. *International Journal of Research Studies in Education*, Volume 6, Issue Number 4, October 2017. https://doi.org/10.5861/ ijrse.2016.1644.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Stockwell, R. P., Bowen, D. J. and Martin, J. W. (1965a). The Sounds of English and Spanish. Chicago: University of Chicago Press.

Stockwell, R. P., Bowen, D. J. and Martin, J. W. (1965b). The Grammatical Structures of English and Spanish. Chicago: University of Chicago Press.

Shackelford, T. K., & Weekes-Shackelford, V. A. (Eds.). (2020). *Encyclopedia of Evolutionary Psychological Science*. Springer International Publishing. https://doi.org/10.1007/978-3-319-16999-6

Tajeddin, Z. & Chiniforoushan, N. (2011). Visual intelligence and lexical enhancement tasks: Their impacts of EFL learners' receptive and productive vocabulary. *The Journal of Asia TEFL*, 8(3), 109-134.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Wang, S., Allen, R. J., Fang, S.-Y., & Li, P. (2017). Cross-modal working memory binding and L1-L2 word learning. *Memory & Cognition*, *45*(8), 1371–1383. https://doi.org/10.3758/s13421-017-0731-2

Wolter, B., & Gyllstad, H. (2011). Collocational Links in the L2 Mental Lexicon and the Influence of L1 Intralexical Knowledge. *Applied Linguistics*, *32*(4), 430–449. https://doi.org/10.1093/applin/amr011

Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (2011). Infants learn about objects from statistics and people. *Developmental Psychology*, *47*(5), 1220–1229. https://doi.org/10.1037/a0024023

Vilke, M. (1988). Some psychological aspects of early second-language acquisition. *Journal of Multilingual and Multicultural Development*, *9*(1–2), 115–128. https://doi.org/10.1080/01434632.1988.9994323

Yamada, J. , Takatsuka, S. , Kotake, N. and Kurusu, J. (1980). On the optimum age for teaching foreign vocabulary to children. International Review of Applied Linguistics in Language Teaching, 18, 245-247.

Yamashita, J., & Jiang, N. (2010). L1 Influence on the Acquisition of L2 Collocations: Japanese ESL Users and EFL Learners Acquiring English Collocations. *TESOL Quarterly*, *44*(4), 647–668. https://doi.org/10.5054/tq.2010.235998

Yu, C., & Smith, L. B. (2007). Rapid Word Learning Under Uncertainty via Cross-Situational Statistics. *Psychological Science*, *18*(5), 414–420. https://doi.org/10.1111/j.1467-9280.2007.01915.x

Yu, C., & Smith, L. B. (2011). What you learn is what you see: using eye movements to study infant cross-situational word learning. *Developmental Science*, *14*(2), 165–180. https://doi.org/10.1111/j.1467-7687.2010.00958.x

Zhang, L., Kornbluth, L., Knoeferle, P. The role of recent versus future events in children's comprehension of referentially ambiguous sentences: Evidence from eye tracking. *Proceedings of the Annual Meeting of the Cognitive Science Society*. Boston, USA. 1227-1232 (2012).

Zobl, H. (1980). The formal and developmental selectivity of L1 influence on L2 acquisition. *Language Learning*, *30*(1), 43–57. https://doi.org/10.1111/j.1467-1770.1980.tb00150.x

Appendices

APPENDIX A

The list of 32 Vietnamese verb-noun phrases in Experiment 1

Item	*Similar verb mapping*		*Item*	*Different verb mapping*	
	Vietnam -ese	*English & German*		*Vietnam -ese*	*English & German*
1	đạp xe	(to) ride (a) bike Fahrrad fahren	17	mặc áo	(to) wear (a) shirt Hemd tragen
2	đạp xích lô	(to) ride (a) cyclo Cyclo fahren	18	mặc váy	(to) wear (a) skirt Kleide tragen
3	lau bàn	(to) clean (a) table Tisch reinigen	19	đội mũ	(to) wear (a) hat Hut tragen
4	lau nhà	(to) clean (a) floor Boden reinigen	20	đội nón	(to) wear (a) scarf Schal tragen
5	đọc sách	(to) read (a) book Buch lessen	21	đeo đồng hồ	(to) wear (a) watch Armbanduhr tragen
6	đọc báo	(to) read (a) newspaper Zeitung lesen	22	đeo kính	(to) wear (a) eyeglass Brille tragen
7	chụp phim	to) take (a) X-ray X-ray machen	23	đi tất	(to) wear (a) sock Socke tragen
8	chụp ảnh	to) take (a) photo Foto machen	24	đi giày	(to) wear (a) shoe Schuh tragen
9	pha cà phê	to) make (a) coffee Kaffe kochen	25	bê ghế	(to) carry (a) chair Stuhl tragen
10	pha trà	to) make (a) tea Tee kochen	26	bê mâm (cơm)	(to) carry (a) tray (of rice) Servierbrett tragen

11	nâng chén	(to) lift (a) cup Tasse heben	27	khiêng tủ	(to) carry (a) wardrobe Schrank tragen
12	nâng ly	(to) lift (a) glass Glas heben	28	khiêng giường	(to) carry (a) bed Bett tragen
13	xây nhà	(to) build (a) house Haus bauen	29	bưng chén (trà)	(to) carry (a) cup (of tea) Tasse tragen
14	xây cầu	(to) build (a) bridge Brücke bauen	30	bưng bát (cơm)	(to) carry (a) bowl Schüssel tragen
15	bắt cá	(to) catch (a) fish Fisch fangen	31	xách túi	(to) carry (a) bag Tasche tragen
16	bắt bóng	(to) catch (a) ball Ball fangen	32	xách vali	(to) carry (a) suitcase Koffer tragen

APPENDIX B

The list of 32 Vietnamese verb-noun phrases in Experiment 2 and 2R

Item	*Similar verb mapping*		*Item*	*Different verb mapping*	
	Vietnam -ese	*English & German*		*Vietnam -ese*	*English & German*
1	dắt xe	(to) lead (a) bike Fahrrad führen	17	mặc áo	(to) wear (a) shirt Hemd tragen
2	đạp xích lô	(to) ride (a) cyclo Cyclo fahren	18	đội mũ	(to) wear (a) hat Hut tragen
3	đóng bàn	(to) make (a) table Tisch machen	19	đeo đồng hồ	(to) wear (a) watch Armbanduhr tragen
4	quét nhà	(to) sweep (a) floor Boden fegen	20	đi tất	(to) wear (a) sock Socke tragen
5	đọc sách	(to) read (a) book Buch lessen	21	giặt váy	(to) wash (a) skirt Kleide waschen
6	bán báo	(to) sell (a) newspaper Zeitung verkaufen	22	rửa ô tô	(to) wash (a) car Auto waschen
7	chụp phim	to) take (a) X-ray X-ray machen	23	thái thịt	(to) cut (a piece of) meat Fleisch schneiden
8	in ảnh	(to) print (a) photo Foto drücken	24	cắt bánh	(to) cut (a) cake Kuchen schneiden
9	rang cà phê	to) roast (a) coffee (beans) Kaffe braten	25	khiêng tủ	(to) carry (a) wardrobe Schrank tragen
10	pha trà	to) make (a) tea Tee kochen	26	bê ghế	(to) carry (a) chair Stuhl tragen
11	nâng chén	(to) lift (a) cup (of wine) Tasse heben	27	bưng bát (cơm)	(to) carry (a) (rice) bowl Schüssel tragen

12	treo ly	(to) hang (a) glass Glas hangen	28	xách túi	(to) carry (a) bag Tasche tragen
13	lái tàu	(to) build (a) house Haus bauen	29	chẻ rau	(to) split (a) stick (of vegetables) Gemüse spalten
14	lau kính	(to) clean (a) eyeglass Brille reinigen	30	bổ ổi	((to) split (a) guava Guava spalten
15	bắt cá	(to) catch (a) fish Fisch fangen	31	cõng bé	(to) carry (a) baby Baby tragen
16	tâng bóng	(to) juggle (a) ball Ball jonglieren	32	vác củi	(to) carry (a) firewood Brennholz tragen

APPENDIX C

Links[7] for object photographs and event photographs (Exp.2, 2R, 2N, 3, and 4)

Item	*Object pictures*	*Event pictures in Part 1 (immediate testing)*	*Event pictures in Part 3 (delayed testing)*	*Event pictures as competitors in Part 1 (immediate testing*	*Event pictures as competitors Part 3 (delayed testing)*
1	bike	lead-bike-1	lead-bike-2	ride-bike-1	ride-bike-2
2	cyclo	ride-cyclo-1	ride-cyclo-2	sit-on-cyclo-1	sit-on-cyclo-2
3	table	make-table-1	make-table-2	clean-table-1	clean-table-2
4	floor	sweep-floor-1	sweep-floor-2	clean-floor-1	clean-floor-2
5	book	read-book-1	read-book-2	look-for-book-1	look-for-book-2
6	newspaper	sell-newspaper-1	sell-newspaper-2	read-newspaper-1	read-newspaper-2
7	Xray-film	take-Xray-film-1	take-Xray-film-2	read-Xray-film-1	read-Xray-film-2
8	photo	print-photo-1	print-photo-2	take-photo-1	take-photo-2
9	coffee-(bean)	roast-coffee-(bean)-1	roast-coffee-(bean)-2	drink-coffee-1	dink-coffee-2
10	tea	make-tea-1	make-tea-2	drink-tea-1	drink-tea-2
11	cup*	lift-cup-1	lift-cup-2	hold-cup-(by-a-hand)-1	hold-cup-(by-a-hand)-2
12	wine-glass	hang-wine-glass-1	hang-wine-glass-2	clean-wine-glass-1	clean-wine-glass-2
13	train	drive-train-1	drive-train-2*	go-into-train-1	go-into-train-2

[7] We collected photographs on the Internet used in experiments, and many of them are copyrighted. Links of pictures with the signal * as references were created by Huong Thi Thu Nguyen because the original links are no longer available.

14	eye-glass	clean-eye-glass-1	clean-eye-glass-2	wear-eye-glass-1	wear-eye-glass-2
15	fish	catch-fish-1	catch-fish-2	fishing-1	fishing-2
16	ball	juggle-ball-1	juggle-ball-2	catch-ball-1	catch-ball-2
17	shirt*	wear-shirt-1*	wear-shirt-2*	iron-shirt-1	iron-shirt-2
18	hat*	wear-hat-1*	wear-hat-2	decorate-hat-1	decorate-hat-2
19	watch*	wear-watch-1*	wear-watch-2	look-at-watch-1	look-at-watch-2
20	sock	wear-sock-1	wear-sock-2*	dry-sock-(in the sun)-1	dry-sock-(in the sun)-2
21	dress	wash-dress-1	wash-dress-2	wear-dress-1*	wear-dress-2*
22	car*	wash-car-1	wash-car-2	drive-car-1	drive-car-2
23	meat	cut-meat-1*	cut-meat-2	grill-meat-1	grill-meat-2
24	cake	cut-cake-1	cut-cake-2	bake-cake-1	bake-cake-2
25	wardrobe	carry-wardrobe-(by-two-people)-1	carry-wardrobe-(by-two-people)-2	carry-wardrobe-(on back)-1	carry-wardrobe-(on back)-2*
26	chair	carry-chair-(by-two-hands)-1	carry-chair-(by-two-hands)-2	sit-on-chair-1	sit-on-chair-2
27	bowl-(of rice)	carry-bowl-(of rice-by two hands)-1	carry-bowl-(of rice-by two hands)-2*	take-bowl-(of rice)-1	take-bowl-(of rice)-2
28	bag*	carry-bag-1*	carry-bag-2*	paint-bag-1	paint-bag-2
29	vegetable	spilt-vegetable-1	spilt-vegetable-2	clean-vegetable-1*	clean-vegetable-2
30	apple	split-apple-1	split-apple-2	pick-apple-1	pick-apple-2

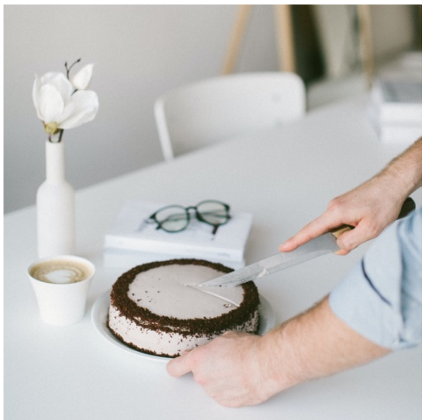
31	baby	carry-baby-(on back)-1	carry-baby-(on back)-2	hold-baby-(on a side)-1	hold-baby-(on a side)-2
32	firewood	carry-firewood-(on a shoulder)-1	carry-firewood-(on a shoulder)-2	carry-firewood-(with bamboo frame)-1	carry-firewood-(with bamboo frame)-2

APPENDIX D

Thirty-two critical Vietnamese verb-noun phrases in the appendix B are depicted by 32 free event photographs (sources: https://unsplash.com; https://www.pexels.com). These event photographs were not used in experiments of the thesis.

Verb-noun phrases & Event pictures	
 dắt-xe	 mặc-áo
 đạp-xe	 đội-mũ
 đóng-bàn	 đeo-đồng hồ

 quét-nhà	 đi-tất
 đọc-sách	 giặt-váy
 bán-báo	 Rửa-ô tô
 chụp-phim	 thái-thịt

 in-ảnh	 cắt-bánh
 rang-cà phê	 khiêng-tủ
 pha-trà	 bê-ghế
 nâng-chén	 bưng-bát cơm

 treo-ly	 xách-túi
 tái-tàu	 chẻ-rau
 lau-kính	 bổ-táo
 bắt-cá	 cõng-bé

 tâng-bóng	 vác-củi