Doctoral Dissertation

How to Learn Vocabulary in a Foreign Language

(Cómo Aprender Vocabulario en una Lengua Extranjera)

International PhD candidate

Ana Belén García Gámez

Advisor

Pedro Macizo Soria

PhD in Psychology

Department of Experimental Psychology

University of Granada

June, 2019

Editor: Universidad de Granada. Tesis Doctorales
Autor: Ana Belén García Gámez
ISBN: 978-84-1306-268-6
URI: http://hdl.handle.net/10481/56575

Front cover created by Jose Manuel Contreras González

Publicaciones de la Tesis

García-Gámez, A. B., & Macizo, P. (2019). Learning nouns and verbs in a foreign language: The role of gestures. *Applied Psycholinguistics, 40*(2), 473-507. doi: 10.1017/S0142716418000656

García-Gámez, A. B., & Macizo, P. (2019). The way in which foreign words are learned determines their use in isolation and within sentences. *Journal of Experimental Psychology: Learning, Memory and Cognition.* Advance online publication. doi: 10.1037/xlm0000721

García-Gámez, A. B., & Macizo, P. (submitted). Seeing or Acting? The Effect of Performing Gestures on Foreign Language Vocabulary Learning. *Language Learning.*

García-Gámez, A. B., & Macizo, P. (in preparation). Semantic and lexical training to learn foreign-language words in adulthood**.** *Second Language Research.*

García-Gámez, A. B., & Macizo, P. (submitted). Lexical and Semantic Training to Acquire Words in a Foreign Language: An Electrophysiological Study. *Brain and Language.*

Agradecimientos

En primer lugar, muchas gracias Pedro, parece que fue ayer cuando entré por primera vez en tu despacho. Recuerdo el camino a la facultad y la inseguridad, pero confiaste en mí y en que podría sacar este trabajo adelante con fe casi ciega. Llegué con muchas lagunas, pero has tenido la paciencia de enseñarme todo lo que sé, y de animarme cuando nos rechazaban una y otra vez…"Venga Ana, ánimo, que esto va en tu tesis". No sé que me deparará el futuro, pero espero contar siempre contigo.

Gracias a mi padre, somos tan parecidos que queremos ser diferentes. Gracias por enseñarme las recompensar del esfuerzo y del trabajo. A mi madre, la diosa de la paciencia, siempre dispuesta a estar al otro lado del teléfono, aunque después de un rato pusiese el piloto automático (normal). A ti bro, no te dejes dominar nunca, eres especial.

Contigo se deshacen los límites del tiempo, 10 años, un gracias se queda corto. Jose, espero seguir escribiendo muchas más páginas de nuestra vida juntos.

A mis amigos, gracias por ser la vía de escape, el punto muerto, la mente en blanco. Se dice que la amistad se desvanece con el tiempo, que los grupos de amigos se acaban separando. Gracias por romper las estadísticas. Mariloli, por tus abrazos. Javi, por escuchar. Miguel Ángel, por borrar las penas. Manoli, por estar siempre ahí. Ramón, por la atención. Isabel…ay…por toda la vida. Rubén, por tus historias. Jose, por ver siempre el lado bueno.

Gracias a los que me rodeáis cada día de trabajo, qué difícil sería todo sin vosotros. Soy afortunada, es muy difícil que personas tan especiales como vosotros coincidan en un mismo lugar. A Chus y Alba, siempre cerca, el ejemplo a seguir. Antonio y Marta, la unión, el buen rollo, la amistad. Gracias por enseñarme a que a veces es mejor tirarse a la piscina. Óscar, supiste darle la vuelta a la tortilla rápidamente, gracias por ser la paz, la calma y el consejo. Nuria, por estar siempre dispuesta a escuchar y dar ánimos. Tania y Ana por compartir tanto, por escuchar. Cristina, por estos meses de complicidad. Fernando, por todo lo compartido. Borja, por todo tu trabajo, tu empeño y tu buen humor.

A todas las demás personas que forman el "Memory and Language Research Group", gracias por compartir estos últimos años. Especialemente, muchas gracias a Teresa, Daniela, Sandra y Paqui con acorgerme en esta péqueña gran familia y enseñarme tanto.

Nadie dijo que fuese a ser fácil, y no lo fue, especialmente al principio, pero como en toda buena historia, el final siempre debe estar a la altura.

Siempre recordaré estos momentos, mil gracias por hacerlo posible.

Introductory Note

This Research has been supported by the grant PSI2016-75250-P from the Spanish Ministry of Economy, Industry and Competitiveness to Pedro Macizo Soria.

The content of this doctoral dissertation has been drawn up following the regulations of the University of Granada to obtain the International Doctorate Mention in the Psychology Doctoral Program. According to the regulation, the majority of the thesis has been written in English. A chapter including a summary and conclusions of the thesis is presented in Spanish (Capítulo VIII).

Index at a Glance

Table of Contents

Preface

I always wanted to be a doctor, but not the kind of doctor that I'm trying to be now. Studying medicine was my dream and my objective when I was 18. But things are not as we desire and the destiny had another option ready for me. At the beginning it was not easy to assume, but after that, I realize that everything happens for a reason. When a door is closed, we always can find an opened window. What it really motivates me was helping people and discovering the Speech Therapist degree was a real gift. At some point I started to perceive that there were many aspects to improve and I fixed a new objective, being a Speech Therapy teacher. Seeing the shortcomings of the teaching system, I thought I could do something to improve it. Until know, I did not reach my goal, but I'm on my way and I'm going to continue fighting for it.

I can only add my little grain of sand, and this is for now my chance to help people. I'm proud of this work; this is not a purely theoretical research. Based on different theoretical models we have proposed different learning methods able to facilitate the second language acquisition in first language speaking contexts. Today, this is an important research topic because many people, not only children but also adults, have to face with language learning. Although this research is rooted in theoretical bases of linguistic processing in monolingual and bilingual population, it is also easily applicable in the second language teaching context. I hope this research serves to make people's lives easier.

"If you can't fly then run, if you can't run then walk,
if you can't walk then crawl, but whatever you do
you have to keep moving forward"

Dr. Martin Luther King Jr

-PART 1-

INTRODUCTION

CHAPTER I

Second Language Acquisition

INTRODUCTION

The word is continuously changing in any aspect that we can imagining, for this reason, the rule for all the species is "to adapt to survive". Humans are not the exception, as animals have to modulate their behaviour or their appearance to adapt themselves to the surrounding environment, we have to change too. Regarding the issue that concern us today is not so serious; our life is not in danger. But, is it not absolutely necessary to speak different languages in the multicultural societies we are involved nowadays?

Bilingualism was born as a consequence of this multicultural immersion. But, what does it exactly mean? At first sight the concept seems to be easy to understand. Have you ever try to define it properly? If you try to create or even to find a definition, you will realize how difficult to describe it is. Not only the individual have to be considered, other global factors that affect the bilingual condition should be taken into account. Different disciplines as psychology, neurology, bilingual education or linguistics describe it focusing on the factors which are more relevant in their fields and do not get to an agreement about a general definition (Beardsmore, 1986). In 1961, Webster's dictionary defines bilingualism as "the constant use of two languages". Bloomfield (1933) considers that being bilingual is "the native-like control of two languages", including only "perfect bilinguals". In 1967, Macnamara considered bilingualism as a global concept including four competences, listening comprehension, speaking, reading and writing. He considered that anyone able to minimally control at least one of this linguistic skills could be considered as a bilingual. The definition proposed by the Weber’s dictionary could seem too over-generic while other explanations are only based on individual skills. Extremes are not usually the best choices, and obviously, lots of definitions between these two points of view were proposed. For instance, Titone (1972) defines bilingualism as the individual’s ability to speak a second language (L2) while following the concepts and structures of that language without remembering their first language (L1). In 1985, Grosjean proposed that a bilingual is much more than the sum of two monolinguals. A bilingual develops specific skills based on his/her language behaviour. Finally, with the aim of adopting a more global concept of bilingualism, Hamers and Blanc (1983) include the distinction between bilingualism and bilinguality. They assume that “bilingualism” is the situation of a linguistic community in which two languages are in contact and both codes can be used in a same interaction. The concept of “bilinguality” considers the psychological state of a speaker. This state allows the individual to use different linguistic codes as mean of communication.

All these definitions conceive the bilingual skills as crucial factors but, how can we measure them? When can I consider myself as a bilingual? There is a lack of precision; we never know where the limit is. In general, literature about L2 acquisition assumes different steps across L2 development. Novel L2 learners are usually mentioned as “novel bilinguals” whilst participants with a medium L2 level are referred to as “relatively fluent bilinguals” and the terms “expert, fluent or highly proficient bilinguals” are generally used.

Although there is not a consensus about a single definition of bilingualism and their limits seem to be blurred, bilingualism is a real fact in our daily lives. It’s very common and much more the rule than the exception in most places. It is estimated that around two-thirds of the people in the world are bilinguals (Crystal, 2003). In European countries, most of the children are able to speak at least two languages and two thirds of working-age Europeans between 25 and 64 years old know a foreign language (FL) (Statista Portal, 2017). However, following the information of the Internet World Stats (2017), in Spain the scenery is not so good. Only 51% of population between 25 and 64 years old claims to know another language apart from Spanish and only 19% of them are really fluent. In addition, Spain is among the seven

European countries with worse English level (Statista Portal, 2017). This issue is a real problem, English is the official international language and this fact could affect the Spanish society in many aspects. For this reason, Spain and other countries in a similar situation have to face and solve this linguistic impasse as soon as possible over the next few years.

Linguistic immersion seems to be the best choice to learn a language (Linck, Kroll, & Sunderman, 2009; Segalowitz & Hulstijn, 2005). In different studies, participants included in linguistic immersion programs has proven to obtain better learning results relative to students in regular L2 learning programs. These results have been obtained in different types of immersion (early, late, delayed, total and partial) and for students with different kinds of educational necessities (Genesee, 2014; Lambert & Tucker, 1972; Swain & Lapkin, 1982). Unfortunately, it is not always possible to learn a language abroad. For this reason, a critical issue in L2 vocabulary learning investigation is the search of strategies able to make the L2 learning process easier in L1 contexts not only at school for children but also for adults.

Our objective with this work is to present strategies able to enhance the L2 learning process. If the reader wants to be able to really understand the reasons why different strategies help or hinder the acquisition process, it is necessary to know the cognitive architecture underlying the development of fluency in L2. To achieve this goal, in the next section we provide theoretical and experimental background about the topic. Two lines of strategies will be explored in this thesis. On the one hand, the role of gestures in L2 vocabulary acquisition. The use of gestures seems to be a good way to enhance vocabulary acquisition because movements promote the formation of a mental image of the word meaning (Goldin-Meadow & Alibali, 2013; Hostetter & Alibali, 2008; see McCafferty & Stam, 2008, for a review). On the other hand, learning results of semantic and lexical based instructions will be compared. Semantic training methods based on the presentation of L2 words coupled with pictures denoting their meanings (picture-word association method) seem to enhance the learning process relative to lexical strategies where L1 and L2 words are presented together as learning tool (word-word association method) (Comesaña, Perea, Piñeiro, & Fraga, 2009; Lotto & De Groot, 1998; Poarch, Van Hell, & Kroll, 2014; Tonzar, Lotto, & Job, 2009).

In the next subsections, different theoretical and experimental approaches about the role of gestures, semantic methodologies and lexical programs in L2 words learning are discussed. The reader is completely free of continue reading this work as he/she desires but I'm going to explain how I would do it. Firstly, the two lines of research previously mentioned follow a common path in order to draw up a theoretical framework on L2 acquisition and processing. After that, we need to split them for a deeper individual analysis. Theoretical information and previous experimental works will be presented separately before introducing the experimental series proposed for both research lines. After this long walk, both paths will find a convergence point in a general discussion and conclusion section. To clarify, this idea is reflected in Figure 1. I propose the reader to continue with the gestures experimental part (page 67) once she completes the reading of the gestures as L2 learning tool section (page 34). After that, the reader would continue with the presentation of the semantic vs. lexical strategies part (page 46) and its corresponding experimental section (page 135) before dealing with the general discussion and conclusions.

We hope you enjoy reading.

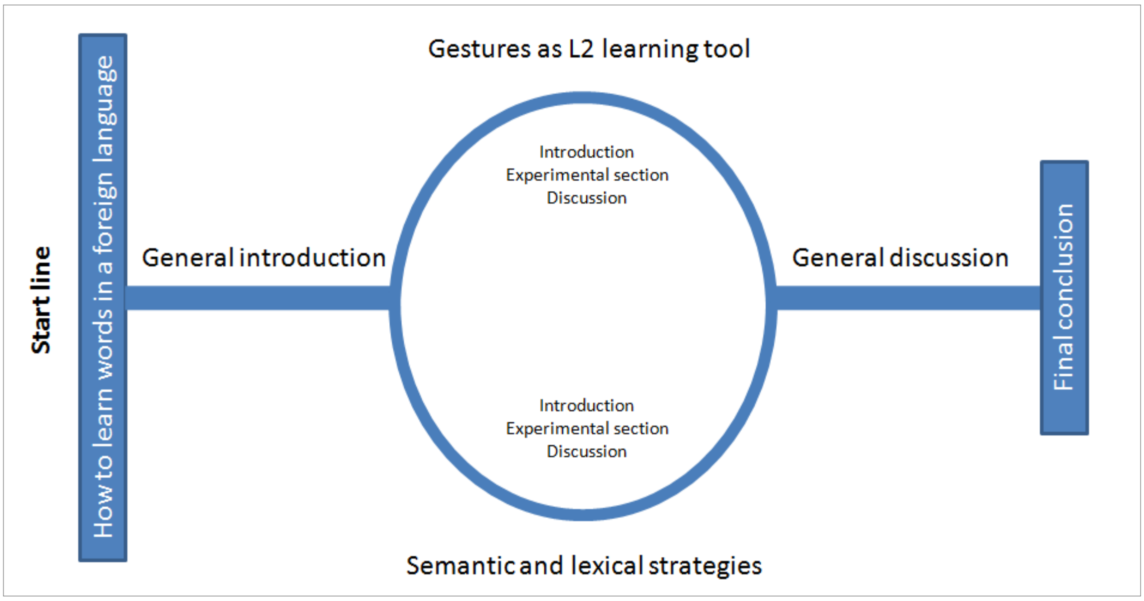


Figure 1. General overview of the thesis.

L2 PROCESSING IN NOVEL AND EXPERTS BILINGUALS

The most efficient way of producing a word in L2 is to retrieve the word in that language from the concept the speaker wants to produce. Therefore, fluent bilinguals are characterized by strong connections between semantic representations and L2 lexical information (Kroll & Stewart, 1994). When bilinguals want to communicate in L2, they access the L2 words directly from their concepts. The retrieval of the translations in the L1 would be an unnecessary step that might add noise to the communication process (Poarch et al., 2014). These ideas were embodied in one of the most influential models of L2 acquisition, the Revised Hierarchical Model (RHM) (Kroll & Stewart, 1994).

The Revised Hierarchical Model

The RHM combines asumptions from two different word-to-concept mapping models (Poarch et al., 2014), the word association model and the concepts mediation model (Potter, So, Von Eckhardt, & Feldman, 1984). The RHM was initially developed to explain L2 processing in novel bilinguals who initiated their L2 exposure after early childhood. In the model, the L1 and L2 lexicons are two separate systems which are interconnected. The last node corresponds to the semantic system (SS), the space where the conceptual information is stored (see Figure 2). Moreover, the L1 and L2 lexical systems are connected to the SS. In the RHM, all connections are flexible and their strength is modulated depending on the level of proficiency in L2 (De Groot & Poot, 1997). In novel bilinguals, the links between the SS and the L2 words are weak and L2 learners preferentially use a lexical route of processing from L2 to L1 when they need to access the conceptual information. As proficiency increases, the links between L2 words and the SS become stronger while the reliance on the lexical route decreases (see Kroll, Van Hell, Tokowicz, & Green, 2010, for a review). Hence, in fluent bilinguals, the L1 and the L2 lexicons are more symmetrically connected to the SS relative to novel learners.

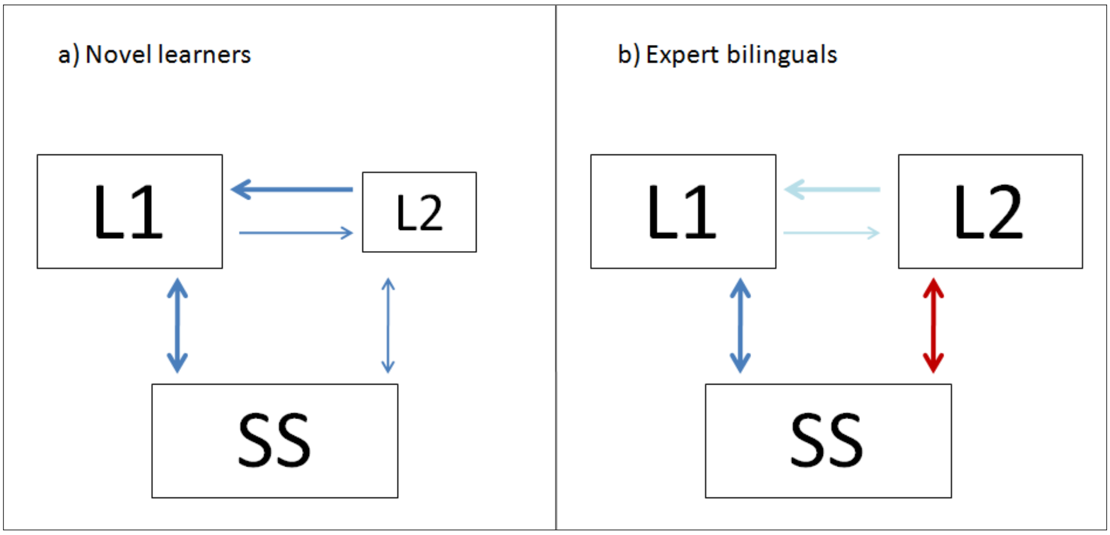


Figure 2. Revised Hierarchical Model. Representation of the first (L1) and the second language (L2) lexicons and their connections to the semantic system (SS). Cognitive architecture of (a) novel bilinguals, (b) expert bilinguals.

The RHM has been widely supported by previous research (see Kroll, Dussias, Bice, & Perrotti, 2015, for a review). Strong evidence in favor of increased semantic processing in fluent bilinguals relative to L2 learners comes from studies in which the translation recognition task is used (Talamas, Kroll, & Dufour, 1999). In this task, participants receive a word in one language followed by a word in another language and they have to indicate whether this last word is the translation of the former. Talamas and colleagues found that English (L1) speakers who were learners of Spanish (L2) took more time to reject L2-L1 words as correct translation when these words were similar in form (e.g., *man-hambre*; *hambre* means hungry and *hombre* means man in Spanish). In contrast, fluent bilinguals suffered more interference when the word pairs were incorrect translations but were semantically related (e.g., *man-mujer*; *mujer* means woman in Spanish). These results suggest that novel learners seem to use the lexical route of processing while fluent bilinguals seem to rely more on semantic information when the L2 word is presented.

Translation tasks can also support the RHM assumptions. In this task, L1 or L2 words are presented for translation (forward and backward translation, respectively). The RHM argues that the forward translation direction (L1-L2) is more likely to occur via semantic information while the backward translation (L2-L1) would preferably occur via between language lexical connections (Kroll et al., 2010). In 1995, Sholl, Sankaranayanan and Kroll, implemented a transfer paradigm experiment between picture naming and translation. After naming pictures (a task that has been proved to require semantic processing), participants performed a forward and backward translation task. Some of the words presented for translation were previously practiced in the naming task. The results showed that when participants translated from L1 to L2, their performance was affected by the items that were presented in the naming task while the transference was absent in the L2-L1 translation direction. In the forward translation direction, participants need more time to respond because the L1 word does not immediately activate the L2 translation, but would cause the activation of the semantic representation of such word. This access to the semantic information slows down the forward translation process. On the contrary, the backward translation direction activates the direct lexical route between languages, resulting in faster performance relative to the forward translation.

Moreover, supporting the RHM, in a picture naming task, L1 naming is faster than L2 naming in novel and experts bilinguals. As mentioned above, the underlying assumption is that the processing of a picture requires access to its meaning. The response time difference between languages in the naming task is larger in low proficient speakers (Kroll, Michael, Tokowicz, & Dufour, 2002). In novel learners the L2 naming task is usually mediated via the L1 word and hence, the access to the L1 lexical information before accessing the SS slows down the naming process (see Kroll et al., 2010 about the asymmetrical connections between L2 and SS for word processing and production).

Challenges to the Revised Hierarchical Model

Although most of research on L2 acquisition can be accommodated within the RHM, some studies have found mixed results (Dufour & Kroll, 1995; Ferré, Sánchez-Casas, & Guasch, 2006; Kroll & Tokowicz, 2001, 2005; Sunderman & Kroll, 2006).

Semantic access in low proficient bilinguals. In this context, different studies have shown that even less proficient bilinguals are sensitive to semantic information. Empirical evidence demonstrates that L2 word-to-concept links might be available for L2 learners at an earlier point in their L2 acquisition than previously assumed (Altarriba

& Mathis, 1997; De Groot & Poot, 1997; Ferré et al., 2006; Frenck-Mestre & Prince, 1997; Potter et al., 1984; Talamas et al., 1999; Sunderman & Kroll, 2006). To illustrate, Frenck-Mestre and Prince (1997; Experiment 2) examined priming of the dominant and subordinate meanings of biased homographs (e.g., *seal*) in a L2 lexical decision task. The priming effect for the subordinate meaning was only found in proficient bilinguals; however, less and more proficient bilinguals showed priming effects for the dominant meaning of the homograph in L2. Thus, L2 learners used a semantic route of processing (although to a lesser extent than fluent bilinguals).

Studies conducted with the translation recognition task have also found evidence for the use of a semantic route of processing in early stages of L2 acquisition. In a study conducted by Talamas and colleagues (1999), there was evidence that less proficient bilinguals processed L2 words semantically. In particular, L2 learners showed semantic interference when the L1-L2 word pairs shared high semantic similarity (while fluent bilinguals showed the effect for both high and low semantic similarity).

Moreover, Sunderman and Kroll (2006) obtained evidence of semantic processing in L2 learners with the translation recognition task (see also, Altarriba & Mathis, 1997). In particular, the authors compared two groups of native English (L1) speakers, one with a greater level of proficiency and the other with less proficiency in Spanish (L2). The two groups of participants were slower to reject the pairs that were related in meaning compared with a control condition tested with unrelated word pairs. Thus, even the less proficient L2 learners appeared to be sensitive to conceptual information. These results seem to play against the RHM assumptions, however if the task's features are analyzed, this fact is easily accommodated within the model by assuming the existence of asymmetrical connections between L2-SS and SS-L2. It should be mentioned that novel learners can directly access the semantic system during comprehension tasks whilst participants with the same level of proficiency are not able to verbalize concepts in L2 in production tasks (Duyck & Brysbaert, 2004; Dufour & Kroll, 1995). Whereas the L2-SS links seems to be more sensitive to the initial steps in L2 acquisition, more effort must be done to strength the SS-L2 links (Kroll et al., 2010; Schwieter & Sunderman, 2009). Recently, this obstacle has been interpreted as a consequence of deficient abilities at the lexical selection stage in novel learners.

Behavioral and neuroimaging studies have shown that the activity of the non-target lexical candidates have to be reduced in order to carry out production tasks. Proficient bilinguals are more experts at that and a more symmetric pattern of results for comprehension and production processes appears at this level (Costa & Santesteban, 2004; Kroll, Bobb, Misra, & Guo, 2008).

Translation directions and semantic sensibility. The RHM posits that the L1-L2 translation direction is more likely to occur via semantic processing than the L2-L1 direction. In fact, the model was initially proposed to explain the asymmetries between translation directions that are usually present in this type of tasks. However, evidence confirms that both translation directions are sensitive to conceptual variables such as the words concreteness (Van Hell & De Groot, 1998). One more time, proficiency of individuals in L2 seems to plays a crucial role. These results were obtained in relatively proficient bilinguals which is not surprising, because, as the RHM explains, the more proficiency in L2 the more available the L2-SS connections are, and the lesser differences between translation directions are expected (Kroll et al., 2002).

The more controversial findings about conceptual processing in novel L2 learners is the observation of several effects (number magnitude effect and concreteness effect) which are due to semantic processing of the material (De Groot & Poot, 1997; Duyck & Brysbaert, 2008). The type of material used in the experiments could explain the presence of semantic effects in novice L2 learners. In particular, the ease with which words are processed would determine whether or not they access the semantic system. Thus, the translation asymmetry would be observed when words are difficult to process (abstract vs. concrete words, low vs. high frequency words, non-cognate vs. cognate words). For example, when abstract non-cognate words are presented, even proficient bilinguals show an asymmetry between translation directions (Van Hell & De Groot, 2008). To conclude, even for highly skilled bilinguals, the type of words determines whether they are processed semantically (Altarriba & Mathis, 1997; Duyck & Brysbaert, 2008; Kroll & Tokowicz, 2005; Kroll et al., 2010). Note that the presence of semantic effects in novel L2 learners can be accommodated within the RHM because, in the model, the conections between the SS and L2 lexicon are weak but they are present at the beginning of the L2 learning.

To summarize, previous studies seem to indicate that both the lexical route and the semantic route of processing are present in L2 learners. However, the weight of these routes appears to depend on the fluency of bilinguals in L2; individuals in early stages of L2 learning would primarily use the lexical route whilst processing would be predominantly semantic in fluent bilinguals. Thus, if we consider the cognitive processing of fluent bilinguals, it would be desirable to implement L2 learning methods that favour the establishment of connections between concepts and L2 words.

At this point, you have theoretical information about L2 processing and a rough idea about the importance of the semantic integration during L2 words learning. Now, I would like you to guess, what strategy would you follow to learn a new language? We encourage the reader to continue reading to solve this question and you may find an answer in next chapter.

CHAPTER II

L2 Learning Strategies

INTRODUCTION

There are several questions about L2 acquisition that needs to be answered. What is the best way to learn a new language? Are there strategies able to enhance this learning process? It seems that immersion programs are good alternatives to learn a new language (Genesee, 2014) but they are not always available for people interested in acquiring a new language. Studies on the acquisition of L2 in childhood and adulthood are, therefore, of particular interest. It would be desirable to implement learning methods that are theoretically based and supported by scientific studies to demonstrate their effectiveness in L2 acquisition.

Early methods of L2 vocabulary acquisition used a word association approach that fostered the establishment of connections between newly acquired L2 words and their translation equivalents in the L1. For example, in the keyword method (Atkinson & Raugh, 1975; Raugh & Atkinson, 1975), a L1 word that sounds like some part of a L2 word is used in the learning process (the keyword). Firstly, learners have to associate the spoken foreign word with the keyword and, afterwards, to associate the keyword with the L1 translation of the word that has to be learned in L2. Previous research has proven the efficacy of these methods for early stages of L2 learning due to the establishment of lexical connections between L1 and L2 (Atkinson, 1975). However, when fluent bilinguals want to express themselves in L2, the best route of processing would be the direct access to L2 words from their concepts. The use of between-languages lexical connections and the retrieval of L1 words would be an unnecessary step when bilinguals communicate in L2. So, if we know how fluent bilinguals process their L2, looking for learning methods able to mimic this pattern of processing could be a good learning strategy.

Previous research has confirmed that the acquisition of connections between L2 words and concepts is fostered by the use of training protocols that involve semantic processing (Barcroft, 2002; Comesaña, Soares, Sánchez-Casas, & Lima, 2012; De Groot & Poot, 1997; Finkbeiner & Nicol, 2003; Kroll, Michael, & Sankaranarayanan, 1998; Poarch et al., 2014; Wimer & Lambert, 1959). To illustrate, it has been proposed that gestures are a good L2 learning tool because speech and movements interact and they produce an integrated representation of the meaning of the word in memory (Kelly, Özyürek, & Maris, 2010; McNeill, 1992, 2005; McNeill, Levy, & Duncan, 2015). In addition, the presentation of L2 words with pictures denoting their meanings (picture association method, Finkbeiner & Nicol, 2003) favors the learning process relative to the learning of L2 words presented with their translations in the L1 (word association method, Altarriba & Mathis, 1997; Lotto & De Groot, 1998; Van Hell & Candia-Mahn, 1997). Similarly, imagining the meaning of words to be learned in a FL enhances the acquisition process (Ellis & Beaton, 1993; Wang & Thomas, 1995). In fact, the Dual Coding Theory (Paivio, 1971) suggests that the formation of mental images during the learning process might favor the acquisition of new words. According to this theory, verbal information, visual images and movements are integrated, and this increases the probability of remembering new words compared to the use of verbal glosses alone.

GESTURES AS L2 LEARNING TOOL

Gestures have been proposed as a good tool that facilitates the L2 vocabulary acquisition process (Kelly et al., 2010; McNeill, 1992, 2005; McNeill et al., 2015). In all spoken languages, speakers accompany their speech with visual-manual communication (Özyürek, 2014). This type of multimodal interaction, called co-speech gestures, involves spoken language, facial expressions, body movements and, especially, hands movements. All these visual and auditory aspects act as an integrated stream of information which improves the communication process (Holle & Gunter, 2007). In fact, Marstaller and Burianová (2014) showed that the right auditory cortex and the left posterior superior temporal brain areas seem to reflect the multisensory integration of spoken language sounds and gestures.

Previous studies have demonstrated the role of iconic gestures in language comprehension (e.g., speech comprehension and gestures, Straube, Green, Weis, & Kircher, 2012; see Yang, Andric, & Mathew, 2015, for a review), and production (Goldin-Meadow & Alibali, 2013, for a review of gestures in speaking). It has been observed that performing gestures during the learning process facilitates the acquisition and recall of words in a FL (Macedonia, Müller, & Friederici, 2011; Quinn-Allen, 1995; So, Chen-Hui, & Low Wei-Shan, 2012; Tellier, 2008; see Macedonia, 2014; and Macedonia & Kriegstein, 2012, for reviews). But, is this effect only observed when L2 learners perform gestures? Previous research indicates that the mere observation of movements is able to activate motor areas at brain level which are related with motor actions (Rizzolatti & Craighero, 2004; Stefan et al., 2005). Thus, although movement performance appears to improve learning, it is not clear whether a learning protocol that involves self-generated actions would have an additional benefit to the mere observation of movements. This issue would be addressed in the following section.

The role of movements on learning

In the educational system, the possible advantages associated to the learning through actions relative to the observation-based learning have been discussed for decades (Goldin-Meadow, 1999). The “learning-by-doing” perspective defends the active participation of the individual in the learning process by performing actions during educational development. Learning-by-doing can positively affect the formation of neural networks underlying the acquisition of knowledge and the performance of many cognitive skills (Goldin-Meadow et al., 2012). This beneficial effect has been demonstrated in several educational fields such as online courses, language learning, or technology use (Aleven & Koedinger, 2002; Bessen, 2015).

Moreover, the relevance of movement in language processing has been confirmed in many studies (Glenberg, Gutiérrez, Levin, Japuntich, & Kaschak, 2004; Koriat & Pearlman-Avion, 2003). For instance, Glenberg, Sato, and Cattaneo (2008) observed that the performance of movements interacts with language comprehension. In their study, participants were required to move 600 beans (one at a time) from a big recipient to a narrow container toward or away from their bodies depending on the recipients' location. Afterwards, meaningful and meaningless sentences describing movements were presented to participants and they judged their plausibility. Results revealed that the time required to judge the sentences depended on the congruency between the beans' movement direction and the direction of the acts described in the sentence (toward or away from the body). Therefore, the performance of actions determined language comprehension.

In this context, several studies have examined the differences between self-performed tasks and experimenter-performed tasks (Cohen, 1981; Engelkamp & Zimmer, 1983). In 2012, Goldin-Meadow and colleagues directly compared the role of self-performed gestures versus seeing another individual producing them when children learned a mental transformation task. In their study, children were required to perform a mental rotation task in which they had to indicate whether two shapes presented in different orientations were the same figure or not. They used this task because previous studies demonstrate the close relationship between mental rotation and motor processing. When participants are instructed to mentally rotate a target, premotor areas involved in the planning of actions become active (Ganis, Keenan, Kosslyn, & Pascual-Leone, 2000; Glenberg et al., 2008) and participants spontaneously make gestures when they are required to explain how they solve this task (Chu & Kita, 2008). Goldin-Meadow et al. (2012) showed that children obtained better results when they were instructed to produce the gesture needed to solve the transformation task rather than when they observed the experimenter doing the movements. Thus, the performance of movements improved the learning compared to the mere exposure to gestures associated with the mental rotation of objects.

Other empirical studies have confirmed the importance that self-generation of movements have in the learning of linguistic material. Engelkamp, Zimmer, Mohr, and Sellen (1994) asked participants to learn sentences while performing the actions described in the sentences or by simply listening to and memorizing the material. The results revealed that the recall of sentences was higher when participants performed actions during the learning phase. The authors interpreted that the performance of actions favored the formation of a motor trace that benefited the retention of information. Empirical evidence supporting this interpretation comes from the study conducted by James and Swain (2011). The authors taught children action words associated to concrete toys. Some of the children manipulated the objects while learning and the remaining children observed the experimenter manipulating the same objects. When children listened to the words they had previously learned, motor brain areas were activated only in children who perform the toys manipulation themselves. Thus, the performance of motor actions favor the learning of new words and the benefit associated with the performance of movement during learning seems to be due to the formation of a motor trace that would be activated during the subsequent retrieval of information.

However, other works have found similar pattern of results when participants produce actions and when they only see actions produced by others (e.g. Rizzolatti & Craighero, 2004). For example, Stefan and colleagues (2005) observed the involvement of the motor cortex during the observation of movement (simple repetitive thumb movements) which produced a specific memory trace in the motor cortex similar to the pattern of activation that occurs when people perform a motor action. In a study addressing the effect of movements on sentences reading comprehension in children, Glenberg and colleagues (2004, Experiment 3) found intermediate results. Children were exposed to histories happening in a particular scenario (a farm) where different referents appeared (a sheep or a tractor). For the first group of children, objects referred to in the text were present and they were instructed to simulate the sentence content by manipulating the objects. The second group of children was required to imagine they were manipulating the toys. The results showed a beneficial effect of the manipulation condition while the imagined condition presented a modest improvement compared to an only-read condition. Thus, although movement performance appears to improve learning, it is not clear whether a learning protocol that involves self-generated actions would have an additional benefit to the mere observation of movements.

In general, it is widely accepted that movements produce a clear impact in the learning process. However, before attempting an in-depth analysis of the role of gestures in L2 learning, a brief summary of the theoretical frameworks concerning the relationship between gestures and speech is offered.

Theoretical frameworks regarding the relationship between gestures and speech

There are several frameworks to explain the connections between gestures and speech. All these frameworks address the nature of representations underlying the processing of gestures. One way to differentiate between models is to consider the relevance of visuospatial and linguistic information. Some views suggest that the representation of gestures is based on visuospatial images (e.g., the sketch model, de Ruiter 2000; the interface model, Kita & Özyürek, 2003; the gestures-as-simulated-action (GSA) framework, Hostetter & Alibali, 2008). In other models, the emphasis is placed on the close relationship between the representation of gestures and linguistic information (e.g., the interface model, Kita & Özyürek, 2003; the growth point theory, McNeill, 1992, 2005). Another way to differentiate between models is by attending to the way gestures and speech are processed (separate vs. unitary processing). In some models, it is considered that gestures and speech are processed in two separate systems (e.g., lexical gesture process model, Krauss, Chen, & Gottesman, 2000) that interact when communicative intentions are formed (sketch model, de Ruiter, 2000), or at the conceptualization stage (interface model, Kita & Özyürek, 2003) to produce effective communication. In other models, however, it is assumed that gestures and speech work together as two parts of the same system (the growth point theory, McNeill, 1992, 2005; the GSA, Hostetter & Alibali, 2008). The gesture-in-learning-and-development (GLD) model (Goldin-Meadow, 2000, 2003), for example, considers that children process gestures and speech independently and they become part of an integrated system in mature speakers (Butcher & Goldin-Meadow, 2000; Özçaliskan, Gentner, & Goldin-Meadow, 2014).

In addition to the specific mechanism by which gestures and speech are processed, a relevant question concerns the role of gestures in communication. It has been confirmed that listeners glean information from gestures (Alibali, Flevares, &

Goldin-Meadow, 1997; Cassell, McNeill, & McCullough, 1999; Goldin-Meadow, 2003; Goldin-Meadow, Wein, & Chang, 1992; Holler, Shovelton, & Beattie, 2009; Singer & Goldin-Meadow, 2005; see Hostetter & Alibali, 2008 for a review; Hostetter, 2011, for a meta-analysis demonstrating the benefits of gestures for communication). In fact, since gestures usually arise during speech planning, many models defend the communicative value of gestures (e.g., the sketch model, de Ruiter, 2000; the interface model, Kita & Özyürek, 2003; the growth point theory, McNeill, 1992, 2005; and the GLD framework, Goldin-Meadow, 2000, 2003).

Here we have sketched the role of gesture in communation. However, in our research, we are interested in the impact that gestures have on L2 acquisition. This issue is directy addressed in the next section.

Gestures and L2 learning

The role of different types of gestures in L2 learning has been emphasized in several studies (e.g., De Grauwe, Willems, Rueschemeyer, Lemhöfer, & Schriefers, 2014; Gullberg, 2014; Kelly et al., 2010; Macedonia & Knösche, 2011; Macedonia & Kriegstein, 2012; Morett, 2014; So et al., 2012; McCafferty & Stam, 2008, for reviews). In general, it is widely assumed that gestures have a positive effect on vocabulary acquisition and they should be used in FL instruction and embedded in a natural approach of language teaching (Asher & Price 1967; Carels, 1981; Krashen & Terrell, 1983; Macedonia, Bergmann, & Roithmayr, 2014; however, see Hirata, Kelly, Huang, & Manansala, 2014; Kelly, Hirata, Manansala, & Huang, 2014, for evidence about the limited effect of gestures on learning segmental phonology).

In cognitive psychology, three main perspectives might account for the facilitative role of iconic gestures in L2 vocabulary learning. The *self-involvement explanation* states that gestures might favor the involvement of the participant in the learning task and therefore, gestures would have a key role in the L2 vocabulary learning process (Helstrup, 1987) and they could facilitate enhanced attention to learning material. In particular, the impact of gestures on vocabulary learning is caused primarily by increased perceptual and attentional processes occurring during proprioception of movements associated with gestures or when individuals use objects to perform the action (Bäckman, Nilsson, & Chalom, 1986). However, under this view, the benefits associated with the use of gestures does not come from enactment itself because the motor component is not crucial (Kormi-Nouri & Nilsson, 2001); rather, it is the multisensory information conveyed about a word that leads to greater semantic processing and higher attention level (Knopf, 1992; Knudsen, 2007). Therefore, according to the self-involvement explanation, the learning of new words with gestures facilitates vocabulary acquisition regardless of whether a gesture is usually produced within a language or it denotes the same meaning of the word to be learned. Attention increases the retention of words (Craik & Tulving, 1975; see Muzzio et al., 2009, for the role of attention in the encoding and retrieval of information at the neurobiological level). Thus, individuals using gestures would learn L2 words in an attentive manner, which would favor the learning process. The *motor-trace perspective* suggests that the physical component of gestures is coded in the learning process, leaving a motor trace in memory that aids the acquisition of new words in L2 (Engelkamp & Zimmer, 1984, 1985). According to this view, the physical enactment is crucial because it allows the formation of a motor trace associated with the meaning of the word. There is recent evidence from neuroscientific studies (e.g., repetitive transcranial magnetic stimulation) supporting the role of the motor cortices in the understanding of written words (Vukovic, Feurra, Shpektor, Myachykov, & Shtyrov, 2017). Moreover, there is evidence that the use of familiar gestures might engage procedural memory since they involve well-defined motor programs (Macedonia & Mueller, 2016). Thus, the involvement of procedural memory along with declarative memory used for word processing might enhance vocabulary learning. From this perspective, familiar gestures that have been routinely practiced and are within the repertoire of the speaker (e.g., the gesture of answering the phone) have strong motor traces so that they produce more facilitation than less familiar gestures (e.g., the gesture of moving the finger from the mouth to the ear). Thus, the facilitation effect would depend on the extent to which an individual is familiar with certain gestures. However, according to this view, the effect of gestures would operate independently of their meaning; that is, regardless of the congruency between the gesture and the word meaning, familiar (well practiced) gestures would facilitate the learning process relative to less familiar gestures. Finally, the *motor-imagery perspective* indicates that gestures are associated with motor images that form part of a word's meaning (Denis, Engelkamp, & Mohr, 1991). To be more specific, performing a gesture when individuals process a word promotes the creation of a visual image associated with the meaning of this word, which would enrich the semantic content of the word to be learned. This image would be a mental representation of the action associated with the word during encoding (Macedonia & Knösche, 2011; Macedonia et al., 2011). In fact, neurobiological evidence obtained with functional connectivity analyses suggests the involvement of the hippocampal system in binding visual and linguistic representations of words learned with pictures (Takashima, Bakker, Van Hell, Janzen, & McQueen, 2014). Hence, according to this view, the facilitation effect observed with gestures would be greater for iconic gestures related to the meaning of words to be learned than when there is a mismatch between the gestures and the meaning of the word. In addition, the learning of words accompanied by gestures with incongruent meanings would produce semantic interference and reduced recall (Cook, Yip, & Goldin-Meadow, 2012; Yap, So, Yap, Tan, & Teoh, 2010).

It is important to note that the three perspectives described above are not mutually exclusive but instead emphasize different aspects of the effect of gestures in L2 learning. Hence, a gesture accompanying a word might increase self-involvement (co-speech gestures as movements that enhance attention to the L2 learning), create a motor trace (co-speech gestures as meaningful movements) due to the physical enactment associated with the production of gestures, or a semantic visual image that could become integrated with the meaning of the word (co-speech gestures as semantically congruent movements).

Empirical evidence regarding the impact of gestures on L2 learning

The first empirical study concerned with the impact of gestures on L2 learning was conducted by Quinn-Allen (1995). In this study, English participants had to learn French expressions under two conditions. In the control condition, the participants were presented with sentences in French (e.g., *Veux-tu quelque chose à boire?* Do you want something to drink?), which they had to repeat in French. In the experimental condition, the learners received the sentences with an emblematic gesture illustrating the meaning (e.g., pointing the thumb towards the open mouth) and they were then required to reproduce the gesture. The results showed that sentences presented with gestures were associated with better recall in comparison with control sentences. The facilitation effect produced by gestures in Quinn-Allen's (1995) study can be accommodated within the self-involvement account (Helstrup, 1987). Participants would be more engaged in the learning process when they received and produced gestures relative to the control condition without gestures. However, the facilitation effect could also be explained by the motor-trace account and the motor-imagery explanation. The gestures used by Quinn-Allen, such as the act of drinking, are conventional gestures frequently produced in social communication, which could explain the facilitation effect according to the motor-trace view. In addition, this gesture is congruent in meaning with the sentence to be learned (do you want something to drink?) and, hence, a facilitation effect would be expected on the basis of the motor-imagery account. In short, a disadvantage of comparing only a condition with gestures with a condition without gestures is that it is not possible to distinguish whether the benefits observed in the gesture condition come from the use of familiar gestures, gestures with related meanings, and/or the mere fact of performing gestures. Additional experimental work has resolved this problem by using paradigms in which several gesture conditions are compared (Kelly, McDevitt, & Esch, 2009; Krönke, Mueller, Friederici, & Obrig, 2013; Macedonia et al., 2011; Macedonia & Knösche, 2011).

In Macedonia et al.'s (2011) study, a group of German speakers were trained in Vimmi, an artificial language that has the advantage of controlling for several linguistic variables such as word length, familiarity of L2 words and phonotactic factors (see Jalbert, Neath, Bireta, & Surprenant, 2011, for a discussion of these variables in vocabulary learning). The authors compared the learning of concrete nouns presented with iconic gestures (e.g., the word suitcase appeared with the gesture of an actor lifting an imaginary suitcase) or meaningless gestures (e.g., the word suitcase and the gesture of touching one's own head). The results showed better recall for words learned with iconic gestures relative to words accompanied by meaningless gestures. The findings of this study appear to indicate that gestures involve something additional to the self-involvement of the participant in the task, since participants were exposed to gestures in the two experimental conditions. Both the motor-trace and the motor-imagery accounts could explain the better memory performance in the iconic gesture condition compared with meaningless gestures. Iconic gestures would favor L2 learning because they are semantically rich and because they are produced more frequently than meaningless gestures, hence producing strong motor activation. In fact, the results found in Macedonia et al.'s study favor both explanations. The authors observed activity in the premotor cortices for words encoded with iconic gestures, which is compatible with the motor trace account. Moreover, words learned with meaningless gestures produce activity in a network engaged in cognitive control suggesting that individuals detected the mismatch between the meaning of the word and the gesture.

Other studies have made use of additional experimental conditions to differentiate between explanations based strictly on the motor component of gestures as opposed to those based on the motor-imagery account. Studies performed in monolingual contexts have explored congruity effects during communication by mismatching the information between the semantics of words and the meanings of the gestures (Barbieri, Buonocore, Dalla Volta, & Gentilucci, 2009; Bernardis & Gentilucci, 2006; Bernardis, Salillas, & Carameli, 2008; Chieffi, Secchi, & Gentilucci, 2009; Feyereisen, 2006; see Kircher al., 2009 for congruity effects in the context of an unknown language). To illustrate, Kelly, Creigh, and Bartolotti (2010; see also, Kelly, Healey, Özyürek & Holler, 2015) conducted an event-related potential study combined with the use of a Stroop-like paradigm. The participants were presented with words (e.g., "cut") and gestures that could be congruent (e.g., the act of cutting) or incongruent (the act of drinking). The authors found reduced N400 to words accompanied by congruent gestures relative to incongruent gestures (a semantic integration effect, Kutas & Hillyard, 1980). In addition, response times were faster in the congruent condition compared with the incongruent condition. The pattern of results found in Kelly et al.'s study is similar to what is observed in other Stroop tasks and seems to indicate that gestures are integrated with the meaning of words, producing interference when the word meaning does not match the meaning of the gesture (incongruent condition). In our view, the results of this study strongly agree with the motor-imagery account of gestures (see also Macedonia et al., 2011). No differences would be expected according to the self-involvement account since both conditions, congruent and incongruent, involved gestures. Moreover, in the congruent and incongruent conditions, familiar gestures were used so they would promote strong activation of motor trace representations and no differences between the two conditions would be found. On the contrary, clear differences between the congruent and incongruent condition would be predicted based on the motor-imagery view since they would differ in the degree to which the gesture could be integrated with the word meaning.

Nevertheless, there is a limitation with using the results of Stroop-like studies to draw the conclusion that gestures integrate with the meaning of words. In the standard Stroop color-word tasks (see MacLeod, 1991, for a review) three conditions are implemented: a congruent condition, an incongruent condition, and a meaningless condition. The usual finding is a facilitation effect when the congruent condition is compared to the meaningless condition and an interference effect when the incongruent condition is compared to the meaningless condition. However, in some of the studies (cited above) regarding the impact of gestures on L2 learning, the meaningless condition is omitted so it is not possible to determine the amount of facilitation and interference that results from the direct comparison of the congruent and incongruent gesture conditions.

Taken together, the results of previous studies appear to confirm the influence of gestures on L2 vocabulary learning. However, in spite of the merits of previous work in the field, a detailed comparison across several conditions is needed to distinguish the various theoretical explanations. In particular, the acquisition of L2 words should be compared between conditions with and without gestures in order to evaluate the self-involvement account. Furthermore, a condition with familiar gestures (e.g., iconic gestures) might be compared to a condition with unfamiliar gestures in order to evaluate whether the motor-trace of gestures modulates vocabulary learning. Finally, congruent and incongruent conditions need to be compared with a meaningless condition to fully determine the motor-imagery account of gestures in L2 vocabulary acquisition. We acknowledge that some of these comparisons have been made in separate studies. For example, previous work has explored the comparison between congruent and incongruent gestures (e.g., Kelly et al., 2010), as well as the use of a gestural control condition (Cook et al., 2012; Wagner, Nusbaum, & Goldin-Meadow, 2004). However, if we assume that several explanations might work together to explain how gestures interact with L2 vocabulary acquisition, it would be desirable to evaluate in the same study the role of congruent, incongruent, meaningless, and no-gesture conditions.

Moreover, if we consider the overall pattern of results found in previous research, almost all studies have revealed a positive effect of gestures on L2 vocabulary learning. However, this observation contrasts with previous work showing that dual task conditions, in which individuals have to perform two tasks simultaneously, can hinder language learning (e.g., learning at the same time the meaning and form of aural input, see Van Patten, 1990; Wong, 2001). A L2 learning condition in which individuals have to process words and gestures concurrently is a dual task that might negatively influence the learning process. However, in some previous studies concerning the role of gestures in L2 vocabulary acquisition, this possible negative effect has not been captured. In fact, previous studies show beneficial effects of gesturing during the L1 and L2 language processing (Holle & Gunter, 2007; Marstaller & Burianová, 2014; Özyürek, 2014). For example, Quinn-Allen (1995) observed better retrieval of L2 expressions when these were learned with semantically related gestures relative to learning without gestures. In our opinion, the possible negative effect of a dual task condition (the concurrent processing of a L2 sentence and a gesture) might not be observed in this study since it overlapped with a positive effect of congruency between the sentence and the gesture meaning. Therefore, a direct comparison between a no gesture condition and a meaningless gesture condition is required in order to determine the possible negative effect of gestures in L2 vocabulary learning.

Another important question concerns the type of words used as learning material in previous studies. To assess the effect of gestures on vocabulary learning, most of the studies have employed verbs as training material (De Grauwe et al., 2014; Kelly et al., 2009; Kircher et al., 2009). The use of verbs seems to be the best option if we consider that representational gestures depict actions. There is a more direct mapping between gestures and the semantic characteristics of verbs, in comparison with other types of words such as nouns (Childers & Tomasello, 2002; De Grauwe et al., 2014; Kelly et al., 2009; Kircher et al., 2009). However, apart from gestures, it has been confirmed that nouns are easier to learn than verbs, at least for children. Different studies have shown that children are able to learn English nouns easier than verbs in their natural context (Fernald & Morikawa, 1993; Goldfield, 1993; Tardif, Shatz, & Naigles, 1997). It is probably due to the fact that English speakers place special emphasis on nouns when they interact with children while acquiring their L1. However, this advantage is not present in other cultures (see Gentner & Boroditsky, 2001; Gopnik & Choi, 1995; Tardif et al., 1997 for the absence of this advantage in Korean and Mandarin languages). Concerning the role of gestures during the learning of nouns and verbs, the GSA framework makes concrete predictions about what types of words could be more influenced by gestures. This theory states that gestures occur as a result of simulated action and perception, which are the basis of mental imagery and language production (Hostetter & Alibali, 2008). Even if thinking about the size or shape of a particular object (nouns) involves simulating movements, the relationship between verbs and movements is stronger. Gestures would therefore have a greater influence on verb learning than on noun learning.

SEMANTIC VS. LEXICAL LEARNING STRATEGIES

As a house construction must be started at the foundation, this research line begins addressing the role of semantic and lexical learning methods in single L2 words acquisition. However, we know that knowking a language is much more than acquiring isolated words, we cannot reduce our linguistic interchanges to single words. For this reason, we decided to evaluate the role of both learning strategies in sentences processing. Finally, we wanted to go one step ahead and look for the neural correlates of these two training methods with electrophysiological evidence (event related potentials, ERPs).

L2 single words learning

The ways in which we learn determines the cognitive architecture in which lexical and semantic information is represented. There is previous work addressing the efficiency of semantic and lexical learning strategies to acquire foreign-language vocabulary, the majority of which have been conducted with children (Comesaña, et al., 2012; Poarch, et al., 2014). In different studies, the semantic and lexical methods of foreign-language vocabulary acquisition have been implemented with the use of pictures and words, respectively. The underlying assumption is that the processing of a picture requires access to its meaning, and thus the learning of new words in L2 accompanied by pictures denoting their meaning would strengthen the connections between the concepts and the L2 lexical system. Tonzar et al. (2009) examined this issue with Italian (L1) children from 4th to 8th grades. The children learned words in FL (English and German) with two training protocols: a word learning method in which the L2 words and L1 word were presented, and a picture learning method consisting of the L2 word and a picture of an object to which the L1 word referred. The authors observed superiority of the picture-mediated learning over the word-mediated learning method. They concluded that the efficiency of the picture-based method in the acquisition of L2 words might be due to the fact that this training allowed a direct connection between the L2 lexical representation and the corresponding concept. Comesaña et al. (2009, Experiment 2), found direct evidence for this conclusion by examining the semantic interference effect in a translation recognition task. Two groups of Spanish (L1) children learned Basque (L2) words with a word-based method (L1-L2 words were presented) or a picture-based method (L2-picture pairs were used). At testing, the children who learned the L2 words with pictures showed a semantic interference effect and so were slower and they made more errors to semantically related L2-L1 pairs (e.g., *aulkia*, the word chair in Basque, paired with *mesa*, table in Spanish) relative to unrelated pairs of L2-L1 words. This semantic interference effect suggested that the connection between L2 words and the conceptual system in children was a function of the way in which the L2 words had been learned. A picture-mediated L2 learning instruction favored semantic processing in children learning words in a FL. This conclusion was strongly supported in Poarch et al.'s (2014, Experiment 1) study. The authors evaluated a group of Dutch (L1) children aged 10-11 who were L2 learners of English (L2) in contexts enriched by pictures and listening/speaking exercises. The children showed a semantic interference effect in the translation recognition task, which suggests that they were able to exploit conceptual information during L2 processing and map L2 words to concepts.

All of the studies reviewed so far have examined the impact of semantic vs. lexical mediated instruction of L2 in childhood. There are reasons to assume that in adults, the type of learning (e.g., with pictures vs. words) might influence L2 vocabulary learning in a way that is different to what occurs in children. Children are individuals still in the process of creating connections between the semantic system and L1 words, so the semantic/lexical systems are likely to be more sensitive to the use of material that favors the link between concepts and L2 words. In contrast, concepts-L1 links are already formed in adults that decide to learn a L2. Thus, adult individuals might use these concept-L1 connections to access new L2 words (i.e., through a lexical route from L1 words to L2 words) regardless of the training program used to acquire foreign-language vocabulary. If fact, Chen and Leung (1989) showed that Chinese (L1) children were better at retrieving words in English (L2) from pictures (e.g., an L2 picture naming task) while Chinese adults learners of English as L2 performed better when retrieving L2 words from L1 words (i.e., L1-L2 translation task). Thus, adults appear to use a lexical route of processing to learn L2 vocabulary, which might reduce the effect of a picture-mediated L2 learning instruction in adulthood.

There are relatively few studies comparing the impact of semantic vs. lexical methods on the acquisition of L2 words in adulthood. Of these, some have addressed the congruency between the learning method and the evaluation task of foreign-language vocabulary. To illustrate, Chen (1990) asked two groups of Chinese (L1) college students to learn words in French (L2). One group was presented the L2 word paired with a picture depicting its meaning (picture-learning group) while the other received the L2 word with the corresponding L1 translation (word-learning group). At testing, the word-learning group showed better performance on translation tasks with words while the picture-learning groups were faster in naming pictures in L2. Similar picture-learning and word-learning methods were compared by Lotto and De Groot (1998) who asked adult Dutch (L1) speakers to learn Italian words (L2). After learning with pictures-L2 words or L1-L2 words, the learners received either the pictures or the L1 words as cues for the recall of the L2 words. Similar to what was found with children

(Tonzar et al., 2009), adult learners showed that the picture-based methods lead to better performance than the word-based methods. Moreover, the results showed evidence of better recall when the study and test conditions were congruent than in the case in which they were incongruent. Therefore, these studies indicate that the performance of adult L2 learners is better when the learning and test conditions match than when they mismatch.

Finally, evidence about the superiority of semantic vs. lexical L2 learning methods for establishing connections between concepts and L2 words has been gathered in studies where the learning results are only measured after the learning procedure (e.g., translation recognition task, translation task, etc.) (Comesaña et al., 2009). In previous studies, novel learners who follow a semantic instruction have shown advantages relative to lexical learners with a pattern of results more similar to fluent bilinguals (Talamas et al., 1999) in evaluation tasks after training. It suggests that they were able to exploit conceptual information during L2 processing and map L2 words to concepts.

L2 words processing within sentences

Studies conducted with both monolinguals and bilinguals have shown that sentence context modulates single word processing. For example, in electrophysiological studies conducted with monolingual speakers, it is observed that sentence comprehension depends on the lexical properties of the words in the sentence (e.g., lexical frequency, word length, etc.). However, the expectation of an upcoming word within a sentence (cloze probability or semantic constraint) overrides the influence of these lexical factors when the words are highly predicted by the sentence context (Khachatryan et al., 2014; Van Petten, Coulson, Rubin, Plante, & Parks, 1999).

Studies with bilingual individuals have also found that word processing within sentences may differ from processing words in isolation. Specifically, a large number of studies have shown that the L1-L2 lexical links are active even in high proficient bilinguals (Hoshino & Kroll, 2008). This non-selective activation of words across the bilinguals' languages depends on several factors (such as the type of words to be processed or the fluency of bilinguals in their two languages). It should be mentioned that the strengthened of L2 words-to-concepts connections does not necessarily implies skipping over the L1-L2 lexical links between languages. Desirable difficulties, which are initial disadvantages in new words learning that are later translated into benefits for learning and memory (Bjork, Dunlosky, & Kornell, 2013), are present in fluent bilinguals. These problems are caused by the self regulation, that imposes initial costs but foster the learning process overtime (Valian, 2015). In fact, Bogulski, Bice and Kroll (2018), demonstrated that bilinguals advantages in new words learning based on improved regulatory skills in their dominant language. Bilinguals are experts on regulate their L1 to successfully use their L2 depending on the context. Hence, in bilinguals the new acquired vocabulary is founded on the regulatory skills when words are learned via the dominant language. Generally speaking, between-language co-activation has been found across a wide variety of tasks when individuals process out-of-context words (e.g., picture naming, lexical decision, and translation tasks) (Cristoffanini, Kirsner, & Milech, 1986; De Groot & Nas, 1991; Dijkstra, Grainger, & Van Heuven, 1999; Dijkstra, Van Jaarsveld, & ten Brinke, 1998; Dijkstra & Van Heuven, 2002; Lalor & Kirsner, 2001). Moreover, this non-selective activation reflects the use of L1-L2 lexical connections. For instance, words that share lexical overlap across languages (e.g., cognate words such as *tren,* meaning train in Spanish) relative to control words without L1-L2 lexical overlap facilitate performance on single word processing tasks. However, non-selective activation is attenuated when bilinguals process words in a sentence context (see Van Assche, Duyck, & Hartsuiker, 2016; Lauro & Schwartz, 2017, for reviews). For instance, Schwartz and Kroll (2006) considered the processing of cognate words as an index of the use of L2-L1 lexical connections. The authors observed cross-language facilitation when Spanish (L1) – English (L2) bilinguals processed cognate words (e.g., the word *piano*, meaning the same in Spanish and English) in low semantic constraint sentences (e.g., *when we entered the dining hall we saw the piano*…) but not in high semantic constraint sentences (e.g., *before playing, the composer first wiped the key of the piano*…). Thus, these studies suggest that the use of L2-L1 connections is affected by the semantic constraint of sentences because an effect associated to between-language lexical connections (cognate facilitation) is observed in low predictability sentences but not in high predictability sentences (although see, Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011; Libben & Titone, 2009; Pivneva, Mercier, & Titone, 2014; Titone, Libben, Mercier, Whitford, & Pivneva, 2011, for cognate facilitation in both low- and high-constraint sentences). Finally, it has been shown that even in contexts without lexical overlap between languages (non-cognate words) bilinguals are able to effectively anticipate upcoming words (Foucart, Martin, Moreno, & Costa, 2014).

ERPs associated to L2 learning

It has been shown that in early stages of L2 acquisition, although behavioral measures might not reveal differences associates to the learning process, ERPs show patterns of activity compatible with L2 sensibility (McLaughlin, Osterhout, & Kim, 2004; Tokowicz & MacWhinney, 2005). In this context, an electroencephalography (EEG) study could shed light on this issue. Electrical activity at the surface of the scalp can be experimentally measured and time locked to events of interest. The activity is averaged across similar trials and the resulting waveform can inform about underlying neural processes. Nowadays, after a vast number of studies implementing ERPs, different components have been identified and associated to specific brain processes when participants perform experimental tasks. A growing number of studies use ERPs measures to study bilingual language processing (Elgort, Perfetti, Rickles, & Stafura, 2015; Hahne, 2001; Ojima, Nakata, & Kakigi, 2005; Wu & Thierry, 2017). In our study, we pay special attention to different brain components that have previously demonstrated to be directly related to lexico-semantic processing in novel and experts bilinguals. Two electrophysiological components have been used to index vocabulary acquisition in L2, the N400 component and the late positivity component (LPC) (Midgley, Holcomb, & Grainger, 2009; Yum, Midgley, Holcomb and Grainger, 2014).

The N400 component was observed for the first time by Kutas and Hillyard (1980) in a sentence reading task in which semantically incorrect words elicited a larger negativity peaking around 400 ms post-stimulus onset at central-parietal areas. The N400 component is a negative-going waveform peaking at approximately 350-450 ms after stimulus onset, whose amplitude is sensitive to the processing of lexico-semantic information (Kutas & Hillyard, 1980). In bilingual studies, it has been shown that the processing of L2 words elicit smaller N400 than the processing of L1 words.

Moreover, L2 words also produce larger N400 in high proficiency L2 bilinguals compared with L2 learners with low proficiency (Midgley et al., 2009). Thus, modulations of N400 amplitude when individuals process L2 words have been considered in previous studies as an index of L2 proficiency.

Many behavioral studies have found that word retrieval is easier in backward translation than in forward translation due to the difficulty associated with semantic processing in L1-L2 translation (Kroll & Stewart, 1994). In electrophysiological terms, an easy retrieval of lexical information would be associated with an attenuation of the N400 component. For instance, word frequency is one of the main indicators of difficulty in lexical access (e.g., Hudson & Bergman, 1985; Monsell, Doyle, & Haggard, 1989), and this lexical factor produce a N400 attenuation (Rugg, 1990, Van Petten & Kutas, 1990) with reduced brain-wave negativity during the processing of high vs. low frequency words. However, electrophysiological studies with the translation task are very limited and the results are mixed. Christoffels, Ganushchak, and Koester (2013) observed a greater amplitude of the N400 component when Dutch (L1) - English (L2) bilinguals translated words in forward vs. backward direction. In contrast, Jost, Radman, Buetler, and Annoni (2018) did not find differences related to the translation direction in the N400 time-window.

The LPC component is a late-onset sustained posterior positivity (about 500 ms) that has been related to reanalysis of information and response monitoring in the field of language processing (Kolk & Chwilla, 2007). In translation tasks, the LPC shows an inverse polarity between parietal and frontal regions and this component has been linked to the reprocessing of information between the input and the output language. For example, in a translation recognition task, a greater LPC amplitude is observed in later regions when word pairs are not translations but are semantically related compared to unrelated word pairs (e.g., Guo, Misra, Tam, & Kroll, 2012). Other authors have related the LPC observed in translation tasks to the process of establishing connections between words across languages (Jackson, Swainson, Cunnington, & Jackson, 2001).

On the other hand, the LPC has been observed when bilinguals perform picture naming tasks (Martin et al., 2013). For instance, the LPC amplitude is greater when the difficulty in the retrieval of the picture names increases (e.g., naming tasks that involve language switching across trials). In general, studies of lexical processing in bilinguals seem to show that the more complex the processing of the stimuli (e.g., L2 naming vs. L1 naming) the greater the mean amplitude of the LPC component (Jackson et al., 2001; Khateb et al., 2007; Kieffaber, Kruschke, Cho, Walker, & Hetrick, 2013; Moreno, Federmeier, & Kutas, 2002).

CHAPTER III

Organization, objectives and hypothesis

Since long time ago, researchers are looking for methods able to facilitate and enhance the L2 acquisition process in children and adults. In their classes, teachers usually implement conservative teaching methods. Teachers have a beautiful and important job; they have in their hands the responsibility of preparing people to real life. Hence, they might have a look to teaching strategies based on scientific research to implement efficient learning methods.

In this doctoral dissertation we present two lines of research in which we evaluate different strategies for the acquisition of vocabulary in a L2. In this section, we present the organization of the experimental series, the hypotheses and the specific goals of each line of research. The objectives associated to both research lines and to specific experimental studies are highlighted in bold font.

The title of this thesis summarizes the main goal of our experimental work. **The main objective of our research is to study strategies to facilitate the acquisition of L2 vocabulary in adult people who learn a foreign language in L1 speaking contexts.**

To accomplish this objective, within the field of bilingualism research, two research lines were developed. Specifically, the focus of attention was established on the role of gestures and lexical-semantic methodologies in L2 vocabulary learning. **As general goal for the gestures line, we wanted to evaluate the effect of performing gestures while learning L2 words. In the case of the lexical-semantic methods, the principal aim was to compare the learning results of lexically based and semantically based instructions.**

Role of gestures line

This research line consisted of three experimental studies. In all of them, we manipulated the relationship between the meaning of gestures and words in order to observe the impact that different types of gestures have in L2 vocabulary acquisition. To this end, we compared four experimental conditions on the learning of different types of words (Experiments 1 to 3). In addition, we compared the learning results of concrete nouns and action verbs after training with our gesture-word association method (Experiments 1 and 2). In teaching research, the possible advantages of learning through actions have been studied for a long time. In Experiment 3, we propose a direct comparison of the effects of self-performed and observed movements while learning vocabulary in L2.

Experiments 1 and 2. Learning Nouns and Verbs in a Foreign Language: The Role of Gestures

Regarding the role of gestures in L2 vocabulary learning, two experiments were conducted. Until now, experiments addressing this topic have used verbs as learning material. Verbs are intrinsically associated with movements and have a clear motor component in their semantic. However, during development, nouns are the first words that children acquire. This may be the reason why nouns are easier to learn than verbs. In our experimental work, we compare the results of learning nouns and verbs coupled with gestures using the same acquisition procedure. In this way, Experiment 1 explores the effect of gestures on L2 nouns learning and Experiment 2 addresses the same issue by using verbs.

1. Previous research seems to indicate that gestures facilitate L2 vocabulary learning. The data suggest that gestures help to create a motor-trace associated with the meaning of words that involves procedural memory, which fosters L2 learning (Macedonia & Mueller, 2016). Nevertheless, we cannot discard other possible explanations, and different accounts might work together to explain the role of gestures in L2 vocabulary acquisition. Moreover, it might be possible that gestures could also have a negative impact on the learning process since individuals learning vocabulary with gestures would be

involved in a dual task situation (Paivio, 1971). **Hence, the first objective was to examine if the use of gestures while learning enhance or hinder L2 vocabulary acquisition.**

2. **The second goal was to elucidate which one of the three theoretical accounts regarding the role of gestures in L2 vocabulary learning, the self-involvement account (Helstrup, 1987), the motor-trace account (Engelkamp & Zimmer, 1984, 1985) or the motor-imagery account (Denis et al., 1991) could better explain the learning results observed in the study.** To this end, participants learned words in four different conditions (no gesture, congruent, incongruent and meaningless condition). The use of these conditions made possible the establishment of specific predictions. If gestures only promote the self-involvement of the participant in the learning tasks, all conditions with gestures would favor L2 vocabulary learning relative to the condition without gestures. If motor-trace of gestures helps participants to acquire new words, familiar gestures would be associated with better L2 vocabulary learning (congruent condition and incongruent condition) relative to the learning of L2 words accompanied by less familiar gestures (meaningless condition). Moreover, the motor-imagery explanation would suggest that the learning of gestures with meaning would produce a facilitation effect or an interference effect depending on the convergence between the meaning of the gesture and the meaning of the word to be learned in L2. To be more specific, congruent gestures could facilitate vocabulary learning while incongruent gestures might impede the acquisition of new words. Alternatively, incongruent gestures might become distinctive and benefit the encoding and recall of L2 words (Worthen, 2006). Furthermore, the positions described above are not mutually exclusive. For instance, the acquisition of L2 words coupled with congruent gestures might involve a balance between the positive effect of promoting the connection between semantic information and L2 words, and the negative impact of engaging the participants in a dual task situation.
3. In two experiments, we compared the learning of L2 nouns (Experiment 1) and verbs (Experiment 2). These word classes were used because it has been

suggested that verbs are more difficult to learn than nouns (Childers & Tomasello, 2002; Gentner, 1981; Gentner & Boroditsky, 2001). In particular, it has been proposed that compared with nouns, action verbs intrinsically contain a gestural/motor component in their representation (Boulenger, Hauk, & Pulvermüller, 2009; Hauk, Johnsrude, & Pulvermüller, 2004). Hence, differences between these two types of words might depend on whether they engage overt body movements (e.g., action verbs). **Thus, our third goal was to determine whether the impact of gestures was greater in the learning of L2 verbs than L2 nouns.**

Experiments 3. Seeing or Acting? The Effect of Performing Gestures on Foreign Language Vocabulary Learning

1. Based on previous studies about the role of gestures in L2 vocabulary learning, we expected to find a positive effect from the use of congruent gestures during the acquisition of L2 words. At the same time, the processing of gestures not related to the meaning of words would impair L2 learning. This pattern of results would confirm both the benefit and the cost of using gestures during vocabulary acquisition in a FL (Feyereisen, 2006; Kelly et al., 2009; Macedonia et al., 2011). **Our objective was to corroborate (replicate) the facilitation and interference effects observed in previous studies regarding the role of congruent and incongruent gestures, respectively.**
2. In general, there is agreement about the facilitative role that gestures have in L2 vocabulary learning (Feyereisen, 2006; Kelly et al., 2009; Macedonia & Klimesch, 2014). However, there is controversy regarding the role that the performance of self-generated movement ("do" learning) versus the mere observation of movement ("see" learning) has in L2 vocabulary acquisition. In our study, we directly evaluated this point. Although movement performance appears to improve learning, it is not clear whether a learning protocol involving self-generated actions would have an additional benefit to the mere observation of movements. If mere observation of gestures were sufficient to modulate vocabulary acquisition, the pattern of outcomes would not depend on the type of training. On the contrary, if active training involving the

performance of gestures maximizes learning, the learning rate would be higher in the "do" training group compared to the "see" learning group. **We wanted to evaluate if self-performed gestures enhance the L2 vocabulary learning process compared to a gestures observation condition.**

Lexical-Semantic methods line

In this research line, three experiments were planned to evaluate the role of lexical and semantic learning methodologies in L2 acquisition. First of all, we wanted to show the effects of both training methods on the reinforcement of connections between L2 words and the semantic system. This is the pattern of processing present in fluent bilinguals and we wanted to see to what extent semantic and lexical learning procedures were able to foster it in novel L2 learners. This issue is addressed in Experiment 4. In Experiment 5, we wanted to go one step ahead. In real life, we do not produce only words in isolation. Instead, we arrange words from different lexical categories within sentences to convey complex messages. So, the consequences of both learning strategies on sentence processing were analyzed. Finally, there are changes at brain level that occur at the beginning of the learning process that might not be captured by behavioral measures. In order to compare the processing differences of participants submitted to lexical and semantic methodologies, we registered electrophysiological measures corresponding to different task performance after learning with lexical or semantic instructions. This topic is presented in Experiment 6.

Experiment 4. Semantic and Lexical Training to Learn Foreign-Language Words in Adulthood

1. **In the this study, our first goal was to explore whether a semantically based L2 learning relative to a lexically based instruction would foster the creation of connections between concepts and L2 words.** Previous studies have shown that in learners of a FL, the type of L2 training modulates the establishment of connections between L1 words, L2 words, and the semantic system (Comesaña et al., 2009). On the one hand, training protocols based on lexical processing (e.g., word association method) favour lexical connections between L1 and L2

words. On the other hand, training protocols based on semantic processing (e.g., picture association method) promote the connections between the words in L2 and the semantic system. This pattern of processing is present in fluent bilinguals. If semantic learners really strengthen the semantic connections between concepts and L2 words, a more robust semantic interference effect is expected in participants submitted to this learning methodology.

2. It is not clear if adult learners are able to exploit the connections between L2 and the semantic system when they acquire vocabulary in a L2. This pattern of processing has been observed in children; however, there is evidence that adults mainly use lexical connections between L1 and L2 when they learn L2. **For this reason, the second objective of this study is to see if adult learners are able to reinforce the links between L2 and the semantic system as children when a semantic learning is followed during L2 instruction.** If adult L2 learners can use pictures to exploit semantic processing of L2 words, they would show interference effects in a translation recognition task when they were subjected to the semantic training protocol relative to the lexical training protocol. In other words, the picture-based method would foster semantic-L2 connections, as confirmed in children (Comesaña et al., 2009; Poarch et al., 2014). In contrast, if adult individuals primarily use lexical connections to acquire L2 words (e.g., Chen & Leung, 1989), no semantic interference effects would be found irrespective of the method used to acquire new words.

Experiment 5. The Way in which Foreign Words are Learned Determines Their Use in Isolation and within Sentences

1. **The aim of this study was to examine to what extend semantic vs. lexical learning strategies promote the use of concepts-to-L2 words links or L1-L2 between languages connections, respectively, when participants process L2 words in isolation and in linguistically rich contexts (sentences).** As far as we know, studies comparing semantic vs. lexical learning methods have only been conducted using out-of-context tasks (Comesaña et al., 2009, 2012; Poarch et al., 2014). Following the same training methodologies presented in experiment

4, participants were submitted to semantic or lexical based instructions. Predictions about the processing of L2 words in sentences were grounded in studies on the modulatory role of the context in the use of L1-L2 lexical connections (e.g., Foucart et al., 2014; Schwartz & Kroll, 2006). According to these studies, fluent bilinguals are able to make use of semantic context to anticipate upcoming words within sentences (e.g., Foucart et al., 2014). In our study, we hypothesized that L2 learners undergoing semantic training would more easily process L2 words in highly predictable sentences than in sentences of low predictability because highly predictable sentences would maximize the use of connections already practiced by these learners during training, that is, the connections between semantic information and L2 words. On the contrary, the predictability effect would be attenuated in the case of L2 learners undergoing lexical training because they would preferably use L1-L2 connections already practiced during the training phase.

2. **The second goal was to evaluate the possible effect that the interval between training and testing can produce in the learning process** (e.g., a time delay between training sessions can produce both costs and benefits in the acquisition process, Rickard, 2007). To this end, participants trained during two sessions that were carried out in two consecutive days with the same set of words. Although the results were evaluated at the end of the second learning session, we collected response time and accuracy measures while participants were performing the training task.

Experiment 6. Lexical and Semantic Training to Acquire Words in a Foreign Language: An Electrophysiological Study

Previous studies indicate that training methods plays a crucial role in L2 acquisition. While semantic strategies as the picture-word method reinforce the conceptual connections between L2 words and the SS, lexical paradigms as the word-word association method strengthen the lexical connections between languages. In addition, previous works posit that sometimes behavioral data are not sensitive enough to capture initial differences that take place at the beginning of the acquisition process (McLaughlin et al., 2004). In this context, ERPs measures are usually implemented to fill this gap. Thus, in this study, we focused on electrophysiological signatures of L2 vocabulary acquisition, so brain-waves were registered during the evaluation tasks. **Hence, our principal objective in this study was to compare ERPs data in learners of a new language submitted to lexical or semantic training procedures.** We hypothesized that if semantic learning strategies are able to foster a pattern of L2 processing similar to that found in fluent bilinguals, the group of participants who learned with the picture-word association method would show a N400 attenuation compared to participants who learned with the word-word association method (Midgley et al., 2009; Monsell et al., 1989). In addition, we expect a less extended LPC in semantic learners than in lexical learners because less reanalysis mechanism would be required if better learning results are observed in this group of training (Kieffaber et al., 2013; Moreno et al., 2002). These signals could be interpreted as an electrophysiological correlate of the deeper semantic processing of participants in the semantic group of training.

In the following section, we will develop the empirical work conducted in the doctoral thesis. All the experimental series have been published or are in the process of publication. In our thesis, we adapt these works to have a homogeneous format throughout the experimental series. However, we have kept the original content of the articles submitted for publication. For this reason, the introductory section of each chapter of the experimental section is similar or closely resembles the introduction section described in previous pages of this doctoral thesis.

-PART 2-

Experimental Section

CHAPTER IV

The Role of Gestures in L2 Vocabulary Learning

Experiments 1 and 2. Learning Nouns and Verbs in a Foreign Language: The Role of Gestures

The impact of gestures on L2 vocabulary learning was evaluated with nouns (Experiment 1) and verbs (Experiment 2). Four training methods were compared: the learning of L2 words with congruent gestures, incongruent gestures, meaningless gestures, and no gestures. Better vocabulary learning was found in both experiments when participants learned L2 words with congruent gestures relative to the no gesture condition. This result indicates that gestures have a positive effect on L2 learning when there is a match between the word meaning and the gesture. However, the recall of words in the incongruent and meaningless gesture condition was lower than that of the no gesture condition. This suggests that gestures might have a negative impact on L2 learning. The facilitation and interference effects we found with the use of gestures in L2 vocabulary acquisition are discussed.

This study has been published as García-Gámez, A. B., & Macizo, P. (2019). Learning nouns and verbs in a foreign language: The role of gestures. *Applied Psycholinguistics, 40*(2), 473-507. doi: 10.1017/S0142716418000656

INTRODUCTION

A critical issue in L2 vocabulary learning research is to find training paradigms to promote faster and efficient vocabulary acquisition. If we consider what fluent bilinguals might do to speak in L2, they would need to retrieve the concepts and the words in that language. The retrieval of the translations in the L1 would be an unnecessary step that might add noise to the communication process. However, learners in the early stages of L2 acquisition seem to activate L1 words when they are processing L2 words. The RHM (Kroll & Stewart, 1994) accounts for the developmental changes occurring during the early stages of L2 acquisition. In this model, L2 words are linked to L1 words and concepts. In the early stages of L2 processing, L1 translation equivalents mediate access to meaning. In contrast, in later stages of L2 development, direct connections between L2 words and concepts become possible. However, empirical evidence demonstrates that L2 word-to-concept links might be available for L2 learners at an earlier point in their L2 acquisition than previously assumed (e.g., Sunderman & Kroll, 2006). It is therefore of interest to search for learning protocols that favor the early establishment of connections between concepts and L2 words.

Previous research has confirmed that the acquisition of connections between L2 words and concepts is fostered by the use of training protocols that involve semantic processing (Barcroft, 2002; Comesaña et al., 2012; De Groot & Poot, 1997; Finkbeiner & Nicol, 2003; Kroll et al., 1998; Poarch et al., 2014; Wimer & Lambert, 1959). To illustrate, the presentation of L2 words with pictures denoting their meanings (picture association method, Finkbeiner & Nicol, 2003) favors the learning process relative to the learning of L2 words presented with their translations in the L1 (word association method, Altarriba & Mathis, 1997; Lotto & De Groot, 1998; Van Hell & Candia-Mahn, 1997). Similarly, imagining the meaning of words to be learned in a FL enhances the acquisition process (Ellis & Beaton, 1993; Wang & Thomas, 1995). In fact, Dual Coding Theory (Paivio, 1971) suggests that the formation of mental images during the learning process might favor the acquisition of new words. According to this theory, verbal information and visual images are integrated, and this increases the probability of remembering new words compared to the use of verbal glosses alone. In this scenario, the use of gestures would be a good way to enhance vocabulary acquisition because such gestures promote the formation of a mental image of the word meaning (Goldin-Meadow & Alibali, 2013; Hostetter & Alibali, 2008; see McCafferty & Stam, 2008, for a review). In fact, it has been proposed that gestures and speech interact and they produce an integrated representation of the meaning of the word (Kelly et al., 2010; McNeill, 1992, 2005; McNeill et al., 2015).

In all spoken languages, speakers accompany their speech with visual-manual communication (Özyürek, 2014). This type of multimodal interaction, called co-speech gestures, involves spoken language, facial expressions, body movements and, especially, hands movements. All these visual and auditory aspects act as an integrated stream of information which improves the communication process (Holle & Gunter, 2007). In fact, Marstaller and Burianová (2014) showed that the right auditory cortex and the left posterior superior temporal brain areas seem to reflect the multisensory integration of spoken language sounds and gestures.

Following the gesture taxonomy proposed by McNeill (1992), representational gestures include iconic gestures used to illustrate what is being said by using the hands to refer to concrete entities and/or actions, and metaphorical gestures, which convey an abstract idea by expressing concrete attributes that can be related to it. This taxonomy considers two additional types of gestures, deictic gestures, which consist of one or more fingers directed to a reference, and beat gestures, which are hands movements that reflect the prosody and emphasize the speech. Iconic gestures can be distinguished from emblematic gestures, which are culturally specific and involve body movements that deliver a message like a word such as “good” (thumb up, hand in fist).

Previous studies have demonstrated the role of iconic gestures in language comprehension (e.g., speech comprehension and gestures, Straube et al., 2012; see Yang et al., 2015, for a review), and production (Goldin-Meadow & Alibali, 2013, for a review of gestures in speaking). It has been observed that performing gestures during the learning process facilitates the acquisition and recall of words in a FL (Macedonia et al., 2011; Quinn-Allen, 1995; So et al., 2012; Tellier, 2008; see Macedonia, 2014; and Macedonia & Kriegstein, 2012, for reviews). For instance, Tellier (2008) evaluated the impact of gestures on the learning of English words in French children. A group of children were presented with words and pictures denoting their meanings, whilst another group of children received the words with gestures illustrating the meaning (e.g., the gesture representing the word 'book' was made by opening and closing hands, palms facing up). The results showed better recall of words in the gesture condition relative to the picture condition. Therefore, a learning process based on the use of gestures might be even more efficient than the picture association method described above (Finkbeiner & Nicol, 2003). However, before attempting an in-depth analysis of the role of gestures in L2 learning, we offer a brief summary of the theoretical frameworks concerning the relationship between gestures and speech. After this, we present a section on the role of gestures in the learning of L2 describing three theoretical perspectives that guided the predictions of our study.

Theoretical frameworks regarding the relationship between gestures and speech

There are several frameworks to explain the connections between gestures and speech. All these frameworks address the nature of representations underlying the processing of gestures. One way to differentiate between models is to consider the relevance of visuospatial and linguistic information. Some views suggest that the representation of gestures is based on visuospatial images (e.g., the sketch model, de Ruiter 2000; the interface model, Kita & Özyürek, 2003; the gestures-as-simulated-action (GSA) framework, Hostetter & Alibali, 2008). In other models, the emphasis is placed on the close relationship between the representation of gestures and linguistic information (e.g., the interface model, Kita & Özyürek, 2003; the growth point theory, McNeill, 1992, 2005). Another way to differentiate between models is by attending to the way gestures and speech are processed (separate vs. unitary processing). In some models, it is considered that gestures and speech are processed in two separate systems (e.g., lexical gesture process model, Krauss et al., 2000) that interact when communicative intentions are formed (sketch model, de Ruiter, 2000), or at the conceptualization stage (interface model, Kita & Özyürek, 2003) to produce effective communication. In other models, however, it is assumed that gestures and speech work together as two parts of the same system (the growth point theory, McNeill, 1992, 2005; the GSA, Hostetter & Alibali, 2008). The gesture-in-learning-and-development (GLD) model (Goldin-Meadow, 2000, 2003), for example, considers that children process gestures and speech independently and they become part of an integrated system in mature speakers (Butcher & Goldin-Meadow, 2000; Özçaliskan et al., 2014).

In addition to the specific mechanism by which gestures and speech are processed, a relevant question concerns the role of gestures in communication. It has been confirmed that listeners glean information from gestures (Alibali et al., 1997; Cassell et al., 1999; Goldin-Meadow, 2003; Goldin-Meadow et al., 1992; Holler et al., 2009; Singer & Goldin-Meadow, 2005; see Hostetter & Alibali, 2008 for a review; Hostetter, 2011, for a meta-analysis demonstrating the benefits of gestures for communication). In fact, since gestures usually arise during speech planning, many models defend the communicative value of gestures (e.g., the sketch model, de Ruiter, 2000; the interface model, Kita & Özyürek, 2003; the growth point theory, McNeill, 1992, 2005; and the GLD framework, Goldin-Meadow, 2000, 2003). In the next section, we address the role of gestures in L2 learning.

Gestures and L2 learning

The role of different types of gestures in L2 learning has been emphasized in several studies (e.g., De Grauwe et al., 2014; Gullberg, 2014; Kelly et al., 2010; Macedonia & Knösche, 2011; Macedonia & Kriegstein, 2012; Morett, 2014; So et al., 2012; McCafferty & Stam, 2008, for reviews). In general, it is widely assumed that gestures have a positive effect on vocabulary acquisition and they should be used in FL instruction and embedded in a natural approach of language teaching (Asher & Price 1967; Carels, 1981; Krashen & Terrell, 1983; Macedonia et al., 2014; however, see Hirata et al., 2014; Kelly et al., 2014, for evidence about the limited effect of gestures on learning segmental phonology).

In cognitive psychology, three main perspectives might account for the facilitative role of iconic gestures in L2 vocabulary learning. The *self-involvement explanation* states that gestures might favor the involvement of the participant in the learning task and therefore, gestures would have a key role in the L2 vocabulary learning process (Helstrup, 1987) and they could facilitate enhanced attention to learning material. In particular, the impact of gestures on vocabulary learning is caused primarily by increased perceptual and attentional processes occurring during proprioception of movements associated with gestures or when individuals use objects to perform the action (Bäckman et al., 1986). However, under this view, the benefits associated with the use of gestures does not come from enactment itself because the motor component is not crucial (Kormi-Nouri & Nilsson, 2001); rather, it is the multisensory information conveyed about a word that leads to greater semantic processing and higher attention level (Knopf, 1992; Knudsen, 2007). Therefore, according to the self-involvement explanation, the learning of new words with gestures facilitates vocabulary acquisition regardless of whether a gesture is usually produced within a language or it denotes the same meaning of the word to be learned. Attention increases the retention of words (Craik & Tulving, 1975; see Muzzio et al., 2009, for the role of attention in the encoding and retrieval of information at the neurobiological level). Thus, individuals using gestures would learn L2 words in an attentive manner, which would favor the learning process.

The *motor-trace perspective* suggests that the physical component of gestures is coded in the learning process, leaving a motor trace in memory that aids the acquisition of new words in L2 (Engelkamp & Zimmer, 1984, 1985). According to this view, the physical enactment is crucial because it allows the formation of a motor trace associated with the meaning of the word. There is recent evidence from neuroscientific studies (e.g., repetitive transcranial magnetic stimulation) supporting the role of the motor cortices in the understanding of written words (Vukovic et al., 2017). Moreover, there is evidence that the use of familiar gestures might engage procedural memory since they involve well-defined motor programs (Macedonia & Mueller, 2016). Thus, the involvement of procedural memory along with declarative memory used for word processing might enhance vocabulary learning. From this perspective, familiar gestures that have been routinely practiced and are within the repertoire of the speaker (e.g., the gesture of answering the phone) have strong motor traces so that they produce more facilitation than less familiar gestures (e.g., the gesture of moving the finger from the mouth to the ear). Thus, the facilitation effect would depend on the extent to which an individual is familiar with certain gestures. However, according to this view, the effect of gestures would operate independently of their meaning; that is, regardless of the congruency between the gesture and the word meaning, familiar (well practiced) gestures would facilitate the learning process relative to less familiar gestures.

Finally, the *motor-imagery perspective* indicates that gestures are associated with motor images that form part of a word's meaning (Denis et al., 1991). To be more specific, performing a gesture when individuals process a word promotes the creation of a visual image associated with the meaning of this word, which would enrich the semantic content of the word to be learned. This image would be a mental representation of the action associated with the word during encoding (Macedonia & Knösche, 2011; Macedonia et al., 2011). In fact, neurobiological evidence obtained with functional connectivity analyses suggests the involvement of the hippocampal system in binding visual and linguistic representations of words learned with pictures (Takashima et al., 2014). Hence, according to this view, the facilitation effect observed with gestures would be greater for iconic gestures related to the meaning of words to be learned than when there is a mismatch between the gestures and the meaning of the word. In addition, the learning of words accompanied by gestures with incongruent meanings would produce semantic interference and reduced recall (Cook et al., 2012; Yap et al., 2010).

It is important to note that the three perspectives described above are not mutually exclusive but instead emphasize different aspects of the effect of gestures in L2 learning. Hence, a gesture accompanying a word might increase self-involvement (co-speech gestures as movements that enhance attention to the L2 learning), create a motor trace (co-speech gestures as meaningful movements) due to the physical enactment associated with the production of gestures, or a semantic visual image that could become integrated with the meaning of the word (co-speech gestures as semantically congruent movements).

Empirical evidence regarding the impact of gestures on L2 learning

The first empirical study concerned with the impact of gestures on L2 learning was conducted by Quinn-Allen (1995). In this study, English participants had to learn French expressions under two conditions. In the control condition, the participants were presented with sentences in French (e.g., *Veux-tu quelque chose à boire?* Do you want something to drink?), which they had to repeat in French. In the experimental condition, the learners received the sentences with an emblematic gesture illustrating the meaning (e.g., pointing the thumb towards the open mouth) and they were then required to reproduce the gesture. The results showed that sentences presented with gestures were associated with better recall in comparison with control sentences.

The facilitation effect produced by gestures in Quinn-Allen's (1995) study can be accommodated within the self-involvement account (Helstrup, 1987). Participants would be more engaged in the learning process when they received and produced gestures relative to the control condition without gestures. However, the facilitation effect could also be explained by the motor-trace account and the motor-imagery explanation. The gestures used by Quinn-Allen, such as the act of drinking, are conventional gestures frequently produced in social communication, which could explain the facilitation effect according to the motor-trace view. In addition, this gesture is congruent in meaning with the sentence to be learned (do you want something to drink?) and, hence, a facilitation effect would be expected on the basis of the motor-imagery account. In short, a disadvantage of comparing only a condition with gestures with a condition without gestures is that it is not possible to distinguish whether the benefits observed in the gesture condition come from the use of familiar gestures, gestures with related meanings, and/or the mere fact of performing gestures. Additional experimental work has resolved this problem by using paradigms in which several gesture conditions are compared (Kelly et al., 2009; Krönke et al., 2013; Macedonia et al., 2011; Macedonia & Knösche, 2011).

In Macedonia et al.'s (2011) study, a group of German speakers were trained in Vimmi, an artificial language that has the advantage of controlling for several linguistic variables such as word length, familiarity of L2 words and phonotactic factors (see Jalbert et al., 2011, for a discussion of these variables in vocabulary learning). The authors compared the learning of concrete nouns presented with iconic gestures (e.g., the word suitcase appeared with the gesture of an actor lifting an imaginary suitcase) or meaningless gestures (e.g., the word suitcase and the gesture of touching one's own head). The results showed better recall for words learned with iconic gestures relative to words accompanied by meaningless gestures. The findings of this study appear to indicate that gestures involve something additional to the self-involvement of the participant in the task, since participants were exposed to gestures in the two experimental conditions. Both the motor-trace and the motor-imagery accounts could explain the better memory performance in the iconic gesture condition compared with meaningless gestures. Iconic gestures would favor L2 learning because they are semantically rich and because they are produced more frequently than meaningless gestures, hence producing strong motor activation. In fact, the results found in Macedonia et al.'s study favor both explanations. The authors observed activity in the premotor cortices for words encoded with iconic gestures, which is compatible with the motor trace account. Moreover, words learned with meaningless gestures produce activity in a network engaged in cognitive control suggesting that individuals detected the mismatch between the meaning of the word and the gesture.

Other studies have made use of additional experimental conditions to differentiate between explanations based strictly on the motor component of gestures as opposed to those based on the motor-imagery account. Studies performed in monolingual contexts have explored congruity effects during communication by mismatching the information between the semantics of words and the meanings of the gestures (Barbieri et al., 2009; Bernardis & Gentilucci, 2006; Bernardis et al., 2008; Chieffi et al., 2009; Feyereisen, 2006; see Kircher al., 2009 for congruity effects in the context of an unknown language). To illustrate, Kelly et al. (2010; see also, Kelly et al., 2015) conducted an event-related potential study combined with the use of a Stroop-like paradigm. The participants were presented with words (e.g., "cut") and gestures that could be congruent (e.g., the act of cutting) or incongruent (the act of drinking). The authors found reduced N400 to words accompanied by congruent gestures relative to incongruent gestures (a semantic integration effect, Kutas & Hillyard, 1980). In addition, response times were faster in the congruent condition compared with the incongruent condition. The pattern of results found in Kelly et al.'s study is similar to what is observed in other Stroop tasks and seems to indicate that gestures are integrated with the meaning of words, producing interference when the word meaning does not match the meaning of the gesture (incongruent condition). In our view, the results of this study strongly agree with the motor-imagery account of gestures (see also Macedonia et al., 2011). No differences would be expected according to the self-involvement account since both conditions, congruent and incongruent, involved gestures. Moreover, in the congruent and incongruent conditions, familiar gestures were used so they would promote strong activation of motor trace representations and no differences between the two conditions would be found. On the contrary, clear differences between the congruent and incongruent condition would be predicted based on the motor-imagery view since they would differ in the degree to which the gesture could be integrated with the word meaning.

Nevertheless, there is a limitation with using the results of Stroop-like studies to draw the conclusion that gestures integrate with the meaning of words. In the standard Stroop color-word tasks (see MacLeod, 1991, for a review) three conditions are implemented: a congruent condition, an incongruent condition, and a meaningless condition. The usual finding is a facilitation effect when the congruent condition is compared to the meaningless condition and an interference effect when the incongruent condition is compared to the meaningless condition. However, in some of the studies (cited above) regarding the impact of gestures on L2 learning, the meaningless condition is omitted so it is not possible to determine the amount of facilitation and interference that results from the direct comparison of the congruent and incongruent gesture conditions.

Taken together, the results of previous studies appear to confirm the influence of gestures on L2 vocabulary learning. However, in spite of the merits of previous work in the field, a detailed comparison across several conditions is needed to distinguish the various theoretical explanations. In particular, the acquisition of L2 words should be compared between conditions with and without gestures in order to evaluate the self-involvement account. Furthermore, a condition with familiar gestures (e.g., iconic gestures) might be compared to a condition with unfamiliar gestures in order to evaluate whether the motor-trace of gestures modulates vocabulary learning. Finally, congruent and incongruent conditions need to be compared with a meaningless condition to fully determine the motor-imagery account of gestures in L2 vocabulary acquisition. We acknowledge that some of these comparisons have been made in separate studies. For example, previous work has explored the comparison between congruent and incongruent gestures (e.g., Kelly et al., 2010), as well as the use of a gestural control condition (Cook et al., 2012; Wagner et al., 2004). However, if we assume that several explanations might work together to explain how gestures interact with L2 vocabulary acquisition, it would be desirable to evaluate in the same study the role of congruent, incongruent, meaningless, and no-gesture conditions.

Moreover, if we consider the overall pattern of results found in previous research, almost all studies have revealed a positive effect of gestures on L2 vocabulary learning. However, this observation contrasts with previous work showing that dual task conditions, in which individuals have to perform two tasks simultaneously, can hinder language learning (e.g., learning at the same time the meaning and form of aural input, see Van Patten, 1990; Wong, 2001). A L2 learning condition in which individuals have to process words and gestures concurrently is a dual task that might negatively influence the learning process. However, in some previous studies concerning the role of gestures in L2 vocabulary acquisition, this possible negative effect has not been captured. In fact, previous studies show beneficial effects of gesturing during the L1 and L2 language processing (Holle & Gunter, 2007; Marstaller & Burianová, 2014; Özyürek, 2014). For example, Quinn-Allen (1995) observed better retrieval of L2 expressions when these were learned with semantically related gestures relative to learning without gestures. In our opinion, the possible negative effect of a dual task condition (the concurrent processing of a L2 sentence and a gesture) might not be observed in this study since it overlapped with a positive effect of congruency between the sentence and the gesture meaning. Therefore, a direct comparison between a no gesture condition and a meaningless gesture condition is required in order to determine the possible negative effect of gestures in L2 vocabulary learning.

Another important question concerns the type of words used as learning material in previous studies. To assess the effect of gestures on vocabulary learning, most of the studies have employed verbs as training material (De Grauwe et al., 2014; Kelly et al., 2009; Kircher et al., 2009). The use of verbs seems to be the best option if we consider that representational gestures depict actions. There is a more direct mapping between gestures and the semantic characteristics of verbs, in comparison with other types of words such as nouns (Childers & Tomasello, 2002; De Grauwe et al., 2014; Kelly et al., 2009; Kircher et al., 2009).

However, apart from gestures, it has been corroborated that nouns are easier to learn than verbs, at least for children. Different studies have shown that children are able to learn English nouns easier than verbs in their natural context (Fernald & Morikawa, 1993; Goldfield, 1993; Tardif et al., 1997). It is probably due to the fact that

English speakers place special emphasis on nouns when they interact with children while acquiring their L1. However, this advantage is not present in other cultures (see Gentner & Boroditsky, 2001; Gopnik & Choi, 1995; Tardif et al., 1997 for the absence of this advantage in Korean and Mandarin languages). Childers and Tomasello (2002) conducted a study where children learned 6 novel nouns, 6 novel verbs and 6 novel actions. Concerning the role of gestures during the learning of nouns and verbs, the GSA framework makes concrete predictions about what types of words could be more influenced by gestures. This theory states that gestures occur as a result of simulated action and perception, which are the basis of mental imagery and language production (Hostetter & Alibali, 2008). Even if thinking about the size or shape of a particular object (nouns) involves simulating movements, the relationship between verbs and movements is stronger. Gestures would therefore have a greater influence on verb learning than on noun learning. However, to our knowledge, there are no previous studies in which the role of gestures is examined during the learning of nouns and verbs. This comparison was performed in the current study.

The current study

Previous research seems to indicate that gestures facilitate L2 vocabulary learning. The data suggest that gestures help to create a motor-trace associated with the meaning of words that involves procedural memory, which fosters L2 learning (Macedonia & Mueller, 2016). Nevertheless, we cannot discard other possible explanations, and these accounts might work together to explain the role of gestures in L2 vocabulary acquisition. Moreover, it might be possible that gestures could also have a negative effect on the learning process since individuals learning vocabulary with gestures would be involved in a dual task situation. The current study aimed at evaluating these questions.

In our study, monolingual speakers of Spanish (L1) learned words in an artificial language (Vimmi, L2) on three consecutive training days. The words to be learned were presented alone (no gesture condition), coupled with meaningless gestures or unfamiliar gestures (meaningless gesture condition), or they were presented with iconic gestures whose meanings were semantically related (congruent condition) or unrelated (incongruent condition) to the meanings of words.

In two experiments, we compared the learning of L2 nouns (Experiment 1) and verbs (Experiment 2). These word classes were used because it has been suggested that verbs are more difficult to learn than nouns (Childers & Tomasello, 2002; Gentner, 1981; Gentner & Boroditsky, 2001). To be more specific, it has been proposed that compared with nouns, action verbs intrinsically contain a gestural/motor component in their representation (Boulenger et al., 2009; Hauk et al., 2004). Hence, differences between these two types of words might depend on whether they engage overt body movements (e.g., action verbs). For example, De Grauwe et al. (2014) found that the comprehension of motor verbs (e.g., to throw) in a L2 produced activation of motor and somatosensory brain areas. Hence, it might be possible that the effect of using gestures in L2 vocabulary learning would be greater with verbs than with nouns.

If we consider the theoretical accounts concerning the role of gestures in L2 vocabulary learning, the self-involvement account (Helstrup, 1987), the motor-trace account (Engelkamp & Zimmer, 1984, 1985) and the motor-imagery account (Denis et al., 1991), then it is possible to generate specific predictions. If gestures only promote the self-involvement of the participant in the learning tasks, all conditions with gestures would favor L2 vocabulary learning relative to the condition without gestures. If motor-trace of gestures helps participants to acquire new words, familiar gestures would be associated with better L2 vocabulary learning (congruent condition and incongruent condition) relative to the learning of L2 words accompanied by less familiar gestures (meaningless condition). Moreover, the motor-imagery explanation would suggest that the learning of gestures with meaning would produce a facilitation effect or an interference effect depending on the convergence between the meaning of the gesture and the meaning of the word to be learned in L2. More specifically, congruent gestures could facilitate vocabulary learning while incongruent gestures impede the acquisition of new words. Alternatively, incongruent gestures might become distinctive and benefit the encoding and recall of L2 words (Worthen, 2006). Furthermore, the positions described above are not mutually exclusive. For instance, the acquisition of L2 words coupled with congruent gestures might involve a balance between the positive effect of promoting the connection between semantic information and L2 words, and the negative impact of engaging the participants in a dual task situation.

Experiment 1. Learning L2 Nouns

METHOD

Participants

Twenty-five individuals participated in the experiment (21 women and 4 men). Their mean age was 21.72 years (*SD* = 3.17). All of them were native Spanish speakers. Each subject provided written informed consent before performing the experiment.

Design and Materials

Four L2 vocabulary learning conditions were manipulated within-participants: No gesture condition: Spanish (L1) - Vimmi (L2) word pairs had to be learned without gestures (e.g., *cuchara*, spoon in Spanish, and *deschoga*, a Vimmi word). Meaningless condition: L1-L2 word pairs to be learned were coupled with unfamiliar gestures (e.g., *cuchara-deschoga* and the gesture of moving the hand from the forehead to the ear). Congruent condition: L1-L2 word pairs were accompanied with gestures that reflected the common use of objects whose names had to be learned in L2 (e.g., *cuchara-deschoga* and the gesture of holding an invisible spoon and raising it to the mouth). Incongruent condition: L1-L2 word pairs were coupled with gestures associated with the use of an object different from that denoted by the L1 word (e.g., *cuchara-deschoga* and the gesture of lighting a match) (see Figure 3) (see Appendix 1 for the complete set of material used in the study).

The congruent and incongruent gestures presented along with the L1-L2 word pairs were iconic gestures (McNeill, 1992), which have been also called representational gestures (Kendon, 1981) that usually illustrate a concrete physical object by using the hands to show the properties or details of the item. For example, for the meaning of pencil, the gesture would involve holding a pencil with the fingers and doing handwriting movements. The gestures used in the meaningless condition were small movements performed with the hand that did not have iconic or metaphoric associations with the meaning of physical items (for example, to form a fist with one hand and raise the fingers of the other hand). We took care to select meaningless gestures with similar properties to those of meaningful gestures (e.g., hand configuration, the use of simple movement trajectory, and spatial location). For the meaningless condition, we selected ten different movements which were the same for all the participants.

In addition, 40 words were selected in Spanish. These words were concrete nouns denoting objects that could be manipulated with the hands (e.g., spoon, pen, etc.). Forty words were also selected from an artificial language, Vimmi (Macedonia et al., 2011; Macedonia & Knösche, 2011). The corpus of Vimmi words is constructed so that it avoids factors that might favor the learning of specific items (co-occurrence of syllables, similarity with words from languages such as Spanish, English, and French). Vimmi words were carefully selected so that they were pseudowords with legal orthography and phonology in Spanish but without meaning. To create the learning material, the forty Spanish nouns were randomly paired with the forty Vimmi words previously selected. The resulting 40 word pairs (L1-Spanish/L2-Vimmi word pairs) were randomly divided into 4 sets of 10 word pairs. Each set of 10 pairs was associated to one gesture condition (congruent condition, incongruent condition, meaningless condition, and no gesture condition). In order to counterbalance the gestures conditions across word sets, a total of 4 lists of material were created. In this way, a word pair (e.g., “cuchara-deschoga”) was coupled with a congruent gesture in list 1, an incongruent gesture in list 2, a meaningless gesture in list 3, and it was presented without gesture in list 4. Each participant was randomly assigned to one of the four lists. Hence, across lists, the 40 words were counterbalanced over the four training conditions, so that all word pairs appeared in all training conditions.

Figure 3. Conditions used during the study of Spanish (L1) – Vimmi (L2) words (nouns in Experiment 1 and verbs in Experiment 2). In the example, cuchara (spoon in Spanish) – dechoga (a Vimmi word) was accompanied by (a) the gesture of holding an invisible spoon and raising it to the mouth (congruent condition); (b) the gesture of lighting a match (incongruent condition), (c) the meaningless gesture of moving the hand from the forehead to the ear (meaningless condition), (d) the word pairs were presented without a gesture (no gesture condition).

We equated the Spanish words across the four sets of word pairs in lexical variables (Davis & Perea, 2005). There were no differences across word sets in terms of word length (number of graphemes, $F(3, 36) = 1.40$, $p = .26$ ($M = 6.57$, $SD = 1.72$), number of phonemes, $F(3, 36) = 1.06$, $p = .38$ ($M = 6.27$, $SD = 1.63$), number of syllables, $F(3, 36) = 1.58$, $p = .21$ ($M = 2.77$, $SD = 0.77$), lexical frequency, $F(3, 36) = 2.67$, $p = .06$ ($M = 16.07$, $SD = 35.74$, per one million count), familiarity, $F(3, 36) = 1.59$, $p = .21$ ($M = 3.95$, $SD = 2.85$), and concreteness, $F(3, 36) = 1.42$, $p = .25$ ($M = 4.06$, $SD = 2.88$). Similarly, Vimmi words in the four sets were equated in number of graphemes, $F < 1$ ($M = 5.15$, $SD = 1.61$), number of phonemes, $F < 1$ ($M = 5.10$, $SD = 1.53$), and number of syllables, $F < 1$ ($M = 2.47$, $SD = 0.78$). Finally, we controlled for the similarity between the Spanish words and Vimmi words across sets of word pairs. The number of shared phonemes between the Spanish and Vimmi words was the same across the four sets both when the phoneme position was considered, $F(3, 36) = 1.59$, $p = .21$ ($M = 0.36$, $SD = .54$), and when it was not, $F < 1$ ($M = 1.82$, $SD = 0.93$).

The gestures used in all gesture conditions involved hand movements. The congruent and incongruent gesture conditions included iconic gestures that represented frequent movements that individuals usually perform when they manipulate objects (e.g., the gesture of playing the flute, the gesture of typing on a keyboard, etc.). The meaningless gestures involved similar small movements with the hands but they did not convey any meaning (e.g., to move a closed hand from right to left in front of the face). These gestures were carefully selected to ensure that they were not emblems or gestures with metaphorical meaning. In addition, we wanted to make sure that the congruent condition, incongruent condition, and meaningless condition differed in the degree to which the semantic of the word was associated with the paired gesture. To this end, a set of 15 Spanish participants that did not participate in the main experiment took part in a pilot study. The participants received a video with a gesture (without sound) at the top of the screen and a word written in Spanish at the bottom of the screen, and they were instructed to rate the degree to which there was a match between the meaning of the word and the gesture, which ranged from 1 (high mismatch) to 9 (high match). There were differences between the congruent condition (M = 8.23, SD = 2.34), meaningless condition (M = 3.96, SD = 1.80), and incongruent condition (M = 1.55, SD = 1.52), $F(2, 28) = 146.39$, $p < .001$, $\eta^2 = .91$. The gesture-word pairs were rated higher in the congruent condition compared to the meaningless condition, $F(1, 14) = 94.80$, $p < .001$, $\eta^2 = .87$, and the incongruent condition, $F(1, 14) = 170.64$, $p < .001$, $\eta^2 = .92$. The incongruent condition and the meaningless condition also differed, $F(1, 14) = 43.62$, $p < .001$, $\eta^2 = .96$. Therefore, the three conditions with gestures used in the study differed in terms of the association between the meaning of the word and the gesture.

Procedure

L2 vocabulary learning involved three training sessions conducted on three consecutive days. In each session, participants performed, firstly, the L2 training and, afterwards, the assessment of the L2 learning. The two phases were separated by a 15-min break. E-prime experimental software was used for stimulus presentation and data acquisition (Schneider, Eschman, & Zuccolotto, 2002).

L2 training. We employed a stimulus presentation procedure grouped by experimental condition similar to that used in other studies with various gesture conditions (e.g., Macedonia et al., 2011). Participants were presented with a block of 40 Spanish-Vimmi word pairs. These word pairs were grouped (10 word pairs in each group) according to the four learning conditions (no gestures, meaningless gestures, congruent gestures, and incongruent gestures). This block was repeated 12 times. Hence, a participant received 480 trials where the 40 word pairs were presented 12 times. A short break was introduced between learning blocks. The word pairs were randomly presented within each condition. In addition, the order in which the learning conditions were presented within a block was counterbalanced. This procedure blocked by learning condition avoids the cognitive cost associated with the continuous change between situations in which the participants have to perform gestures and the learning condition without gestures. In all experimental conditions, on each trial, the participant received a Spanish-Vimmi (L1-L2) word pair visually presented at the bottom of the screen. These word pairs were presented alone or they were accompanied by a gesture (see Figure 3). Gestures were recorded on video by the experimenter and they were congruent, incongruent, and meaningless, depending on the learning condition. The duration of each recorded gesture was five seconds and the gesture was repeated twice. In all learning conditions, participants received verbal instruction to read aloud each L1-L2 word twice. In the gesture conditions, participants had to produce the gesture presented on the trial each time they said aloud the L1-L2 word pair. Participants started the production of the gesture when they began the production of the L1-L2 word pair so that each gesture and L1-L2 word pair was produced twice. For example, when participants received the word pair cuchara-deschoga along with the congruent gesture (Figure 3, left-hand panel), they had to say aloud this word pair at the time they produced the gesture of holding an invisible spoon and raising it to the mouth. Once the participants had produced the word pair twice, they had to press the space bar to continue to the next trial. The training lasted approximately 1 hour.

L2 learning assessment. Two tests were used to evaluate the acquisition of L2 words: Translation from Spanish into Vimmi (forward translation from L1 to L2) and translation from Vimmi into Spanish (backward translation from L2 to L1). These tasks have been used in previous studies to evaluate L2 vocabulary acquisition (Kroll & De Groot, 2005; Poarch et al., 2014). Other L2 learning tasks could have been used, however, previous studies have found a positive correlation between L2 proficiency and performance on translation tasks and lexical decision tasks (Prior, MacWhinney, & Kroll, 2007).

The order in which the translation tests were presented was randomized across the three training sessions and across participants. In each translation task, the 40 words to be learned were presented and participants were instructed to translate them. On each trial, a word was presented in the middle of the screen until the participant produced its translation. Oral translations were recorded for later analyses of recall accuracy. Reaction times (RTs) from the presentation of the word until the beginning of the oral translation were also registered. The learning assessment lasted approximately 10 minutes.

RESULTS

The main index of L2 vocabulary learning was the percentage of words recalled in the forward and backward translation tasks; however, RTs were also examined. Translation accuracy across the three training sessions was 48.73% (48.80% in the forward translation test and 48.67% in the backward translation test). The RTs associated with correct translations were trimmed following the procedure described by Tabachnick and Fidell (2001) to eliminate univariate outliers. Raw scores were converted to standard scores (*z*-scores). Data points which, after standardization were 3 *SD* outside the normal distribution, were considered outliers. After removing outliers from the distribution, *z*-scores were calculated again. The filter was applied in recursive cycles until no observations were outside 3 *SD*. The percentage of outliers was 5.20%. Next, we report the results found in the L2 evaluation tests (forward translation and backward translation) for RTs and correct recall. In all analyses, we adopted a significance level of $\alpha = 0.05$. Only correct responses were included in the analyses of the RTs. Data points were excluded from the RT analyses if: (a) the participants produced nonverbal sounds that triggered the voice key, (b) the participants stuttered or hesitated in producing the word, (b) the participants produced something different than the word required. Some small errors were allowed and considered correct responses depending on the length of the correct word to be produced: (1) For monosyllabic words, the replacement of a vowel, 2) for disyllabic words, the replacement of a vowel or a consonant but not both, 3) for words with three or more syllables, the inversion of a vowel and a consonant or the replacement of a vowel or a consonant.

Reaction times

The RTs were subject to analysis of variance (ANOVA) with translation test (forward translation, backward translation), training session (first session, second session, third session) and learning condition (no gestures, meaningless gestures, congruent gestures, incongruent gestures) as within-participants factors. The order in which participants received the translation tasks (forward-backward translation order vs. backward-forward translation order) was initially entered in the analyses as a between-subject factor. However, the main effect of order was not significant and this variable did not interact with any other factor (all *p*s > .05); thus, the order of the translation task was not considered any further.

Table 1 shows the mean RTs across conditions. The main effect of session was significant, $F(2, 10) = 5.45$, $p = .02$, $\eta_p^2 = .52$. Mean translation latency was 1380 ms (*SE* = 95) in the first session, 1277 ms (*SE* = 84) in the second session, and 1184 ms (*SE* = 77) in the third session. Linear trend analysis was significant, $F(1, 50) = 7.27$, $p = .04$, $\eta^2 = .52$, thus there was a practice effect with faster recall of words at the end of the training relative to the beginning of the learning process. No other main effects or interactions were significant (all *p*s > .05).

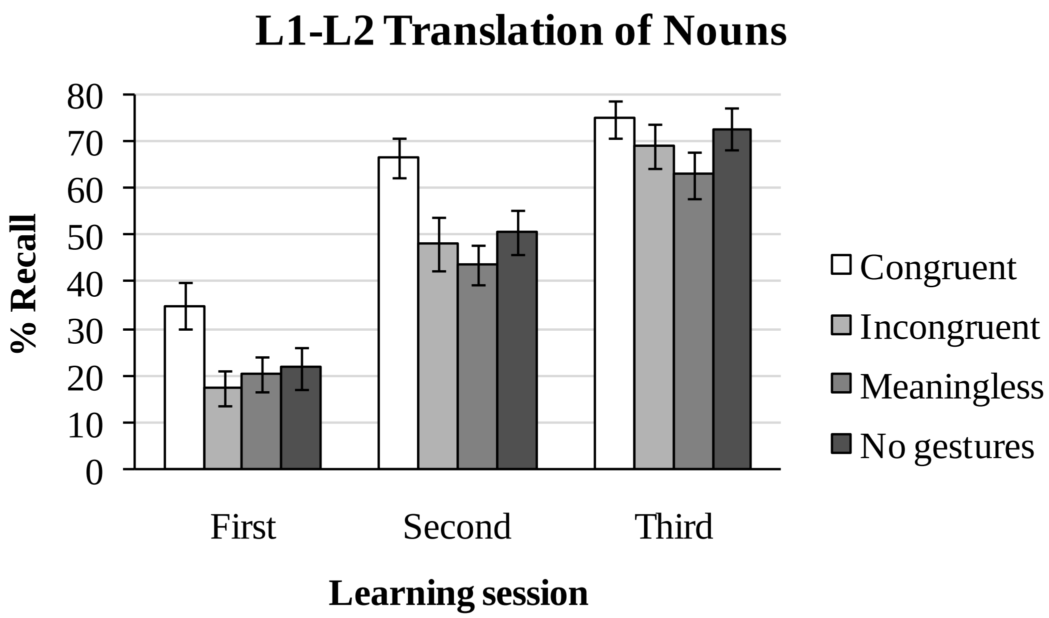
Table 1. *Reaction Times in the Noun Translation Tasks*

	First session	Second session	Third session
	L1 to L2 translation		
Congruent	1392 (57)	1206 (116)	1317 (100)
Incongruent	1289 (93)	1373 (124)	1146 (124)
Meaningless	1549 (165)	1328 (148)	1345 (89)
No gestures	1704 (184)	1370 (135)	1156 (96)
	L2 to L1 translation		
Congruent	1458 (111)	1269 (45)	1121 (45)
Incongruent	1597 (153)	1330 (112)	1232 (67)
Meaningless	1567 (167)	1404 (86)	1329 (78)
No gestures	1390 (130)	1319 (56)	1259 (67)

Note. Mean RTs (in milliseconds) obtained in Experiment 1 (learning L2 nouns) as a function of the translation test (L1 to L2 translation, L2 to L1 translation), the training session (first session, second session, third session) and the learning condition (congruent gestures, incongruent gestures, meaningless gestures, no gestures). Standard errors are in brackets.

Recall performance

The Translation test x Session x Learning condition ANOVA revealed a significant main effect of session, $F(2, 48) = 210.17$, $p < .001$, $\eta_p^2 = .90$. The recall percentage of L2 words was 25% (*SE* = 3.31) in the first session, 52% (*SE* = 4.12) in the second session, and 69% (*SE* = 3.75) in the third session. Linear trend analysis showed that the recall of L2 words was higher on the final session relative to the beginning of training, $F(1, 24) = 263.69$, $p < .001$, $\eta^2 = .92$. The main effect of learning condition was significant, $F(3, 72) = 11.80$, $p < .001$, $\eta_p^2 = .33$; and this effect was modulated by the translation test, so that the Learning Condition x Translation test interaction was significant, $F(3, 72) = 3.99$, $p = .01$, $\eta_p^2 = .14$. No other main effects or interactions were significant (all *p*s > .05). When the Translation test x Learning condition interaction was examined, we found differences due to the translation test in the no gesture condition, $F(1, 24) = 4.74$, $p = .04$, $\eta_p^2 = .16$. In particular, better recall was observed in the backward translation (54%, *SE* = 4.57) relative to the forward translation (49%, *SE* = 4.05). The differences between the two translation directions were not significant in the congruent condition, incongruent condition and meaningless condition (all *ps* > .05). Next, we examined recall performance in each translation test.



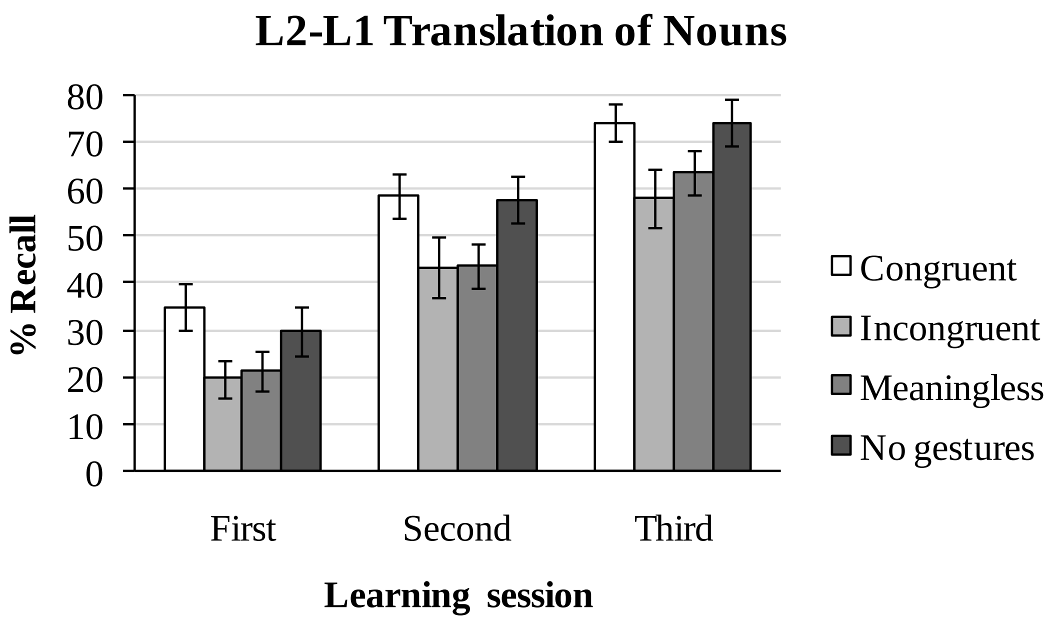


Figure 4. Recall percentages (% Recall) obtained in Experiment 1 during the translation of nouns from L1 to L2 (upper graph) and from L2 to L1 (bottom graph) across training sessions (first, second, third) and gesture conditions (congruent, incongruent, meaningless, and no gestures). Standard error is plotted in vertical lines.

Forward translation (from Spanish-L1 to Vimmi-L2). The results found in the forward translation test are presented in Figure 4. When session and learning condition were entered in the ANOVA, the main effect of session was significant, $F(2, 48) = 217.65$, $p < .001$, $\eta_p^2 = .90$. Recall percentage was 24% ($SE = 3.32$) in the first session, 53% ($SE = 4.02$) in the second session, and 70% ($SE = 3.47$) in the third session. Linear trend analysis was significant, $F(1, 24) = 326.53$, $p < .001$, $\eta^2 = .93$. Thus, L2 vocabulary learning increases as a function of the training. In addition, the main effect of learning condition was significant, $F(3, 72) = 10.09$, $p < .001$, $\eta_p^2 = .30$. The Session x Learning condition interaction was not significant, but there was a trend towards significance, $F(6, 144) = 1.90$, $p = .08$, $\eta_p^2 = .07$. When learning conditions were compared (see Table 2), we observed better recall in the congruent gesture condition (59%, $SE = 3.80$) relative to the meaningless gesture condition (43%, $SE = 3.58$). The recall performance in the incongruent condition (45%, $SE = 4.19$) was similar to the meaningless condition and the no gesture condition. However, compared with the no gesture condition (49%, $SE = 4.05$), the recall percentage was lower in the meaningless condition (see Figure 5).

Table 2. *Comparison between Learning Conditions in the Noun and Verb Translation Tasks*

	Translation of Nouns				Translation of Verbs			
	L1 to L2		L2 to L1		L1 to L2		L2 to L1	
	t(24)	*p*	*t*(24)	*p*	*t*(24)	*p*	*t*(24)	*p*
Congruent vs. No gestures	3.33	.003*	0.64	.528	6.15	.001*	3.05	.005*
Congruent vs. Meaningless	5.02	.001*	3.87	.001*	5.94	.001*	5.54	.001*
Congruent vs. Incongruent	3.77	.001*	3.53	.001*	8.55	.001*	6.71	.001*
Incongruent vs. No gestures	1.10	.282	3.90	.001*	0.37	.716	2.68	.011*
Incongruent vs. Meaningless	0.82	.418	0.78	.442	0.41	.682	1.16	.256
Meaningless vs. No gestures	2.02	.055*	3.68	.001*	0.82	.418	2.09	.050*

Note. Comparison of recall percentages across learning conditions obtained in the translation tasks of Experiment 1 (learning of L2 nouns) and Experiment 2 (learning of L2 verbs). *$p \leq .05$

Backward translation (from Vimmi-L2 to Spanish-L1). The ANOVA conducted with session and learning condition as variables revealed a significant main effect of session, $F(2, 48) = 122.55$, $p < .001$, $\eta_p^2 = .84$. The recall of Spanish words from Vimmi words was 27% (*SE* = 3.90) in the first session, 51% (*SE* = 4.48) in the second session, and 68% (*SE* = 4.26) in the third session. Linear trend analysis indicated that recall performance increased as a function of training, $F(1, 24) = 141.98$, $p < .001$, $\eta^2 = .85$. The main effect of learning condition was also significant, $F(3, 72) = 10.05$, $p < .001$, $\eta_p^2 = .30$. Finally, the Session x Learning condition interaction was not significant, $F < 1$.

When the learning conditions were compared, we observed better recall in the congruent condition (56%, *SE* = 4.07) relative to the meaningless condition (43%, *SE* = 4.10). The difference in recall between the incongruent condition (41%, *SE* = 5.06) and meaningless condition was not significant. However, compared with the no gesture condition (54%, *SE* = 4.57), recall was lower in the incongruent condition and the meaningless condition.

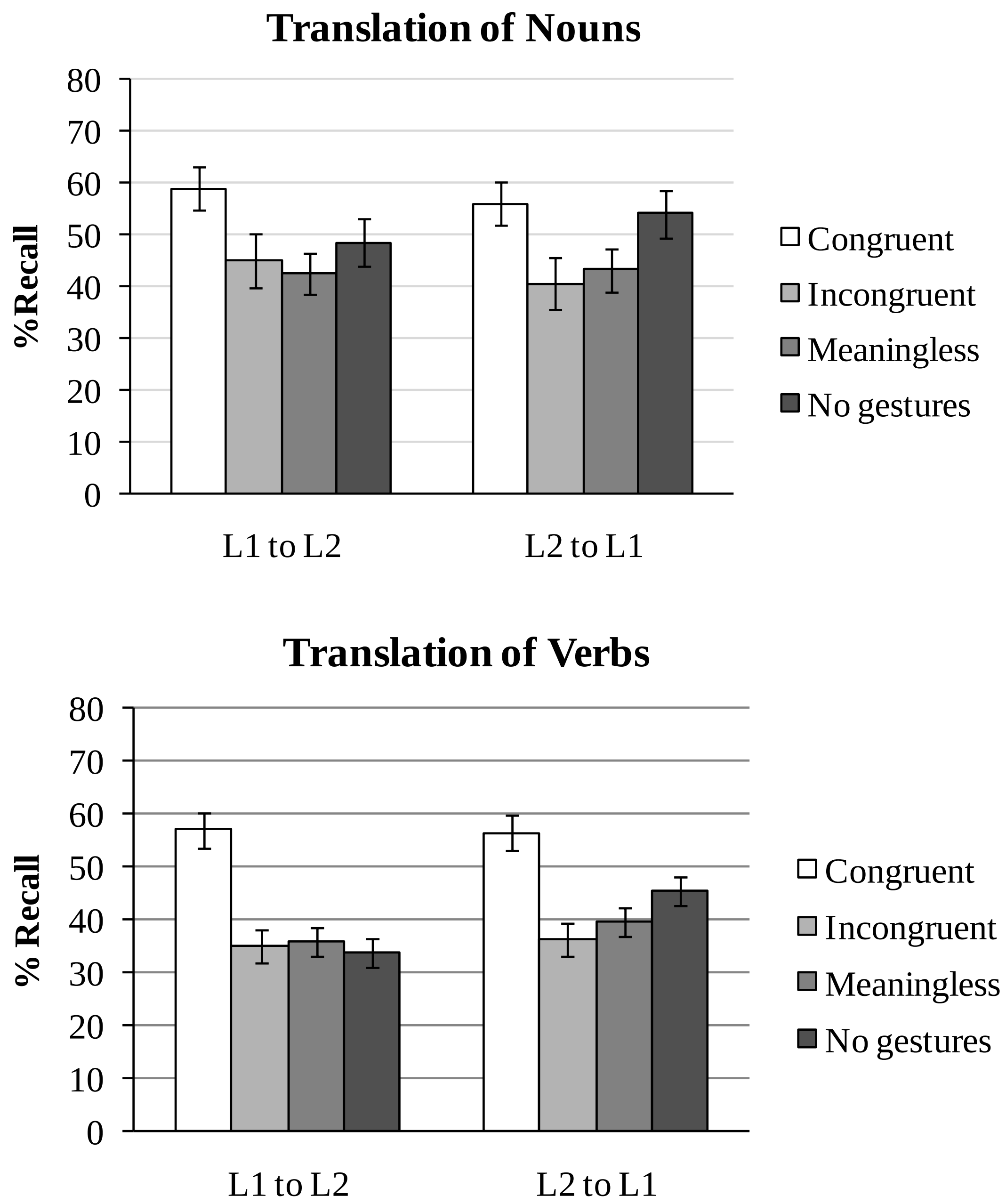


Figure 5. Recall percentage (% Recall) of nouns (Experiment 1, upper graph) and verbs (Experiment 2, bottom graph) as a function of translation direction (L1 to L2, L2 to L1) and gesture conditions (congruent, incongruent, meaningless, and no gestures). Standard error is plotted in vertical lines.

DISCUSSION

Two main effects were found in Experiment 1 when examining the learning of L2 nouns across different training conditions. When participants learned L2 words in the congruent condition, percentage of recall was higher than in the meaningless condition. However, the recall of L2 words was lower in the incongruent and meaningless condition relative to the no gesture condition, as shown in the backward translation task. This finding indicates that these conditions negatively influenced the learning process. Together, two opposing effects were found when several L2 learning methods were compared, i.e. facilitation and interference. The facilitation effect found with congruent gestures seems to support the motor-imagery account of the role of gestures in L2 vocabulary learning (Denis et al., 1991). The shared semantic meaning of gestures and L1 words fostered the acquisition of L2 words. The motor-trace perspective (Engelkamp & Zimmer, 1984, 1985) could also explain the facilitation effect found here, since congruent gestures were also familiar gestures. However, this view would not account for the interference effect found with incongruent gestures relative to the meaningless condition since incongruent gestures were also common gestures. Moreover, the self-involvement explanation could not accommodate the pattern of results found in this experiment since clear differences between conditions involving gestures were observed (Helstrup, 1987). On the other hand, the magnitude of the interference effect was similar in the incongruent condition and the meaningless condition, when comparing both with the learning of nouns in the no gesture condition. This observation seems to indicate that the negative impact of gestures on these two conditions was due to the fact that participants were immersed in a dual task setting, which increased the difficulty of the learning process.

Experiment 2. Learning L2 Verbs

In Experiment 1, we observed that gestures could either benefit or hinder the learning of nouns in a FL. This observation needed replication and further examination, which was the aim of Experiment 2.

When comparing the learning of nouns and verbs, it has been found that verbs are more difficult to acquire than nouns (Childers & Tomasello, 2002; Gentner, 1981;

Gentner & Boroditsky, 2001). A possible way of remediating the intrinsic difficulty associated with the acquisition of verbs would be the use of gestures during the learning process. In fact, it has been theorized that the semantic representation of verbs involves an intrinsic motor component (Boulenger et al., 2009; Hauk et al., 2004). Hence, in Experiment 2 we evaluated the role of gestures when participants learned verbs in L2.

METHOD

Participants

Thirty-two native Spanish speakers from the University of Granada participated in Experiment 2 (6 men and 26 women). The participants had not taken part in Experiment 1. Their mean age was 20.97 years (*SD* = 3.21). Each subject gave written informed consent before performing the experiment. Their participation was rewarded with academic credits.

Design and Materials

The same design as that used in Experiment 1 was employed here, with four learning conditions (no gestures, meaningless gestures, congruent gestures, and incongruent gestures). A new set of 40 Spanish (L1) words was selected. These words were verbs denoting actions that require movements of certain parts of the body (e.g., to eat, to smile, etc. see Appendix 1 for the complete set of material used in the study). In addition, the same 40 Vimmi (L2) words used in Experiment 1 were randomly paired with the L1 verbs to form the L1-L2 word pairs to be learned.

The forty word pairs were randomly sorted into 4 sets (10 word pairs in each) and randomly assigned to one of the four learning conditions. The Spanish verbs across the 4 sets were equated in lexical variables (Davis & Perea, 2005). There were no differences across word sets in the number of graphemes, $F < 1$ ($M = 6.53$, $SD = 1.48$), number of phonemes, $F < 1$ ($M = 6.30$, $SD = 1.44$), number of syllables, $F < 1$ ($M = 2.50$, $SD = 0.68$), lexical frequency, $F < 1$ ($M = 15.48$, $SD = 23.62$, per one million count), familiarity, $F < 1$ ($M = 3.50$, $SD = 3.09$), and concreteness, $F < 1$ ($M = 2.90$, $SD = 2.53$). Finally, we controlled for the similarity between the L1 and L2 words. The Spanish and

Vimmi words across conditions shared the same number of phonemes in the same position, $F < 1$ ($M = 0.37$, $SD = 0.63$), and irrespective of the position within the word, $F < 1$ ($M = 1.45$, $SD = 0.96$).

We took care to control the material used in Experiment 2 (verbs) as was done with the material used in Experiment 1 (nouns). However, verbs intrinsically describe actions so they could have a more direct mapping with representational gestures because they usually depict actions. Thus, the use of gestures may have a greater effect during the acquisition of L2 verbs than for L2 nouns. Nevertheless, as in Experiment 1, we wanted to ensure that the congruent condition, incongruent condition, and meaningless condition differed in the degree to which the meaning of the L1 word was associated with the paired gesture. To this end, a new set of 15 Spanish participants that did not participate in the main experiment took part in a pilot study. The participants received a video with a gesture and a verb written in Spanish and they had to rate the degree to which the gesture and the meaning of the verb matched, ranging from 1 (high mismatch) to 9 (high match). There were differences between the congruent condition ($M = 8.51$, $SD = 1.71$), the meaningless condition ($M = 3.02$, $SD = 2.13$), and the incongruent condition ($M = 1.80$, $SD = 2.02$), $F(2, 28) = 365.09$, $p < .001$, $\eta^2 = .96$. The gesture-word pairs were rated more highly in the congruent condition compared to the meaningless condition, $F(1, 14) = 433.22$, $p < .001$, $\eta^2 = .97$, and the incongruent condition, $F(1, 14) = 1663.89$, $p < .001$, $\eta^2 = .99$. The incongruent condition and the meaningless condition also differed, $F(1, 14) = 13.29$, $p < .001$, $\eta^2 = .48$. Therefore, the three conditions with gestures used in the study differed in terms of the association between the meaning of the verbs and the gestures.

All other details concerning the design and materials were the same as those described in Experiment 1.

Procedure

The procedure was identical to that used in Experiment 1, except that verbs were used instead of nouns.

RESULTS

The recall percentage across the three training sessions was 42.46% (40.44% in the forward translation test, and 44.48% in the backward translation test). The RTs associated with correct translations were trimmed as described in Experiment 1. The percentage of outliers was 7.86%.

Reaction time

An ANOVA was performed with translation test (forward translation, backward translation), training session (first session, second session, third session) and learning condition (no gestures, meaningless gestures, congruent gestures, and incongruent gestures) as within-participant factors. Table 3 shows mean RTs across conditions.

Table 3. *Reaction Times in the Verb Translation Tasks*

	First session	Second session	Third session
	L1 to L2 translation		
Congruent	2924 (159)	2514 (135)	2045 (115)
Incongruent	3011 (169)	2737 (121)	2583 (140)
Meaningless	3078 (171)	2885 (166)	2316 (118)
No gestures	3091 (154)	2784 (152)	2451 (139)
	L2 to L1 translation		
Congruent	2625 (146)	2399 (126)	2146 (121)
Incongruent	3059 (196)	2769 (153)	2528 (136)
Meaningless	2899 (188)	2780 (172)	2557 (148)
No gestures	2918 (159)	2568 (128)	2175 (122)

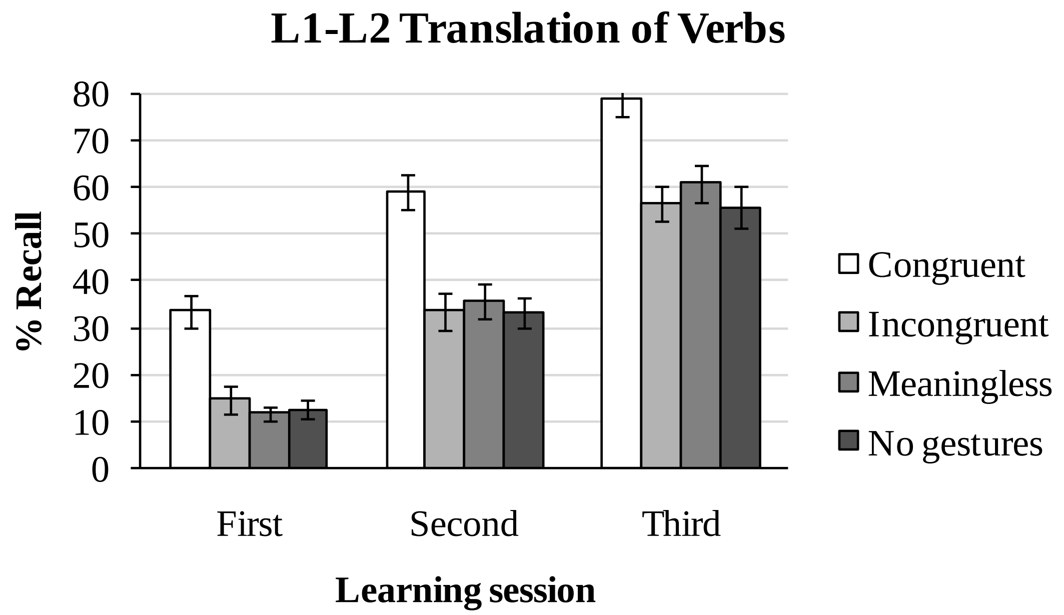
Note. Mean RTs (in milliseconds) obtained in Experiment 2 (learning L2 verbs) as a function of the translation test (L1 to L2 translation, L2 to L1 translation), the training session (first session, second session, third session) and the learning condition (congruent gestures, incongruent gestures, meaningless gestures, no gestures). Standard errors are in brackets.

The main effect of session was significant, $F(2, 62) = 23.05$, $p < .001$, $\eta_p^2 = .43$. Mean translation latency was 2951 ms ($SE = 139$) in the first session, 2680 ms ($SE = 115$) in the second session, and 2350 ms ($SE = 97$) in the third session. Linear trend analysis was significant, $F(1, 31) = 39.37$, $p < .001$, $\eta^2 = .56$, with faster responses at the end of the learning process relative to the beginning of training. The main effect of learning condition was also significant, $F(3, 93) = 9.96$, $p < .001$, $\eta_p^2 = .24$. The mean RTs were 2494 ms (*SE* = 114) in the congruent condition, 2777 ms (*SE* = 111) in the incongruent condition, 2760 ms (*SE* = 130) in the meaningless condition, and 2776 ms (*SE* = 119) in the no gesture condition. RTs were faster in the congruent condition compared with the incongruent condition, $F(1, 31) = 29.96$, $p < .001$, $\eta^2 = .49$, the meaningless condition, $F(1, 31) = 13.45$, $p < .001$, $\eta^2 = .30$, and the no gesture condition, $F(1, 31) = 9.57$, $p = .004$, $\eta^2 = .23$. No differences were found between the incongruent and the meaningless condition, $F < 1$; and the meaningless condition did not differ from the no gesture condition, $F < 1$. No other main effects or interactions were significant (all *p*s > .05).

Recall performance

In the Translation test x Session x Learning condition ANOVA, the main effect of session was significant, $F(2, 62) = 240.80$, $p < .001$, $\eta_p^2 = .89$. The recall percentage was 20% (*SE* = 1.91) in the first session, 43% (*SE* = 2.92) in the second session, and 64% (*SE* = 3.10) in the third session. Linear trend analysis was significant, $F(1, 31) = 325.47$, $p < .001$, $\eta^2 = .91$; the recall of L2 words was higher at the end of training relative to the beginning of learning. The main effect of learning condition was significant, $F(3, 93) = 23.06$, $p < .001$, $\eta_p^2 = .43$, and the main effect of translation test was also significant, $F(1, 31) = 9.23$, $p = .005$, $\eta_p^2 = .23$. Moreover, the Translation test x Learning condition interaction was significant, $F(3, 93) = 9.14$, $p < .001$, $\eta_p^2 = .23$. No other main effect or interactions were significant (all *p*s > .05). When we examined the Translation test x Learning condition interaction, we obtained differences due to the translation test in the no gesture condition, $F(1, 31) = 32.84$, $p < .001$, $\eta^2 = .51$. This analysis revealed better recall in the backward translation (45%, *SE* = 3.09) relative to the forward translation (34%, *SE* = 2.83). There were no significant differences between the two translation directions in the congruent condition, incongruent condition, and meaningless condition (all *p*s > .05). As in Experiment 1, we examined the learning condition effect across training sessions in the two tests used to measure the learning of L2 words.

Forward translation (from Spanish-L1 to Vimmi-L2). The main effect of session was significant, $F(2, 62) = 209.66$, $p < .001$, $\eta_p^2 = .87$. Recall percentage was 18% (*SE* = 1.76) in the first session, 40% (*SE* = 2.72) in the second session, and 63% (*SE* = 3.14) in the third session. Linear trend analysis was significant, $F(1, 31) = 301.32$, $p < .001$, $\eta^2 = .91$. Thus, the vocabulary learning increased throughout the course of training. Moreover, a significant learning condition effect was found, $F(3, 93) = 25.27$, $p < .001$, $\eta_p^2 = .45$. The Session x Learning condition interaction was not significant, $F(6, 186) = 1.22$, $p = .30$, $\eta_p^2 = .04$. When learning conditions were compared (see Table 2), we observed better recall in the congruent gesture condition (57%, *SE* = 3.24), relative to the meaningless gesture condition (36%, *SE* = 2.70) and the no gesture condition (34%, *SE* = 2.83). The recall performance was similar in the incongruent condition (35%, *SE* = 3.10), and the meaningless condition. Finally, compared with the no gesture condition, the recall percentage was similar in the meaningless condition and the no gesture condition (Figure 6).



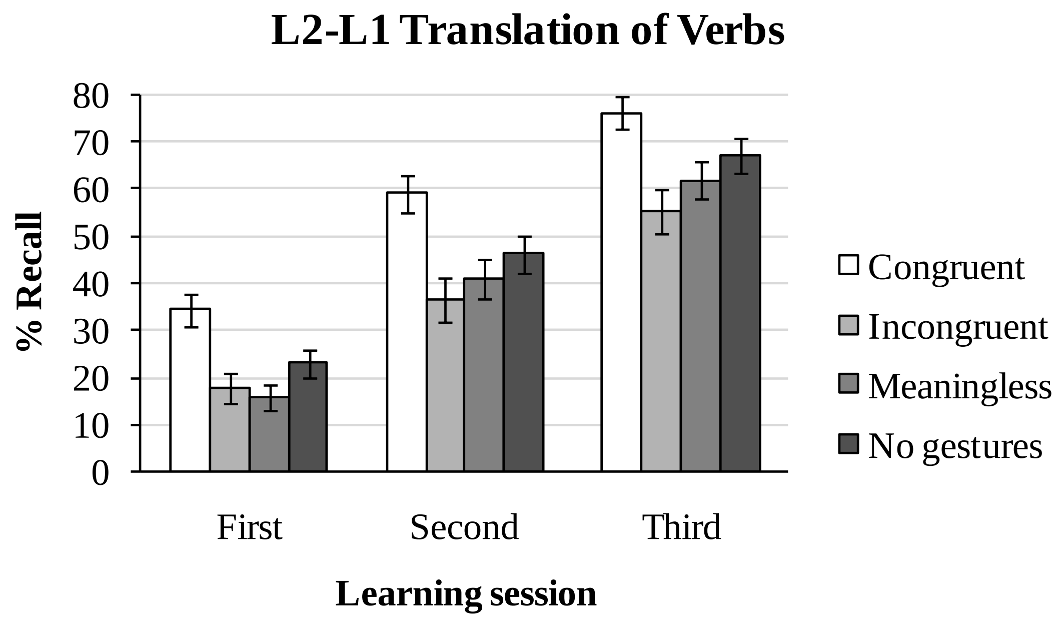


Figure 6. Recall percentages (% Recall) obtained in Experiment 2, during the translation of verbs from L1 to L2 (upper graph) and from L2 to L1 (bottom graph) across training sessions (first, second, third) and gesture conditions (congruent, incongruent, meaningless, and no gestures). Standard error is plotted in vertical lines.

Backward translation (from Vimmi-L2 to Spanish-L1). Figure 6 shows the results obtained in the backward translation test. The ANOVA conducted with session and learning condition as variables revealed a significant main effect of session, $F(2, 62) = 173.80$, $p < .001$, $\eta_p^2 = .85$. The recall of Spanish words from Vimmi words was 23% (*SE*

= 2.26) in the first session, 46% (*SE* = 3.45) in the second session, and 65% (*SE* = 3.33) in the third session. Linear trend analysis indicated that recall performance increased as a function of training, $F(1, 31) = 247.57$, $p < .001$, η^2= 89. The main effect of condition was also significant, $F(3, 93) = 16.32$, $p < .001$, η_p^2= .34. Finally, the Session x Learning condition interaction was not significant, $F(6, 186) = 1.32$, $p = .25$, $\eta_p^2 = .04$. When the learning conditions were compared (Table 2), we observed better recall in the congruent condition (57%, *SE* = 3.30) relative to the meaningless condition (39%, *SE* = 3.22) and the no gesture condition (45%, *SE* = 3.10). The difference in recall between the incongruent condition (36%, *SE* = 3.76) and meaningless condition was not significant. However, compared with the no gesture condition, recall was lower in the incongruent condition and the no gesture condition.

DISCUSSION

In Experiment 2, we evaluated the role of gestures in the acquisition of L2 verbs. The pattern of results we found was very similar to that observed in Experiment 1 (learning of L2 nouns). We found a facilitation effect due to the use of gestures in the learning process. Congruent gestures facilitated the acquisition of L2 verbs relative to the learning with no gestures or meaningless gestures. However, we obtained an interference effect due to the use of incongruent and meaningless gestures relative to the no gesture condition in the L2 to L1 translation task. The two effects found here (facilitation and interference) will be explained further in the next section.

Between experiments results

We evaluated the differential effect of gestures on the learning of L2 nouns (Experiment 1) and L2 verbs (Experiment 2) (Figure 7). In the forward translation task, the main effect of word type was significant, $F(1, 55) = 4.50$, $p = .04$, $\eta_p^2 = .08$. Participants recalled more L2 nouns (49%, *SE* = 2.95) than L2 verbs (40%, *SE* = 2.61). Furthermore, the Condition x Word type interaction was significant, $F(3, 165) = 2.85$, $p = .04$, $\eta_p^2 = .05$. No other interactions involving the word type variable were significant (*p*s > .05).

The recall of verbs was lower relative to the recall of nouns in the no gesture condition, $F(1, 55) = 9.49$, $p = .003$, $\eta_p^2 = .15$, and the incongruent condition, $F(1, 55) = 3.97$, $p = .05$, $\eta_p^2 = .07$.

However, no differences due to word type were found in the meaningless condition, $F(1, 55) = 2.18$, $p = .15$, $\eta_p^2 = .04$, and the congruent condition, $F < 1$. When the backward translation was taken into account, the main effect of word type was not significant, $F < 1$, nor did this variable interact with any other factor (all *p*s> .05).

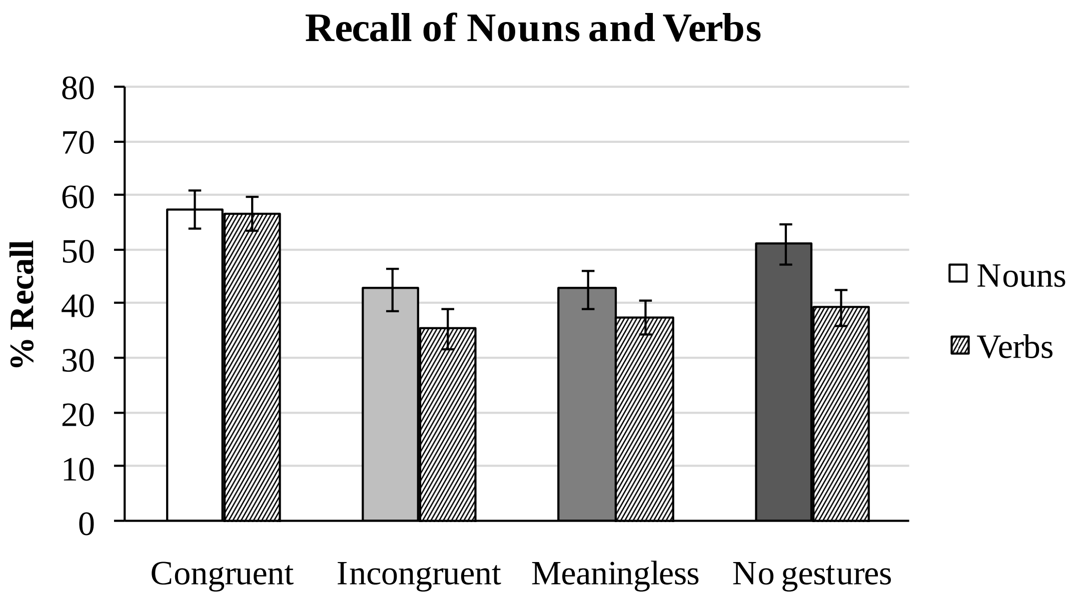


Figure 7. Comparison of recall percentages (% Recall) of nouns (Experiment 1) and verbs (Experiment 2) across gesture conditions (congruent, incongruent, meaningless, no gestures). Standard errors are plotted in vertical lines.

To specify further the nature of the differences in L2 learning depending on the word type, we evaluated possible differences between nouns and verbs in lexico-semantic variables reported in the method section of Experiments 1 and 2. The lexical frequency of nouns (16.07, *SD* = 35.74) and verbs (15.48, *SD* = 23.62) was similar, $t(78) = 0.09$, $p = .93$. The number of graphemes of nouns (6.57, *SD* = 1.72) and verbs (6.53, *SD* = 1.48) did not differ, $t(78) = 0.13$, $p = .89$. Further, the number of phonemes was equated in the case of nouns (6.27, *SD* = 1.63) and verbs (6.30, *SD* = 1.44), $t(78) = 0.07$, $p = .94$. The familiarity of words was similar for nouns (3.95, *SD* = 2.85) and verbs (3.50, *SD* = 3.09), $t(78) = 0.67$, $p = .50$. However, differences between nouns and verbs emerged when concreteness was considered, $t(78) = 1.91$, $p = .05$. The concreteness value was higher for nouns (4.06, *SD* = 2.88) than for verbs (2.90, *SD* = 2.53).

GENERAL DISCUSSION

Native English speakers beginning school have a vocabulary of around 4,000 to 5,000 word families, adding about 1,000 word families each year until reaching adulthood with a vocabulary of around 20,000 word families (Nation & Waring, 1997). This observation means that learners of a L2 face a considerable challenge, and there is hence a need to develop training methods that can facilitate L2 vocabulary acquisition. The use of gestures might be a good tool to foster this learning process (see Gullberg, 2014, for a review). Previous empirical studies support this premise by showing that L2 vocabulary learning is facilitated by the inclusion of gestures in the learning context (Macedonia et al., 2011; Masumoto et al., 2006; Quinn-Allen, 1995; Tellier, 2008).

In previous studies, several conditions have been considered to evaluate the role of gestures. For instance, the comparison between a congruent and incongruent condition using speech or gestures has been addressed during the learning of verbs (Kelly et al., 2009) and nouns (Macedonia et al., 2011). Other studies have compared congruent gestures relative to meaningless gestures presented with written words (Krönke et al., 2013) and sentences (Straube et al., 2012). Previous work has also compared conditions with gestures relative to conditions without gestures (Macedonia & Knösche, 2011). Moreover, congruent, incongruent, and no gesture conditions have also been considered (Feyereisen, 2006). However, to our knowledge, there are no previous studies in which four conditions have been evaluated at the same time (congruent, incongruent, meaningless, and no-gesture condition). The simultaneous comparison of these conditions in a single experiment would be of use in evaluating models about the role of gestures in L2 vocabulary learning. The current study set out to achieve this aim by evaluating four training procedures during the learning of nouns and verbs in a FL.

The results found in the experiments reported here showed facilitation and interference effects associated with the use of gestures in L2 learning (see also, Kelly et al., 2009). This pattern of results runs counter to the self-involvement account of the role of gestures in L2 vocabulary acquisition (Helstrup, 1987). According to this view, gestures might foster the involvement of the participant in the learning task and thus, whenever gestures are used, an improvement in word acquisition would be found. Focusing on the facilitation effect, we observed that participants learned more words in a FL when the word pairs to be learned were accompanied by congruent gestures (Feyereisen, 2006; Kelly et al., 2010; Macedonia & Klimesch, 2014). This finding agrees with the motor-imagery account of the role of gestures in L2 acquisition (Denis et al., 1991). In the congruent condition, the motor trace associated with the gesture was part of the meaning of the word. Hence, semantic processing was enhanced, which promoted the acquisition of L2 words. The facilitation effect observed with congruent gestures could also be explained by the motor-trace perspective (Engelkamp & Zimmer, 1984, 1985). According to this view, familiar gestures would have a strong motor trace representation in memory, which would favor L2 learning. A way of dissociating between the motor-trace and the motor-imagery account would be the use of beats and deictic gestures. Previous studies have evaluated the role of these gestures in L2 learning (Kushch, Igualada, & Prieto, 2018; Morett, 2014; So et al., 2012). For example, Morett showed that the use of beat and deictic gestures facilitated the recall of words in a language not familiar to the participants. Beats and deictic gestures are typically produced with language and they have meaning associated with them. However, their meaning may not be part of a word's meaning (particularly for beats). Therefore, the comparison between iconic gestures versus beats would be of interest to dissociate between theoretical perspectives. A facilitation effect with beats would favor the motor-trace theory, and a greater facilitation effect in case of iconic gestures would be explained by the motor-imagery account.

It is important to stress that different explanations might work together to explain the role of gestures in L2 vocabulary learning. Macedonia and Mueller (2016) found neural evidence of the relationship between higher cognitive processes associated with the learning of words with iconic gestures (i.e., attention, language, sensory and motor processes, declarative and procedural memory). Thus, it would be simplistic and reductionist to limit the effect of gestures on L2 learning to only one cognitive mechanism. Additionally, in our study, the beneficial effect of using gestures in L2 vocabulary acquisition was considered during the processing of isolated words. However, different studies have found that word processing in L2 is affected by local sentence context and more global discourse context (Van Assche et al., 2016, for a review). Thus, future studies should evaluate the effect of gestures on the acquisition and processing of L2 words in semantically rich contexts.

Nevertheless, not all well practiced gestures facilitated vocabulary learning in our study. In particular, incongruent gestures (familiar gestures with an easily recognizable meaning that mismatched the meaning of the word pair to be learned) made L2 learning difficult (Feyereisen, 2006; Kelly et al., 2010; Macedonia et al., 2011). The interference effect found with incongruent gestures relative to the learning of L2 words without gestures might be difficult to explain merely in terms of semantic processing. This account would predict reduced L2 learning in the incongruent condition (meaningful gestures) relative to the meaningless condition (meaningless gestures) because in the incongruent condition there was a mismatch between the meaning of the gestures and the meaning of the words to be learned in L2. However, there were no differences between these two conditions. The similar pattern of results found in the incongruent iconic gesture condition and the meaningless gesture condition can be explained by assuming that meaningless gestures functioned as self-adaptor gestures (self-touching movements unconnected to the content of the speech). In fact, previous studies have shown that both iconic gestures and self-adaptor gestures play a similar role during lexical processing in speech production (Beattie & Coughlan, 1999). In addition, both the incongruent and meaningless condition were associated with lower L2 vocabulary learning than that found in the no gesture condition. The only difference between the incongruent/meaningless conditions compared with the no gesture condition was the involvement of the participant in a dual vs. single learning context, respectively. Specifically, participants in the incongruent and meaningless conditions were immersed in a dual task situation in which they had to code the gesture and the L1 word at the time they learned the corresponding word in the FL. The cost associated with this dual coding is also found in other studies, which have confirmed the difficulty of coding the meaning of a message in L2 at the time at which the learners are required to perform a concurrent task (Bransdorfer, 1991; Van Patten, 1990; Wong, 2001). Thus, the use of meaningless or incongruent gestures in the learning of L2 would result in a dual task that makes L2 learning difficult. However, this idea contrasts with the absence of cognitive cost when words and gestures are processed concurrently in everyday interactions. Previous studies suggest that the cognitive cost of a concurrent task depends on the familiarity that people have with the words that are learned. For example, Padilla, Bajo, and Macizo (2005) showed that the adverse effect of a concurrent task (articulatory suppression) depended on the participants' word knowledge with the learning material. A cognitive cost of the concurrent task was observed in the learning of pseudowords but not in the retention of words already known to the participants. Thus, in normal circumstances, long-term knowledge associated with known words provides support for the retention of information even in dual task situations. In our study, however, this long-term knowledge of words in L2 was absent, so gestures would have an adverse effect on vocabulary acquisition unless the gestures were congruent with the learning material.

In the current study, we also evaluated the differential role of gestures in the learning of two types of words, i.e. nouns (Experiment 1) and verbs (Experiment 2). Verbs are more difficult to learn than nouns, at least for children. Children acquire verbs slower than nouns and adults usually perform better with nouns than verbs on a range of tasks (Gentner, 1981, 1982; Gentner & Boroditsky, 2001). In addition, L2 learners show better acquisition of nouns than verbs (Lennon, 1996). We confirmed this pattern of results in the current study. In the forward translation task, the recall of verbs was lower relative to the recall of nouns in the no gesture condition. In addition, when examining the impact of gestures, we found similar facilitation and interference effects during the learning of L2 nouns and verbs. However, upon closer inspection, we observed that the difficulty associated with the learning of verbs disappeared when congruent gestures were included in the training. Thus, the use of gestures in L2 vocabulary acquisition appears to remediate the intrinsic difficulty associated with the learning of verbs. This effect appears to agree with the motor-imagery account since differences between nouns and verbs only disappeared when the gestures denoted the same meaning as the words to be learned in L2. Moreover, the differential effect of gestures on the acquisition of verbs in a FL can be accommodated within the gestures for conceptualization hypothesis (Kita, Alibali, & Chu, 2017). According to this view, gestures are rooted in practical actions that involve body movements and motoric contents, and the meaning of verbs also intrinsically denotes motoric information. Thus, gestures might involve the direct simulation of the meaning of verbs, which would facilitate the learning of this category of words.

It is important to note that the conclusions drawn from the differential effect of the use of gestures during the learning of verbs and nouns in L2 should be moderated. In Experiment 2, the mapping between representational gestures which involve depicted actions and the semantic characteristics of verbs which refer to actions also is stronger than that between gestures and nouns (Experiment 1). In fact, this may explain why the vast majority of studies on gestures have used verbs as learning material (Childers & Tomasello, 2002; De Grauwe et al., 2014; Kelly et al., 2009; Kircher et al., 2009). Thus, it might be possible that the benefit associated with the use of congruent gestures during the learning of verbs might reflect this strong semantic overlap between the actions described by the gestures and the meaning of the verbs. On the other hand, besides the semantic differences between nouns and verbs, both types of words differ in grammatical category. To explore this point, the role of gestures during the acquisition of L2 could be evaluated by mixing nouns and verbs as learning material. It is possible that the mixing of nouns and verbs produced a greater salience of their grammatical category and possibly it would maximize the differences found in the role of gestures during the learning of these two types of words. Future studies will shed light on this point.

In this section, we argue that the benefit of using congruent gestures in L2 vocabulary learning might be accounted for by a motor-imagery explanation (Denis et al., 1991). According to this view, gestures promote semantic processing of the material to be learned. However, there remains some uncertainty regarding the specific mechanism by which the increased semantic processing associated with the use of gestures promoted the learning of L2 words. The impact of gestures on L2 vocabulary learning can be accommodated within the RHM (Kroll & Stewart, 1994). As described in the introduction section, in this model, L1 words, L2 words, and a shared semantic system are interconnected. However, differences in the weight of these connections emerge depending on the stage of L2 vocabulary acquisition. In the early stages, the links between the semantic system and the L2 words are weak and L2 learners preferentially use a lexical route of processing from L2 to L1. Furthermore, when proficiency increases, the links between L2 words and the semantic system develop while diminishing the weight of the lexical route. The model has been widely supported by previous work. For example, unbalanced bilinguals show an asymmetry in translation tasks with faster performance in backward translation (lexical route from L2 to L1) than in forward translation (semantic route from L1 to L2) (Kroll & Stewart, 1994).

Increased vocabulary learning has been demonstrated with L2 learning methods that foster a semantic route of processing (e.g., presenting a word to learn with a picture denoting its content, Lotto & De Groot, 1998; Van Hell & Candia-Mahn, 1997; or imagining the meaning of a word to be learned, Ellis & Beaton, 1993; Wang & Thomas, 1995). Previous studies demonstrate that gestures enrich the encoding of the words to be learned by adding sensorimotor networks and procedural memory to the semantic/declarative memory associated with the meaning of the words (Macedonia & Mueller, 2016). Hence, gestures enhance semantic processing of words. The findings obtained in the current study have provided evidence suggesting the use of semantic-L2 connections associated with the use of gestures. As mentioned previously, gestures abolished the difficulty associated with the learning of verbs but not nouns. This effect was not captured in backward translation but in forward translation, a task that is semantically mediated to a greater extent than backward translation (see Kroll et al., 2010, for a critical review). Moreover, when we explored the characteristics of nouns and verbs in our material, concreteness produced a difference: Nouns were associated with a higher concreteness value relative to verbs, a variable that has been shown to modulate L2 vocabulary learning. For example, concrete words activate the semantic system more robustly than abstract words (Van Hell & De Groot, 1998), with the former being more readily acquired by L2 learners (Kaushanskaya & Rechtzigel, 2012).

To conclude, the current study suggests that the use of gestures that are congruent in meaning with the word to be learned facilitates vocabulary acquisition in a FL. Furthermore, congruent gestures reduce the difficulty associated with the learning of verbs relative to the learning of nouns. Other methods based on conceptual analysis of the material foster L2 learning (e.g., the use of pictures paired with the words to be learned, e.g., Tellier, 2008). It would be interesting in the future to evaluate whether these two supporting materials might have additive effects on L2 vocabulary acquisition.

Experiment 3. Seeing or Acting? The Effect of Performing Gestures on Foreign Language Vocabulary Learning

We evaluate the impact of gestures during the learning of vocabulary in a FL. Spanish speakers learned words in a L2 in four gesture conditions according to the relationship between the meaning of the words and the gestures (congruent gestures, incongruent gestures, gestures without meaning and no gestures). The participants learned the words by performing gestures ("do" learning group) or by observing the gestures performed by others ("see" learning group). Compared to the meaningless gesture condition, the processing of congruent gestures facilitated the recall of L2 words in the "see" and "do" learning groups. However, the interference effect associated with the processing of incongruent gestures was greater in the "see" learning group than in the "do" learning group. Thus, the performance of gestures seems to mitigate the negative impact that the use of gestures may have on the acquisition of vocabulary in a FL.

This study has been submitted as García-Gámez, A. B., & Macizo, P. (2019). Seeing or Acting? The Effect of Performing Gestures on Foreign Language Vocabulary Learning. *Language Learning.*

INTRODUCTION

Bilingualism has become the rule rather than the exception. To illustrate, in 2016, around two thirds of working-age adults in the European Union knew at least one FL (Eurostat, 2019). However, it is also true that not everyone speaks a L2. Therefore, it is necessary to implement learning strategies that enhance L2 learning. From a theoretical perspective (Kroll & Stewart, 1994), the acquisition of vocabulary in L2 would imply the establishment and reinforcement of connections between semantics and the new words that are learned in L2. Thus, while novice learners preferably translate words through the use of lexical connections between L1 and L2 words, expert learners translate words across languages using connections between the meaning and the lexicon in L2 (Talamas et al., 1999).

In the past, different strategies have been implemented that maximize L2 vocabulary learning by stressing the semantic processing of the material. Increased vocabulary learning has been demonstrated with L2 learning methods that foster a semantic route of processing (e.g., presenting a word to learn with a picture denoting its content, Lotto & De Groot, 1998; Van Hell & Candia-Mahn, 1997; or imagining the meaning of a word to be learned, Ellis & Beaton, 1993; Wang & Thomas, 1995). Moreover, the use of gestures that represent the meaning of words to be learned has a positive effect on vocabulary acquisition (De Grauwe et al., 2014; Gullberg, 2014; Kelly, Manning, & Rodak, 2008; Kelly, et al., 2009; Macedonia & Knösche, 2011; Macedonia & Kriegstein, 2012; Morett, 2014; So et al., 2012; McCafferty & Stam, 2008, for reviews).

In the current study, we focus on the role of gestures in L2 vocabulary acquisition. Specifically, we evaluate whether mere exposure to gestures is sufficient to show the positive effect of gestures on the learning of L2 words. Alternatively, we examine if it is necessary for the participants to perform the gestures in order to observe their effect on learning.

The role of movements on learning

In the educational system, the possible advantages associated to the learning through actions relative to the observation-based learning have been a topic under discussion for decades (Goldin-Meadow, 1999). The "learning-by-doing" perspective defends the active participation of the individual in the learning process by performing actions during educational development. Learning-by-doing can positively affect the formation of neural networks underlying the acquisition of knowledge and the performance of many cognitive skills (Goldin-Meadow et al., 2012). This beneficial effect has been demonstrated in several educational fields such as online courses, language learning, or technology use (Aleven & Koedinger, 2002; Bessen, 2015).

Moreover, the relevance of movement in language processing has been confirmed in many studies (Glenberg et al., 2004; Koriat & Pearlman-Avion, 2003). For instance, Glenberg et al. (2008) observed that the performance of movements interacts with language comprehension. In their study, participants were required to move 600 beans (one at a time) from a big recipient to a narrow container toward or away from their bodies depending on the recipients' location. Afterwards, meaningful and meaningless sentences describing movements were presented to participants and they judged their plausibility. Results revealed that the time required to judge the sentences depended on the congruency between the beans' movement direction and the direction of the acts described in the sentence (toward or away from the body). Therefore, the performance of actions determined language comprehension.

A form of movement directly related to language processing are gestures that usually occur with speech. There are different types of gestures associated to language processing (McNeill, 1992), such as one or more fingers directed to a reference (deictic gestures), hand movements that reflect the prosody and emphasize the speech (beat gestures), gestures culturally specific such us thump up and hand in fist to indicate "good", and representational or iconic gestures used to illustrate the meaning of what is being said. As indicated above, there are a large number of studies showing that the use of iconic gestures benefits the acquisition of vocabulary in L2 (Feyereisen, 2006; Kelly et al., 2009; Macedonia & Klimesch, 2014; Morett, 2014).

Performing gestures during word learning in L2 would have a beneficial effect for different reasons. Firstly, gestures might favor the involvement of the participant in the learning task (Helstrup, 1987) so they could facilitate enhanced attention to the learning material which would increase the retention of words (Craik & Tulving, 1975).

Moreover, the performance of gestures when individuals process new words would promote the creation of a visual image associated with the meaning of this word, which would reinforce the semantic content of the word to be learned (Denis, et al., 1991). Finally, gestures would enrich the encoding of the words to be learned by adding sensorimotor networks and procedural memory to the semantic/declarative memory associated with the meaning of the words (Macedonia & Mueller, 2016).

Previous studies reveal that better vocabulary learning is found when participants learn L2 words accompanied with gestures that reflect the common use of objects whose names have to be learned in L2 (Feyereisen, 2006; Kelly et al., 2009; Macedonia & Klimesch, 2014). Thus, the use of gestures promoted the acquisition of L2 words. However, a negative effect is also found associated to the use of gestures in L2 vocabulary learning. Incongruent gestures (familiar gestures with an easily recognizable meaning that mismatch the meaning of the word to be learned) make L2 learning difficult (Feyereisen, 2006; Kelly et al., 2009; Macedonia et al., 2011). The negative effect of incongruent gestures might be due to a semantic interference effect (Bernardis et al., 2008; Yap et al., 2010). To illustrate, Bernardis et al. used a gesture-word priming paradigm in which participants received a gesture prime followed by a word that they had to name. Compared to a baseline condition without gestures, the authors observed faster naming times when the gesture was congruent (i.e., a facilitation effect) and slower latencies when the gesture was incongruent with the meaning of the word (i.e., an interference effect). The semantic interference effect was interpreted as due to the difficulty to integrate the meaning of the gesture and the word when participants processed gesture-word pairs with different meanings. Thus, the processing of gestures could benefit or impair L2 learning depending on the congruency between the gestures and the meaning of the words (Feyereisen, 2006; Kelly et al., 2009; Macedonia et al., 2011).

Consequences of seeing vs. doing while learning

Several studies have examined the differences between self-performed tasks and experimenter-performed tasks (Cohen, 1981; Engelkamp & Zimmer, 1983). In 2012, Goldin-Meadow and colleagues directly compared the role of self-performed gestures versus seeing another individual producing them when children learned a mental transformation task. In their study, children were required to perform a mental rotation task in which they had to indicate whether two shapes presented in different orientations were the same figure or not. They used this task because previous studies demonstrate the close relationship between mental rotation and motor processing. When participants are instructed to mentally rotate a target, premotor areas involved in the planning of actions become active (Ganis et al., 2000; Glenberg et al., 2008) and participants spontaneously make gestures when they are required to explain how they solve this task (Chu & Kita, 2008). Goldin-Meadow et al. (2012) showed that children obtained better results when they were instructed to produce the gesture needed to solve the transformation task rather than when they observed the experimenter doing the movements. Thus, the performance of movements improved the learning compared to the mere exposure to gestures associated with the mental rotation of objects.

Other empirical studies have confirmed the importance that self-generation of movements have in the learning of linguistic material. Engelkamp et al. (1994) asked participants to learn sentences while performing the actions described in the sentences or by simply listening to and memorizing the material. The results revealed that the recall of sentences was higher when participants performed actions during the learning phase. The authors interpreted that the performance of actions favored the formation of a motor trace that benefited the retention of information. Empirical evidence supporting this interpretation comes from the study conducted by James and Swain (2011). The authors taught children action words associated to concrete toys. Some of the children manipulated the objects while learning and the remaining children observed the experimenter manipulating the same objects. When children listened to the words they had previously learned, motor brain areas were activated only in children who perform the toys manipulation themselves. Thus, the performance of motor actions favor the learning of new words and the benefit associated with the performance of movement during learning seems to be due to the formation of a motor trace that would be activated during the subsequent retrieval of information.

However, other studies have found similar pattern of results when participants produce actions and when they only see actions produced by others (e.g. Rizzolatti & Craighero, 2004). For example, Stefan and colleagues (2005) observed the involvement of the motor cortex during the observation of movement (simple repetitive thumb movements) which produced a specific memory trace in the motor cortex similar to the pattern of activation that occurs when people perform a motor action. Finally, in a study addressing the effect of movements on sentences reading comprehension in children, Glenberg and colleagues (2004, Experiment 3) found intermediate results. Children were exposed to histories happening in a particular scenario (a farm) where different referents appeared (a sheep or a tractor). For the first group of children, objects referred to in the text were present and they were instructed to simulate the sentence content by manipulating the objects. The second group of children was required to imagine they were manipulating the toys. The results showed a beneficial effect of the manipulation condition while the imagined condition presented a modest improvement compared to an only-read condition. Thus, although movement performance appears to improve learning, it is not clear whether a learning protocol that involves self-generated actions would have an additional benefit to the mere observation of movements.

The current study

In general, there is agreement about the facilitative role that gestures have in L2 vocabulary learning (Feyereisen, 2006; Kelly et al., 2009; Macedonia & Klimesch, 2014). However, there is controversy regarding the role that the performance of self-generated movement ("do" learning) versus the mere observation of movement ("see" learning) has in L2 vocabulary acquisition. In our study, we directly evaluated this point. To this end, monolingual Spanish (L1) speakers learned new words in an artificial language (Vimmi, L2) in three consecutive days. Participants were randomly assigned to two learning conditions; the "see" and the "do" learning groups. The "see" learning group was required to read aloud Spanish-Vimmi word pairs (L1-L2) and to observe gestures that were presented on a video at the same time. The "do" learning group was instructed to read aloud the word pairs in Spanish and Vimmi (L1-L2) and to imitate the gestures that were presented on the screen. Furthermore, in order to explore both the benefit and the potential cost associated with the use of gestures in L2 vocabulary learning, four conditions were evaluated: (a) the L1-L2 word pair to be learned was presented alone (no gesture condition), (b) coupled with a gesture that matched the meaning of the word (congruent gesture condition), (c) coupled with a gesture which was semantically unrelated to the word meaning (incongruent gesture condition) or (d) coupled with a gesture that did not denote any specific meaning (meaningless gesture condition). Verbs were selected as learning material for this study. Most of research regarding the role of gestures in L2 vocabulary learning has used this type of words because their semantic representations intrinsically involve a gestural/motor component (Boulenger et al., 2009; Hauk et al., 2004).

Based on previous studies about the role of gestures on L2 vocabulary learning, we expected to find a positive effect from the use of congruent gestures during the acquisition of L2 words. At the same time, the processing of gestures not related to the meaning of words would impair L2 learning. This pattern of results would confirm both the benefit and the cost of using gestures during vocabulary acquisition in a FL (Feyereisen, 2006; Kelly et al., 2009; Macedonia et al., 2011). However, the most relevant predictions referred to the role of performing vs. observing gestures during L2 learning. If mere observation of gestures were sufficient to modulate vocabulary acquisition, the pattern of outcomes would not depend on the type of training. On the contrary, if active training involving the performance of gestures maximizes learning, the learning rate would be higher in the "do" training group compared to the "see" learning group.

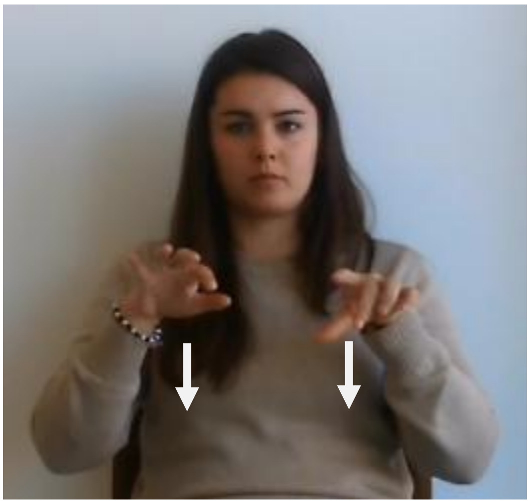
METHOD

Participants

Thirty one individuals participated in the experiment (29 women and 3 men). Sixteen of them (15 women and 1 man) were assigned to the "do" learning group, their mean age was 21.12 years (*SD* = 2.53). The remaining fifteen participants (13 women and 2 men) were assigned to the "see" learning group, their mean age was 21.13 years (*SD* = 2.72). Each subject provided written informed consent before performing the experiment.

Design and Materials

The participants were randomly divided into two experimental groups. Half of them learned words with the “see” training and the rest with the “do” training. Four L2 vocabulary learning conditions were manipulated within-participants: No gesture condition: Spanish (L1) - Vimmi (L2) word pairs had to be learned without gestures (e.g., *peinar*, to brush in Spanish, and *tola*, a Vimmi word). Meaningless condition: L1-L2 word pairs to be learned were coupled with meaningless gestures (e.g., *peinar-tola* and the gesture of moving the hand from the forehead to the ear). Congruent condition: L1-L2 word pairs were accompanied with gestures that reflected the common use of objects whose names had to be learned in L2 (e.g., *peinar-tola* and the gesture of holding an imaginary comb in the right hand and comb the hair from the front to the back). Incongruent condition: L1-L2 word pairs were coupled with a gestures associated with an action verb different from that denoted by the L1 word (e.g., *peinar-tola* and the gesture of moving both hands fingers as if typing on a keyboard) (see Figure 8).

Congruent	Incongruent
	
peinar-tola	*peinar-tola*

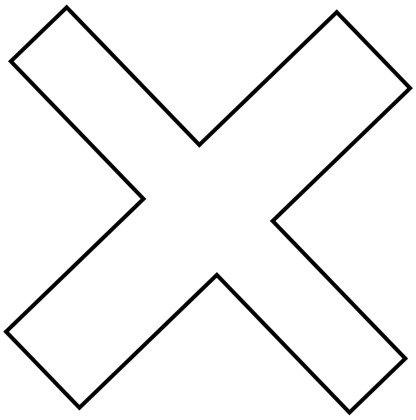
Meaningless	No gesture
	
peinar-tola	*peinar-tola*

Figure 8. Learning conditions used in the study. Spanish (L1) – Vimmi (L2) words (verbs) are coupled with different gesture conditions. In the example, peinar (to brush in Spanish) – tola (a Vimmi word) was accompanied by (a) the gesture of holding an invisible comb in the right hand and comb the hair from the front to the back (congruent condition); (b) the gesture of moving both hands fingers as if typing on a keyboard (incongruent condition), (c) the meaningless gesture of moving the hand from the forehead to the ear (meaningless condition), (d) the word pairs were presented without a gesture (no gesture condition).

The material (word pairs and gestures) used in the current study were the same as those described in García-Gamez and Macizo (2019; Experiment 2). The congruent and incongruent gestures presented along with the L1-L2 word pairs were iconic gestures (McNeill, 1992), which have been also called representational gestures

(Kendon, 1981), that usually illustrate a concrete physical object or movements associated with a known action. For example, for the meaning of eating, the gesture would involve holding an imaginary spoon and putting it to the mouth. The gestures used in the meaningless condition were small movements performed with the hand that did not have iconic or metaphoric associations with the meaning of physical items (for example, to form a fist with one hand and raise the fingers of the other hand). We took care to select meaningless gestures with similar properties to those of meaningful gestures (e.g., hand configuration, the use of simple movement trajectory and spatial location).

In addition, 40 words were selected in Spanish. These words were verbs denoting familiar actions performed with hands and face (e.g., to kiss, to pray, etc.) (see Appendix 1 for the complete set of materials). Forty words were also selected from an artificial language, Vimmi (Macedonia et al., 2011; Macedonia & Knösche, 2011). The corpus of Vimmi words is constructed so that it avoids factors that might favor the learning of specific items (co-occurrence of syllables, similarity with words from languages such as Spanish, English, and French). Vimmi words were carefully selected so that they were pseudowords with legal orthography and phonology in Spanish but without meaning. The forty Spanish words were paired randomly with the forty Vimmi words. The resulting 40 word pairs were randomly sorted into 4 sets of 10 word pairs. Four lists of 40 word pairs were created. Each list was composed of 10 word pairs in each of the four training conditions (no gesture condition, meaningless condition, congruent condition, and incongruent condition). Each participant received one list, with all participants being randomly assigned to one of the four lists. Across lists, the 40 words were counterbalanced over the four training conditions, so that all word pairs appeared in all training conditions.

The forty word pairs were randomly sorted into 4 sets (10 word pairs in each) and randomly assigned to one of the four learning conditions. The Spanish verbs across the 4 sets were equated in lexical variables (Davis & Perea, 2005). There were no differences across word sets in the number of graphemes, $F < 1$ (M = 6.53, SD = 1.48), number of phonemes, $F < 1$ (M = 6.30, SD = 1.44), number of syllables, $F < 1$ (M = 2.50, SD = 0.68), lexical frequency, $F < 1$ (M = 15.48, SD = 23.62, per one million count), familiarity, $F < 1$ ($M = 3.50$, $SD = 3.09$), and concreteness, $F < 1$ ($M = 2.90$, $SD = 2.53$). Finally, we controlled for the similarity between the L1 and L2 words. The Spanish and Vimmi words across conditions shared the same number of phonemes in the same position, $F < 1$ ($M = 0.37$, $SD = 0.63$), and irrespective of the position within the word, $F < 1$ ($M = 1.45$, $SD = 0.96$).

Procedure

L2 vocabulary learning involved three training sessions conducted on three consecutive days. In each session, participants performed, firstly, the L2 training and, afterwards, the assessment of the L2 learning. The two phases were separated by a 15-min break. E-prime experimental software was used for stimulus presentation and data acquisition (Schneider et al., 2002). Participants were informed that the training sessions would be recorded on video to be sure that they followed the instructions provided by the experiment, the participants agreed to the procedure and signed a written consent.

“See” L2 training. Participants were presented with a block of 40 Spanish-Vimmi word pairs. These word pairs were grouped (10 word pairs in each group) according to the four learning conditions (no gestures, meaningless gestures, congruent gestures, and incongruent gestures). This block was repeated 12 times. Hence, a participant received 480 trials where the 40 word pairs were presented 12 times. A short break was introduced between learning blocks. The word pairs were randomly presented within each condition. In addition, the order in which the learning conditions were presented within a block was counterbalanced. On each trial, the participant received a Spanish-Vimmi (L1-L2) word pair visually presented at the bottom of the screen. These word pairs were presented with a video where an actor performed the iconic gestures (see Figure 8). Gestures were recorded on video by the experimenter and they were congruent, incongruent, and meaningless, depending on the learning condition. The duration of each recorded gesture was five seconds and the gesture was repeated twice. The participants were instructed to read aloud each L1-L2 word twice. In the three gesture conditions (congruent, incongruent and meaningless), participants were instructed to imagining themselves imitating the actor gestures but they did not have to do any movement. They had to mentally produce the gesture presented each time they said aloud the L1-L2 word pairs, so they repeated

the gestures twice as the actor does in the videos. In the no gesture condition, the actor did not perform any movement and participants only repeated aloud the words pairs twice (L1 Spanish-L2 Vimmi). For example, when participants received the word pair *peinar-tola* along with the congruent gesture, they had to say aloud this word pair at the time they mentally produced the gesture of holding an imaginary comb in the right hand and to comb the hair from the front to the back. Once the participants had produced the word-pair twice, they had to press the space bar to continue to the next trial. Each training session lasted approximately 1 hour.

"Do" L2 training. As in the "see" training, participants were presented with a block of 40 Spanish-Vimmi word pairs. The learning material was exactly the same used in the "see" L2 training. The participants were instructed to read aloud each L1-L2 word twice. In the three gesture conditions, participants had to produce the gesture presented in the video each time they said aloud the L1-L2 word pair. Hence, they repeated the words pairs and the gesture twice. Participants started the production of the gesture when they began the production of the L1-L2 word pair. Once the participants had produced the word-pair twice, they had to press the space bar to continue to the next trial. The training lasted approximately 1 hour.

L2 learning assessment. Two tests were used to evaluate the acquisition of L2 words in the "see" and "do" learning groups: Translation from Spanish into Vimmi (forward translation from L1 to L2) and translation from Vimmi into Spanish (backward translation from L2 to L1). These tasks have been used in previous studies to evaluate L2 learning (Kroll & De Groot, 2005; Poarch et al., 2014) and in studies about the role of gestures in L2 vocabulary acquisition (García-Gámez & Macizo, 2019). The order in which the translation tests were presented was randomized across the three training sessions and across participants. In each translation task, the 40 Spanish words and the 40 Vimmi words were presented in the forward and backward translation tasks, respectively. On each trial, a word was presented in the middle of the screen until the participant produced its translation. Oral translations were recorded for later analyses of recall accuracy. Response times (RTs) from the presentation of the word until the beginning of the oral translation were also registered. The learning assessment lasted approximately 10 minutes depending on the participants' performance.

RESULTS

Translation direction (forward translation, backward translation), training session (first session, second session, third session) and learning condition (no gestures, meaningless gestures, congruent gestures, incongruent gestures) were considered as within-participants factors while the learning group ("see", "do") was introduced in the analyses as a between groups variable. Although recall percentages (Recall %) are the main index of vocabulary acquisition, RTs were also analyzed in this study. The RTs associated with correct translations were trimmed following the procedure described by Tabachnick and Fidell (2001) to eliminate univariate outliers. Raw scores were converted to standard scores (z-scores). Data points which, after standardizatiown, were 3 *SD* outside the normal distribution, were considered outliers. After removing outliers from the distribution, z-scores were calculated again. The filter was applied in recursive cycles until no observations were outside 3 *SD*. In all analyses, we adopted a significance level of $\alpha = 0.05$. The percentage of outliers was 10.62% in the "see" and 11.51% in the "do" learning groups. Only correct responses were included in the analyses of the RTs. Data points were excluded from the RT analyses if: (a) the participants produced nonverbal sounds that triggered the voice key, (b) the participants stuttered or hesitated in producing the word, (c) the participants produced something different than the word required. Some small errors were allowed and considered correct responses depending on the length of the correct word to be produced: (1) For monosyllabic words, the replacement of a vowel, (2) for disyllabic words, the replacement of a vowel or a consonant but not both, (3) for words with three or more syllables, the inversion of a vowel and a consonant or the replacement of a vowel or a consonant. The mean of recall was 48.39% (47.00% in the forward translation direction and 49.78% in the backward translation direction) in the "see" learning group and 59.98% (59.32% in the forward translation direction, and 60.63% in the backward translation direction) in the "do" learning group.

Table 4. *Statistical Analyses Performed on Recall Percentages and RTs. Main Effect and Interactions between the "See" and "Do" Learning Group.*

Effects	**Recall (%)**		**RT (ms)**	
	F	***p***	***F***	***p***
Learning Group	2.61	.12	6.72	.01*
Translation Direction	2.80	.10	17.71	.00*
Translation Direction x Learning Group	.37	.55	1.18	.29
Session	153.65	.00*	34.26	.00*
Session x Learning Group	2.11	.13	5.90	.00*
Learning Condition	18.96	.00*	6.91	.00*
Learning Condition x Learning Group	2.24	.09~	1.48	.23
Translation Direction x Session	12.45	.00*	1.90	.16
Translation Direction x Session x Learning Group	.02	.98	.11	.90
Translation Direction x Learning Condition	9.37	.00*	.92	.44
Translation Direction x Learning Condition x Learning Group	4.98	.00*	1.36	.26
Session x Learning Condition	1.32	.25	2.13	.05~
Session x Learning Condition x Learning Group	1.55	.16	2.13	.05~
Translation Direction x Session x Learning Condition	.26	.95	1.11	.36
Translation Direction x Session x Learning Condition x Learning Group	.52	.79	.85	.53

Note: *p < .05, ~p < .10.

Recall performance

Recall percentages were submitted to a Session x Learning Condition x Learning Group x Translation Direction analysis of variance (ANOVA). Table 4 shows the complete pattern of results obtained in the analysis. There was a significant main effect of session, $F(2, 58) = 153.65$, $p < .001$, $\eta_p^2 = .84$. The recall percentage was 30.19% ($SE = 4.02$) in the first session, 59.61% ($SE = 3.89$) in the second session, and 72.74% ($SE = 3.68$) in the third session. Linear trend analysis showed that the recall of L2 words was higher in the final session relative to the beginning of the training, $F(1, 29) = 177.52$, $p < .001$, $\eta^2 = .86$. The main effect of learning condition was significant too, $F(3, 87) = 18.96$, $p < .001$, $\eta_p^2 = .40$. Mean recall percentage was 63.72% ($SE = 3.64$) in the congruent condition, 50.07% ($SE = 4.11$) in the incongruent condition, 47.73 ($SE = 3.81$) in the meaningless condition, and 55.20% ($SE = 3.84$) in the no gesture condition. Compared to the meaningless condition, participants responded better in the congruent condition, $t(30) = 6.91$, $p < .001$, and the no gesture condition, $t(30) = 3.55$, $p = .001$. The difference between the meaningless condition and the incongruent condition was not significant, $t(30) = .86$, $p = .40$.

Although the overall recall percentage was higher in the "do" learning group (59.97%, $SE = 4.99$) than in the "see" learning group (48.39%, $SE = 5.15$), the main effect of learning group was not significant, $F(1, 29) = 2.60$, $p = .12$, $\eta_p^2 = .08$. However, the Learning Group x Learning Condition x Translation Direction was significant, $F(3, 87) = 4.98$, $p = .003$, $\eta_p^2 = .15$. No other three-way interactions showed significant differences (all *p*s > .05). The Learning Condition x Translation Direction interaction was analyzed further for each learning group separately (Figure 9).

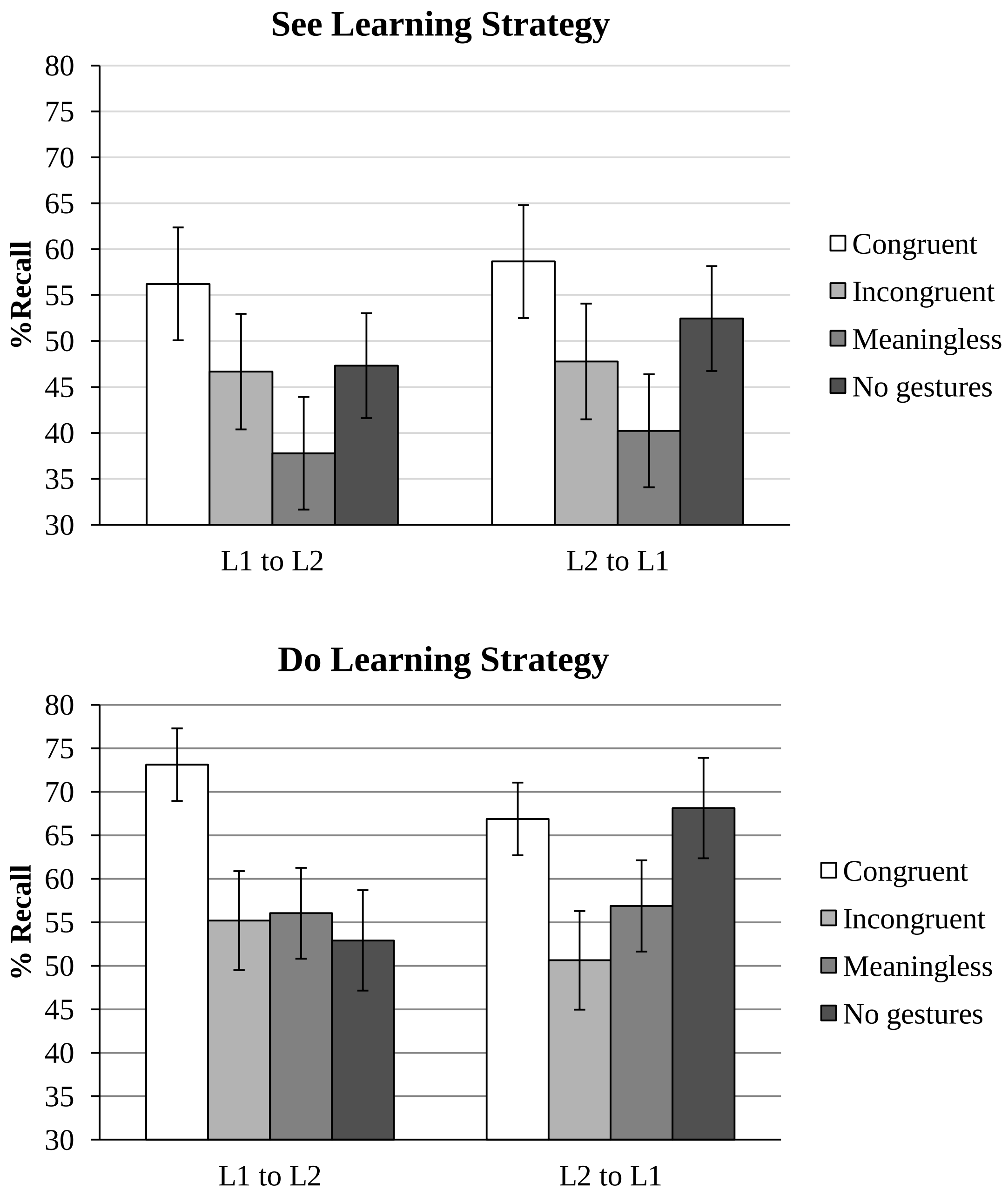


Figure 9. Recall percentage (% Recall) of the "See" (upper graph) and "Do" (bottom graph) learning groups as a function of translation direction (L1 to L2, L2 to L1) and the gesture conditions (congruent, incongruent, meaningless, and no gestures). Standard error is plotted in vertical lines.

"See" learning group. The learning condition (congruent, incongruent, meaningless, and no gestures) and the translation direction (forward, backward) were entered for ANOVA analysis. The main effect of learning condition was significant, $F(3, 42) = 8.97$, $p < .001$, $\eta_p^2 = .39$. The recall percentage was 57.44% (SE = 5.61) in the congruent condition, 47.22% (SE = 5.72) in the incongruent condition, 39.00% (SE = 5.71) in the meaningless, and 49.89% (SE = 5.55) in the no gesture condition.

Compared to the meaningless condition, there was a facilitation effect with congruent gestures so the recall percentage was higher in the congruent condition, $t(14) = 5.08$, $p < .001$. In addition, the difference between the meaningless gesture condition and the no gesture condition was significant, $t(14) = 3.70$, $p = .002$, thus, the processing of a meaningless gesture reduced the recall of L2 words compared with the learning of the words without gestures. Finally, the comparison between the meaningless gesture condition and the incongruent gesture condition was marginally significant, $t(14) = 2.06$, $p = .06$; therefore, there was a trend towards an interference effect due to the presence of meaningless gestures relative to the learning condition without gestures. The remaining main effects or interactions did not reach significance (all *p*s > .05)

"Do" learning group. As in the "see" learning group, the learning condition and the translation direction were subjected to analysis of variance (ANOVA). There was a main effect of learning condition, $F(3, 45) = 12.76$, $p < .001$, $\eta_p^2 = .46$, which was modulated by the Translation Direction x Learning Condition interaction, $F(3, 45) = 13.94$, $p < .001$, $\eta_p^2 = .48$. This interaction was analyzed further by examining the learning condition effect in each translation direction separately.

In the forward translation direction (L1-L2), the learning condition effect was significant, $F(1, 15) = 146.21$, $p < .001$, $\eta^2 = .91$. There was a facilitation effect so participants were more accurate in the congruent gesture condition (73.13%, $SE = 4.18$) relative to the meaningless gesture condition (56.04%, $SE = 5.24$), $t(15) = 6.28$, $p < .001$. No differences were found between the meaningless and the incongruent (55.21%, $SE = 5.68$) conditions, $t(15) = .36$, $p = .73$; thus, the interference effect found in the "see" training group in the incongruent vs. the meaningless gesture condition was not observed when participants produced gestures during the learning phase. Finally, no differences were found between the meaningless and the no gesture conditions (52.92%, $SE = 5.78$), $t(15) = 1.07$, $p =.30$. Therefore, the interference associated to the processing of meaningless gestures vs. no gesture condition obtained in the "see" training group was not observed in the "do" training group.

In the backward translation direction, the differences between learning conditions were significant too, $F(1, 15) = 139.35$, $p < .001$, $\eta^2 = .90$. The mean of recall was 66.87% ($SE = 5.67$) in the congruent condition, 50.63% ($SE = 6.38$) in the incongruent condition, 56.88% ($SE = 5.39$) in the meaningless condition, and 68.13%

(*SE* = 5.00) in the no gesture condition. The difference between the meaningless and the congruent conditions reach significance, $t(15) = 2.61$, $p = .02$, as the comparison between the meaningless and no gesture conditions, $t(15) = 2.84$, $p = .01$. The comparison between meaningless and incongruent gestures was not significant, $t(15) = 1.82$, $p = .09$.

Response times

As in the recall analyses, the Session x Learning Condition x Learning Group x Translation Direction variables were entered in an ANOVA (see Table 4 for the complete set of results and Table 5 for the mean RTs in each treatment).

Table 5. *Response Times in the "See" and "Do" Learning Groups*

	Congruent	Incongruent	Meaningless	No gestures
	"See" Learning Group			
L1-L2 (1st Session)	2546 (181)	3006 (176)	3061 (184)	2802 (213)
L1-L2 (2nd Session)	2464 (156)	2505 (141)	2501 (169)	2536 (162)
L1-L2 (3rd Session)	1961 (142)	2083 (123)	2482 (191)	2140 (140)
L2-L1 (1st Session)	2696 (193)	2783 (131)	2930 (146)	2494 (203)
L2-L1 (2nd Session)	2068 (91)	2207 (155)	2464 (169)	1997 (190)
L2-L1 (3rd Session)	1740 (125)	2075 (156)	1220 (110)	1879 (140)
	"Do" Learning Group			
L1-L2 (1st Session)	2474 (175)	2326 (170)	2543 (178)	1756 (206)
L1-L2 (2nd Session)	2055 (151)	2328 (136)	2278 (164)	2469 (157)
L1-L2 (3rd Session)	1902 (137)	2130 (119)	2069 (185)	2071 (136)
L2-L1 (1st Session)	2262 (187)	2538 (224)	2325 (141)	2030 (196)
L2-L1 (2nd Session)	1720 (88)	2237 (150)	2100 (164)	2116 (184)
L2-L1 (3rd Session)	1729 (121)	2256 (152)	1753 (106)	1835 (136)

Note. Mean RTs (in milliseconds) obtained in the "see" and "do" learning groups as a function of the translation direction (L1 to L2 translation, L2 to L1 translation), the training session (first session, second session, third session) and the learning condition (congruent gestures, incongruent gestures, meaningless gestures, no gestures). Standard errors are in brackets.

The session effect was significant, $F(2, 58) = 34.26$, $p < .001$, $\eta_p^2 = .54$. The mean response time in the first session was 2536 ms (*SE* = 64), 2252 ms (*SE* = 65) in the second session, and 2020 ms (*SE* = 58) in the last session. Linear trend analyses showed significant differences between the first and the third session, $F(1, 29) = 49.53$, $p < .001$, $\eta^2 = .63$. The main effect of translation direction was significant too, $F(1, 29) = 17.71$, $p < .001$, $\eta_p^2 = .38$, with participants responding faster in the backward (2185 ms, *SE* = 55) than in the forward (2354 ms, *SE* = 54) translation direction. The main effect of learning condition was significant, $F(3, 87) = 6.91$, $p < .001$, $\eta_p^2 = .19$. Mean response times were 2135 ms (*SE* = 68) in the congruent condition, 2372 ms (*SE* = 62) in the incongruent condition, 2393 ms (*SE* = 57) in the meaningless condition, and 2177 ms (*SE* = 80) in the no gesture condition. Participants responded significantly faster in the congruent condition compared to the meaningless condition, $t(30) = 3.35$, $p = .002$. The comparison between meaningless and no gesture conditions was marginal, $t(30) = 1.97$, $p = .06$. No differences were observed between the meaningless and the incongruent gesture conditions, $t(30) = .15$, $p = .89$.

The main effect of learning group was significant, $F(1, 29) = 6.72$, $p = .01$, $\eta_p^2 = .19$. Participants in to the “do” learning group responded faster (2137 ms, *SE* = 71) than the “see” learning group (2401 ms, *SE* = 73). Finally, the Session x Learning Group interaction was significant, $F(2, 58) = 5.90$, $p = .005$, $\eta_p^2 = .17$. Participants were significantly faster in the “do” learning group (2282 ms, *SE* = 89) compared to the “see” learning group (2790 ms, *SE* = 92) in the first learning session, $F(1, 29) = 15.86$, $p < .001$, $\eta^2 = .35$; however, the difference between the learning groups was not significant in the second learning session, $F(1, 29) = 1.92$, $p = .18$, $\eta^2 = .06$, nor in the third learning session, $F(1, 29) = .80$, $p = .38$, $\eta^2 = .03$.

DISCUSSION

Movements seem to play a role in many cognitive processes. A facilitative effect has been observed with different types of movements not only in educational settings, but also in clinical context (e.g., developmental disorders, aphasia treatments, etc.) (Botting, Riches, Gaynor, & Morgan, 2010; Hogrefe, Ziegler, Wiesmayer, Weidinger, & Goldenberg, 2013; Kelly et al., 2008). For example, pointing movements, defined as deictic gestures, and beat gestures, that reflect the prosody and emphasize the speech, have showed positive effects on language learning and development (Kushch et al., 2018; Morett, 2014; So et al., 2012). Iconic gestures that make reference to concrete entities or actions are especially remarkable in this context. These gestures have been used in many studies to explore how acts enhance memory consolidation in language production and comprehension (Goldin-Meadow & Alibali, 2013; Straube et al., 2012) and in L2 vocabulary acquisition (So et al., 2012; Tellier, 2008). However, to our knowledge, no previous studies have evaluated the effect of performing vs. observing gestures during the acquisition of L2 vocabulary in adulthood. Thus, in our study, the main interest was the direct comparison between the consequences of self-performed gestures and gestures observation while learning vocabulary in L2. To this end, participants learned L2 verbs accompanied by different types of gestures (congruent, incongruent, meaningless, no gesture) in two different conditions. The "do" learning group was instructed to perform the gestures that accompanied the L2 words while the "see" learning group observed the gestures without performing any movement.

The results found in the study revealed higher recall of L2 words in the "do" learning group (60%) than in the "see" learning group (48%), although the difference between types of training did not reach significance. However, the participants retrieved L2 words more rapidly in the "do" learning group (2137 ms) than in the "see" learning group (2401 ms). Thus, the training based on the self-generation of gestures facilitated the retrieval of vocabulary in a FL.

Concerning the effect of the different gesture conditions used in our study, the results revealed a facilitation effect on L2 learning when the words to be learned were accompanied by congruent gestures compared to the meaningless gesture condition. Specifically, the congruent gesture condition was associated with higher recall percentage and faster response time compared to the meaningless gesture condition. This improvement in L2 learning due to the processing of gestures whose meaning is congruent with that of the words to be learned confirms the results observed in previous research (Feyereisen, 2006; Kushch et al., 2018; Macedonia & Klimesch, 2014; Macedonia & Knösche, 2011). For example, in Macedonia et al.'s (2011) study, a group of German speakers learned Vimmi words presented with iconic gestures (e.g., the word *suitcase* appeared with the gesture of an actor lifting an imaginary suitcase) or meaningless gestures (e.g., the word *suitcase* and the gesture of touching one's own head). The results showed better recall for words learned with iconic gestures relative to words accompanied by meaningless gestures.

The L2 learning facilitation effect due to the processing of congruent gestures was independent of the type of training ("see" learning vs. "do" learning) which suggests that mere exposure of gestures is sufficient to observe the beneficial effect of gestures on the acquisition of vocabulary in a FL. This pattern of results is in line with the outcomes of different studies in which it is demonstrated that mere observation of actions produces a pattern of brain activation similar to that found during the performance of motor actions (motor cortex activation, Stefan et al., 2005). Thus, the processing of gestures, both observing and performing them, might enrich the encoding of the words to be learned by adding sensorimotor networks and procedural memory to the semantic/declarative memory associated with the meaning of the words (Macedonia & Mueller, 2016). Hence, gestures would enhance semantic processing of words.

However, not all gesture conditions improved the learning process. Specifically, in the "see" learning condition, the performance of gestures incongruent with the meaning of L2 words made the acquisition of vocabulary difficult. Previous studies have shown the cost associated with processing incongruent gestures (Feyereisen, 2006; Kelly et al., 2009; Macedonia et al., 2011).

The negative effect of gestures during L2 vocabulary learning might be due to semantic interference produced by the mismatch between the gesture and the meaning of the word to be learned (Bernardis et al., 2008; Yap et al., 2010). Semantic interference with incongruent gesture-word pairs would reflect the difficulty to integrate the meaning of the word and the gesture in working memory (Bernardis et al., 2008, p. 1126). In other words, the lack of correspondence between the information activated by the gesture and the word would lead to a conflict situation in working memory, which would hinder the learning and later recall of L2 words.

On the other hand, the "see" training group showed an additional interference effect, with lower recall of words in the meaningless gesture condition than in the no-gesture condition. The gestures used in the meaningless condition were small movements performed with the hand that did not have iconic or metaphoric associations with the meaning of physical items (for example, the gesture of moving the hand from the forehead to the ear). Thus, the interference found in the meaningless gesture condition would stem from the conflict between motor traces activated by the observation of meaningless movements and the processing of action verbs. In fact, previous studies have shown that both the processing of action verbs and the observation of actions produce activation of the motor cortex similar to the one that would be expected if the participants were performing a movement. Regarding the processing of verbs, Hauk et al. (2004) found that the passive reading of action verbs (e.g., to lick, pick or kick) differentially activated areas along the associated motor strip that overlapped with areas activated by actual movement of the tongue, fingers or feet. Concerning the observation of movements, Buccino et al. (2001) reported that the observation of both object and non-object related actions produced a somatotopically organized activation of premotor cortex similar to that of the classical motor cortex homunculus. Thus, the results found in the "see" training group suggest that the processing of incongruent and meaningless gestures produced interference due to the mismatch between the semantic and motor information associated with these gestures and the words to be learned in L2.

However, the interference effects found in the "see" training group was not obtained when participants performed the gestures during learning; that is, the "do" training group. Hence, we might ask about the cognitive mechanism responsible for minimizing the negative impact of incongruent and meaningless gestures in the "do" training group. The benefit associated to the production of gestures while learning has been reported in previous studies. To illustrate, Cook and Goldin-Meadow (2006) found advantages in solving problems associated to self-performed gestures in children. They argued that the facilitative role attributed to acting while learning would be due to the reduction of working memory load. Thus, in the current study, the performance of gestures during learning would reduce cognitive effort in working memory which would attenuate the conflict produced by the mismatch between the meaning and the motor trace of the gestures and the words in the incongruent and meaningless condition. In fact, it has been broadly confirmed that conflict resolution capacity strongly depends on the availability of resources in working memory (e.g., in bilingual population; Morales, Calvo, & Bialystok, 2013).

Moreover, in the "do" training group, the attenuation of the interference effects was more evident in the forward translation task (L1-L2 translation) than in the backward translation task (L2-L1) (see Figure 9). Thus, interference effects were found in the incongruent and the meaningless condition when participants performed the backward translation task; however, in forward translation, the percentage of recall was similar in the incongruent gesture condition (55.21%), the meaningless gesture condition (56.04%), and the non-gesture condition (52.92%). The main difference between forward and backward translation is task difficulty. In fact, it is well established that L2 learners show worse performance in forward than in backward translation (Kroll & Stewart, 1994). Similarly, in the no gesture condition of our study, the recall of L2 words was lower and the response latencies higher in the L1-L2 translation than in L2-L1 translation; thus confirming the greater difficulty associated with forward vs. backward translation.

The effect of gestures on learning and retrieval of information seems to depend on task difficulty. For example, Marstaller and Burianová (2014) observed that a letter memorization task was more difficult to perform and the recall was lower in participants with low vs. high working memory amplitude. Importantly, the use of gestures facilitated the recall of letters compared to a condition without gestures only in those participants in which the task was more difficult (participants with low working memory amplitude). In our study, the performance of gestures during learning reduced the interference associated with the processing of incongruent and meaningless gestures to a greater extent when the task demands were high (forward translation task). Thus the facilitative effect of performing gestures during learning seems to be more evident when the retrieval task involves high cognitive effort.

To conclude, the results obtained in our study show the beneficial effect of processing gestures in L2 vocabulary acquisition when they are congruent with the words to be learned. This positive effect in learning does not require learners to perform gestures but the effect is shown by mere observation of gestures. However, performing gestures has an additional effect on learning: They seem to mitigate the adverse effect of processing gestures not related to the meaning of the words to be learned. Thus, if one had to choose between different learning programs (with or without gestures), the recommended alternative would a training protocol in which participants produce gestures congruent with words to be learned.

CHAPTER V

The Role of Semantic and Lexical Learning Methodologies in L2 Learning

Experiment 4. Semantic and Lexical Training to Learn Foreign Language Words in Adulthood

We evaluated the role of two L2 learning methods (semantic training and lexical training) in the acquisition of L2 words in adulthood. In the semantic method, L2 words and pictures denoting their meanings were presented and participants learned by practicing a semantic categorization task (to indicate whether new words were exemplars of semantic categories). In the lexical methods, adult individuals received L1-L2 word pairs and they performed a letter identification task (to indicate whether L1-L2 words contained a grapheme). After training, participants decided if L1-L2 word pairs were translation equivalents. When the L2 word was not the L1 translation but was semantically related, interference was found in adult learners subjected to the ssemantic method but not in those that had learned words with a lexical method. This result suggests that semantic training in adulthood favors the acquisition of connections between concepts and L2 words, connections that are usually present in fluent bilinguals.

This study has been reviewed as García-Gámez, A. B., & Macizo, P. (in preparation). Semantic and lexical training to learn foreign-language words in adulthood.

INTRODUCTION

The efficient way of producing a word in a L2 is to retrieve the word in that language from the concept the speaker wants to produce. Therefore, fluent bilinguals are characterized by strong connections between semantic representations and L2 lexical information. This idea was embodied in one of the most relevant views of L2 language acquisition, the RHM (Kroll & Stewart, 1994). In this model, L1 words, L2 words, and a shared semantic system are interconnected. However, differences in the weight of these connections emerge depending on the stage of L2 vocabulary acquisition. In the early stages, the links between the semantic system and the L2 words are weak and L2 learners preferentially use a lexical route of processing from L2 to L1. Furthermore, when proficiency increases, the links between L2 words and the semantic system develop while diminishing the weight of the lexical route.

The RHM has been widely supported by previous work (see Kroll et al., 2015, for a review). For example, unbalanced bilinguals show an asymmetry in translation tasks with faster performance on backward translation (direct lexical route from L2 to L1) than on forward translation (semantic route from L1 to L2) (Kroll & Stewart, 1994). Moreover, strong evidence in favor of increased semantic processing in fluent bilinguals relative to L2 learners comes from studies in which the translation recognition task is used (Talamas et al., 1999). In this task, participants receive a word in one language followed by a word in another language and they have to indicate whether the second word is the translation of the first word. Talamas et al. found that English (L1) speakers who were learners of Spanish (L2) took more time to reject L2-L1 words as correct translation when these words were similar in form (e.g., *man-hambre*; *hambre* means hungry and *hombre* means man in Spanish). In contrast, fluent bilinguals suffered more interference when the word pairs were incorrect translations but were semantically related (e.g., *man-mujer*; *mujer* means woman in Spanish). These results suggest that fluent bilinguals retrieved semantic information when the L2 word was presented.

Therefore, previous research suggests that fluent bilinguals are characterized by making strong connections between the semantic system and L2 words while L2 learners preferentially make use of a lexical route. However, there is evidence suggesting that a semantic route might be functional relatively early in L2 learning (De Groot & Poot, 1997; Frenck-Mestre & Prince, 1997; Potter et al., 1984; see Brysbaert & Duyck, 2010, for a review). To illustrate, Frenck-Mestre and Prince (1997; Experiment 2) examined priming of the dominant and subordinate meanings of biased homographs (e.g., *seal*) in a L2 lexical decision task. The priming effect for the subordinate meaning was only found in proficient bilinguals; however, less and more proficient bilinguals showed priming effects for the dominant meaning of the homograph in L2. Thus, L2 learners used a semantic route of processing (although to a lesser extent than fluent bilinguals).

Studies conducted with the translation recognition task have also found evidence for the use of a semantic route of processing in the early stages of L2 acquisition. In the study conducted by Talamas et al. (1999), described above, there was evidence that less proficient bilinguals processed L2 words semantically. In particular, L2 learners showed semantic interference when the L1-L2 word pairs shared high semantic similarity (while fluent bilinguals showed the effect for both high and low semantic similarity). Moreover, Sunderman and Kroll (2006) obtained evidence of semantic processing in L2 learners with the translation recognition task (see also, Altarriba & Mathis, 1997). In particular, the authors compared two groups of native English (L1) speakers, one with a greater level of proficiency and the other with less proficiency in Spanish (L2). The two groups of participants were slower to reject the pairs that were related in meaning compared with a control condition tested with unrelated word pairs. Thus, even the less proficient L2 learners appeared to be sensitive to conceptual information.

To summarize, previous studies seem to indicate that both the lexical route and the semantic route of processing are present in L2 learners. However, the weight of these routes appears to depend on the fluency of bilinguals in L2; individuals in the early stages of L2 learning would primarily use the lexical route whilst processing would be predominantly semantic in fluent bilinguals. Thus, if we consider cognitive processing of fluent bilinguals, it would be desirable to implement L2 learning methods that favour the establishment of connections between concepts and L2 words. In fact, increased L2 vocabulary learning has been demonstrated with training protocols that foster a semantic route of processing (e.g., presenting a word to learn with a picture denoting its content, see Lotto & De Groot, 1998; Van Hell & Candia-Mahn, 1997; or imagining the meaning of a word to be learned, Ellis & Beaton, 1993; Wang & Thomas, 1995). In the next section, we provide a brief review of the work conducted to address this issue.

Semantic and lexical methods to learn L2 vocabulary

There is previous work addressing the efficiency of semantic and lexical learning strategies to acquire foreign-language vocabulary, the majority of which have been conducted with children (Comesaña et al., 2012; Poarch et al., 2014). In different studies, the semantic and lexical methods of foreign-language vocabulary acquisition have been implemented with the use of pictures and words, respectively. The underlying assumption is that the processing of a picture requires access to its meaning, and thus the learning of new words in L2 accompanied by pictures denoting their meaning would strengthen the connections between the concepts and the L2 lexical system. Tonzar et al., (2009) examined this issue with Italian (L1) children from 4th to 8th grades. The children learned words in FL (English and German) with two training protocols: a word learning method in which the L2 words and L1 word were presented, and a picture learning method consisting of the L2 word and a picture of an object to which the L1 word referred. The authors observed superiority of the picture-mediated learning over the word-mediated learning method. The authors concluded that the efficiency of the picture-based method in the acquisition of L2 words might be due to the fact that this training allowed a direct connection between the L2 lexical representation and the corresponding concept. Comesaña et al. (2009, Experiment 2) found direct evidence for this conclusion by examining semantic interference effects with the translation recognition task. Two groups of Spanish (L1) children learned Basque (L2) words with a word-based method (L1-L2 words were presented) and a picture-based method (L2-picture pairs were used). At testing, the children that had learned the L2 words with pictures showed a semantic interference effect and so were slower and they made more errors to semantically related L2-L1 pairs (e.g., *aulkia*, the word chair in Basque, paired with *mesa*, table in Spanish) relative to unrelated pairs of L2-L1 words. This semantic interference effect suggested that the connection between L2 words and the conceptual system in children was a function of the way in which the L2 words had been learned. A picture-mediated L2 learning instruction favored semantic processing in children learning words in a FL. This conclusion was strongly supported in Poarch et al.'s (2014, Experiment 1) study. The authors evaluated a group of Dutch (L1) children aged 10-11 who were L2 learners of English (L2) in contexts enriched by pictures and listening/speaking exercises. The children showed a semantic interference effect in the translation recognition task, which suggests that they were able to exploit conceptual information during L2 processing and map L2 words to concepts.

All of the studies reviewed so far have examined the impact of semantic vs. lexical mediated instruction of L2 in childhood. In the current study, we evaluated the efficacy of these two learning methods in adult L2 learners. There are reasons to assume that in adults, the type of learning (e.g., with pictures vs. words) might influence L2 vocabulary learning in a way that is different to what occurs in children. Children are individuals still in the process of creating connections between the semantic system and L1 words, so the semantic/lexical systems are likely to be more sensitive to the use of material that favors the link between concepts and L2 words. In contrast, concepts-L1 links are already formed in adults that decide to learn a L2. Thus, adult individuals might use these concept-L1 connections to access new L2 words (i.e., through a lexical route from L1 words to L2 words) regardless of the training program used to acquire foreign-language vocabulary. If fact, Chen and Leung (1989) showed that Chinese (L1) children were better at retrieving words in English (L2) from pictures (e.g., an L2 picture naming task) while Chinese adults learners of English as L2 performed better when retrieving L2 words from L1 words (i.e., L1-L2 translation task). Thus, adults appear to use a lexical route of processing to learn L2 vocabulary, which might reduce the effect of a picture-mediated L2 learning instruction in adulthood.

There are relatively few studies comparing the impact of semantic vs. lexical methods on the acquisition of L2 words in adulthood. Of these, some have addressed the congruency between the learning and the retrieval of foreign-language vocabulary. To illustrate, Chen (1990) asked two groups of Chinese (L1) college students to learn words in French (L2). One group was presented the L2 word paired with a picture depicting its meaning (picture-learning group) while the other received the L2 word with the corresponding L1 translation (word-learning group). At testing, the word-learning group showed better performance on translation tasks with words while the picture-learning groups were faster in naming pictures in L2. Similar picture-learning and word-learning methods were compared by Lotto and De Groot (1998) who asked adult Dutch (L1) speakers to learn Italian words (L2). After learning with pictures-L2 words or L1-L2 words, the learners received either the pictures or the L1 words as cues for the recall of the L2 words. Similar to what was found with children (Tonzar et al., 2009), adult learners showed that the picture-based methods lead to better performance than the word-based methods. Moreover, the results showed evidence of better recall when the study and test conditions were congruent than in the case in which they were incongruent. Therefore, these studies indicate that the performance of adult L2 learners is better when the learning and test conditions match than when they mismatch. However, the aim of these studies was not to determine the role of L2 learning methods (picture-based vs. word-based instruction) on the establishment of a connection between the semantic system and L2 words. The purpose of the present study, therefore, is to address this issue directly.

The current study

In the current study we aimed at exploring whether a picture-mediated L2 learning relative to a word-mediated instruction would foster the creation of connections between concepts and L2 words when adults learn vocabulary in a FL. To this end, adult speakers of Spanish (L1) learned words in Vimmi (L2), an artificial language that has the advantage of controlling for several linguistic variables such as word length, familiarity of L2 words, and phonotactic factors (Jalbert et al., 2011). Two learning methods were implemented, one including pictures-Vimmi words (semantic training) and the other involving Spanish words and Vimmi words (lexical training). Additionally, in order to favor semantic/lexical processing in the picture/word based methods, respectively, participants practiced with two retrieval tasks (see Figure 10). In the semantic L2 learning, participants received a category name along with a Vimmi word that had been previously learned and they were required to decide whether the word was an exemplar of the semantic category. During lexical training, a grapheme was presented followed by a Vimmi word and participants decided whether the grapheme was present in the L1 translation of the L2 word. After training, participants performed a translation recognition task in which they decided if L1-L2 word pairs were translation equivalents. The critical condition was composed of a Vimmi word (e.g., *boreda*, meaning donkey in Vimmi) paired with a semantically related Spanish word (e.g., *caballo*, meaning horse in Spanish). This condition was compared to a control condition with unrelated L1-L2 word pairs. The possible interference effect (performance in the related condition relative to the unrelated condition) was considered to be an index of semantic mediation.

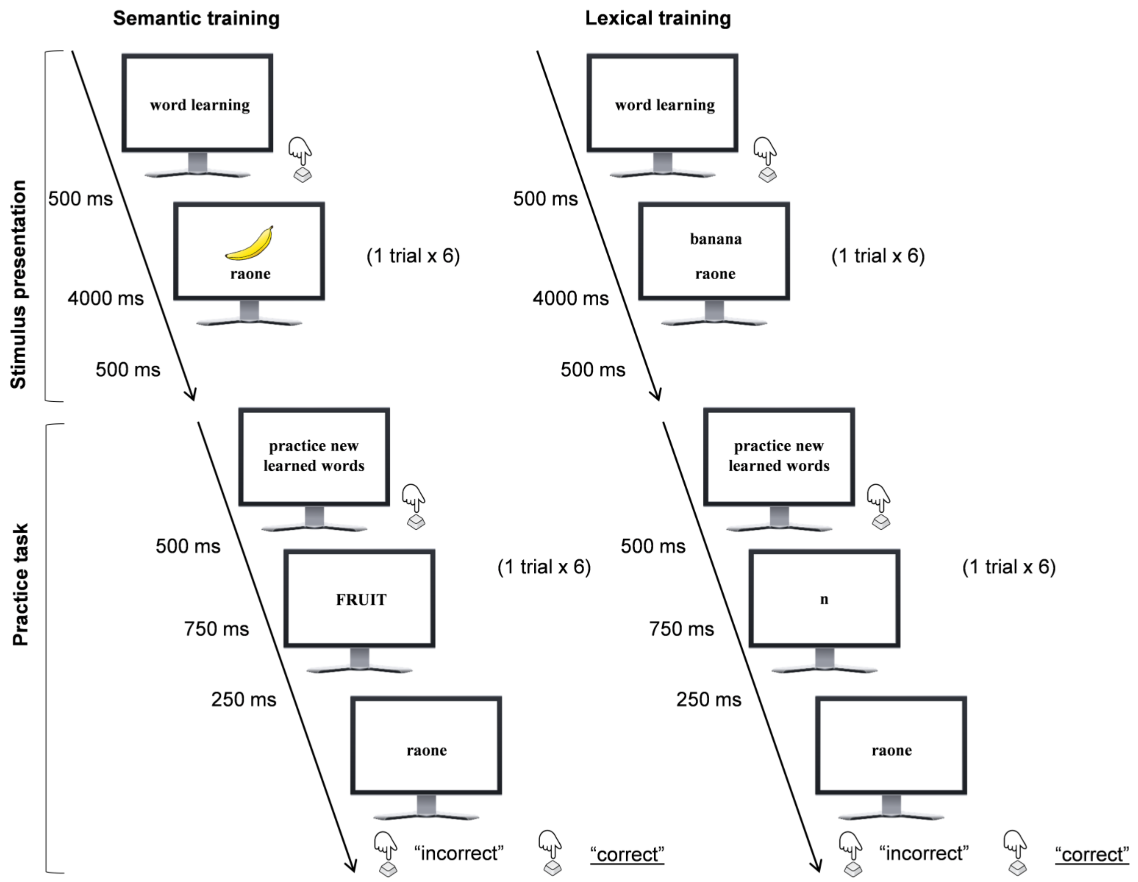


Figure 10. Description of the L2 vocabulary learning methods used in the current study.

Different outcomes were predicted. If adult L2 learners can use pictures to exploit semantic processing of L2 words, they would show an interference effect in the translation recognition task when they were subjected to the semantic training protocol relative to the lexical training protocol. In other words, the picture-based method would foster semantic-L2 connections, as confirmed in children (Comesaña et al., 2009; Poarch et al., 2014). In contrast, if adult individuals primarily use lexical connections to acquire L2 words (e.g., Chen & Leung, 1989), no semantic interference effects would be found irrespective of the method used to acquire Vimmi words.

METHOD

Participants

Thirty students from the University of Granada took part in the study. They gave informed consent before performing the experiment, and their participation was rewarded with academic credits. They reported no history of language disabilities and they had normal or corrected-to-normal visual acuity. All students were native speakers of Spanish. Fifteen participants were randomly assigned to the L2 lexical training condition (13 women, 2 men) and 15 students were assigned to the L2 semantic training condition (13 women, 2 men). The mean age of participants was equated in the L2 lexical training (M = 20.73 years, SD = 4.03), and the L2 semantic training (M = 21.40 years, SD = 2.29), $t(28) = 0.56$, $p = .58$, $d = 0.20$. The number of left/right-handed participants was also equated between the L2 lexical training and the L2 semantic training conditions (2/13 and 1/14, respectively), $X^2(1, N = 30) = 0.37$, $p = .54$.

Design and Materials

In the current study, participants learned a set of 60 Spanish-Vimmi words across ten blocks of training. Half of the participants received lexical training and the remaining participants received the semantic training. Thus, when the training phase is considered, a 2 x 10 mixed design was used with type of L2 training (semantic training, lexical training) as a between-participants factor and block of training (10 levels, from the first block to the last block of training) as a within-participants variable. In the translation recognition task, the Spanish-Vimmi word pairs were correct translations, or they could be an incorrect translation that was semantically related or unrelated. Next, we describe the experimental tasks.

L2 vocabulary learning task. A set of 60 Spanish words were selected. These Spanish words referred to exemplars of six semantic categories (Batting & Montague, 1969; Moreno-Martínez, Montoro, & Rodríguez-Rojo, 2014): three categories were from the living domain (four-footed animals, body-parts, fruits) and three from the non-living domain (kitchen utensils, musical instruments, vehicles). Within each semantic category, ten words were selected. In addition, for each word, a picture was chosen denoting the selected concepts (Pérez & Navalón, 2003; Snodgrass & Vanderwart, 1980). Each Spanish word and its corresponding picture were randomly paired with a Vimmi word. Table 6 displays the lexical characteristics of Spanish words taken from from Cuetos, Glez-Nosti, Barbón, and Brysbaert (2011), and Pérez and Navalón (2003) (length, word frequency, orthographic neighborhood, age of acquisition, familiarity, manipulability, typicality, imageability, and concreteness of words), lexical properties of Vimmi words (length, orthographic neighborhood, and shared graphemes with the Spanish words), and visual properties of pictures taken from Pérez and Navalón (visual complexity of pictures, image agreement, image variability, and picture-name agreement). The complete set of stimuli is presented in Appendix 2.

Table 6. *Characteristics of the Stimuli Used in the Study*

	Mean	SD
Spanish words		
Length	5.68	1.71
Frequency	38.90	61.80
Orthographic neighbourhood	3.36	5.00
AoA	2.47	0.74
Word familiarity	4.16	2.89
Word manipulability	3.36	0.97
Word typicality	4.50	0.47
Word imageability	4.00	2.98
Word concreteness	4.10	2.86
Vimmi words		
Length	5.66	1.45
Orthographic neighbourhood	1.11	2.24
Shared graphemes	2.11	0.97
Shared graphemes (positional)	0.50	0.67

Pictures		
Visual complexity	2.59	0.82
Image agreement	3.94	0.51
Image variability	2.38	0.67
Picture-name agreement	88.2	15.0

Note. Lexical characteristics of words and visual properties of pictures used in the study. *Length:* Number of graphemes of the word. *Frequency:* Spanish word frequency per one million words. *Orthographic neighborhood:* Number of words that can be formed by substituting a single letter at any of the letter position within the string. *Age of Acquisition:* the estimated age at which a word is acquired on a 7-point rating scale (0–2, 3–4, 5–6, 7–8, 9–10, 11–12, and 13 years or more). *Word familiarity:* degree to which the concept denoted by the word is encountered in real life in a 5-point scale (1 = very unfamiliar, 5 = very familiar). *Word manipulability:* degree to which use of the human hand is necessary for the object denoted by the word to perform its function in a 5-point scale (1 = low manipulability, 5 = high manipulability). *Word typicality:* degree to which a concept denoted by a word is a representative exemplar of its category on a 5-point scale (1 = least typical, 5 = most typical). *Word imageability:* how easy it is for a word to arouse mental images on a 7-point scale (1 = low imageability, 7 = high imageability). *Word concreteness:* the degree to which the concept denoted by a word refers to a perceptible entity on a 7-point scale (1 = low concreteness, 1 = high concreteness). *Visual complexity*: of the pictures on a 5-point scale (1 = drawing very simple, 5 = drawing very complex). *Image agreement:* degree to which the picture is similar to the real object depicted in the drawing on a 5-point scale (1 = low agreement, 5 = high agreement). *Image variability:* whether the name of the object evokes few or many different images on a 5-point scale (1 = few images, 5 = many images). *Picture-name agreement:* how close the picture matches the name given to the object in percentage (higher values, higher match).

The 60 Spanish-Vimmi words were randomly grouped into 10 learning sets of 6 words each. On each trial of the semantic training phase, a picture denoting the meaning of the Spanish word (e.g., *plátano*, banana in Spanish) was presented along with its Vimmi translation (e.g., *raone*). After presenting a set of 6 picture-Vimmi word pairs, participants performed a practice task with these Vimmi words: A category name was presented (e.g., *fruit*) followed by a Vimmi word (e.g., *raone*) and participants had to indicate whether or not the word denoted an exemplar of the category previously presented ("yes" in the *plátano-raone* pair). Participants received six trials; on half of the trials the Vimmi word was an exemplar of the semantic category and on the remaining trials it was not. Across the entire task, there were an equal number of "yes" and "no" responses in each of the six semantic categories. In addition, across participants all Vimmi words were assigned to the "yes" and "no" responses.

During lexical training, on each trial a Spanish word and its Vimmi translation were presented, one above the other (e.g., *plátano* and *raone*). After presenting 6 Spanish-Vimmi word pairs, participants completed a practice task with these Vimmi words: A grapheme was presented (e.g., *n*) followed by a Vimmi word that contained it (e.g., *raone*) and participants had to indicate whether this grapheme was also present in the Spanish translation of the Vimmi word ("yes" in the *plátano-raone* pair). Each grapheme was randomly selected from each Vimmi word. On three trials the grapheme was presented in the Spanish translation and on half of the trials it was not. Across the task, half of the target graphemes were vowels and the remaining graphemes were consonants. Across participants all graphemes were assigned to "yes" and "no" responses.

Each participant received either the semantic or lexical training. All participants performed ten blocks with the learning and practice tasks described above. In each block, the 60 Vimmi words to be learned were presented once.

Translation recognition task. A total of 60 trials were presented in the task. Each trial consisted of a Vimmi word followed by a Spanish word. On 20 trials, the Spanish word and Vimmi words were translation equivalents ("yes" responses), on the remaining trials the Spanish-Vimmi word pairs were not translation equivalents ("no" responses). Among the word pairs that were not translation equivalents, 20 trials were composed of words that referred to different semantic categories, one in the living domain and the other in the non-living domain (the unrelated condition, e.g., *plátano-tedo*, banana-cup in Spanish and Vimmi, respectively). In the remaining 20 trials, the word pairs referred to exemplars from the same semantic category (the related condition, e.g., *plátano-deschoga*, banana-pineapple in Spanish and Vimmi, respectively). Each participant received the 60 Vimmi words and 60 Spanish words once (20 translation equivalent word pairs, 20 unrelated word pairs and 20 related word pairs randomly presented). Across participants, each Vimmi word appeared in the translation equivalent condition, the unrelated condition, and the related condition.

Procedure

Participants were tested individually. They first performed a familiarization task, after which the L2 vocabulary-learning task was presented. Finally, they performed the translation recognition task. The complete experiment lasted approximately 2 hours.

In the familiarization phase, the participants received 60 trials randomly presented with a picture and its Spanish name. On each trial, a picture was presented in the centre of the screen with its Spanish name below. The average size of the pictures was 7 x 7 cm. The words were presented in black color (Courier New, bold font, 18 point size). The stimulus remained on the screen until the participants pressed the space bar to continue with the next picture and Spanish words. The participants were instructed to view each picture and its more common name in Spanish and to press the space bar any time they wished to see another picture.

In the L2 vocabulary-learning phase, the task began with the message “word learning”. Afterwards, participants pressed the space bar and, 500 ms later the stimuli to be learned were presented. During semantic training, a picture appeared in the centre of the screen (7 x 7 cm average size) and the Vimmi word below (Courier New, bold font, 18 point size). During lexical training the Spanish word and the Vimmi words were presented one above the other (Courier New, bold font, 18 point size). The stimulus to be learned remained on the screen for 4000 ms and 500 ms later, the next stimulus appeared. After the presentation of 6 stimuli, “practice new learned words” appeared on the screen, the participants pressed the space bar to continue and 500 ms later they performed the practice task. In semantic training, a category name was presented in capital letters for 750 ms, and a Vimmi word appeared in the center of the screen 250 ms later until the participant made a response. Participants pressed the Z and M key to indicate whether or not the Vimmi word denoted an exemplar of the category previously presented. Whether the Z and M keys were assigned to “yes” and “no” responses was counterbalanced across participants. The practice task in lexical learning consisted of the presentation of a grapheme in the center of the screen for 750 ms. Following a 250 ms delay, a Vimmi word was presented until the participants made a response. Participants had to indicate by pressing the Z and M keys whether the Spanish translation of the Vimmi word contained the grapheme previously presented. Whether the Z and M keys were assigned to “yes” and “no” responses was counterbalanced across participants.

The translation recognition task began with the presentation of a fixation point for 1000 ms. After this, a Vimmi word was presented for 250 ms. Following a 250 ms delay, a Spanish word was presented for 500 ms followed by a blank screen for 2000 ms. Participants were instructed to indicate by pressing the Z and M key whether or not the Vimmi word and Spanish word were translation equivalents. Whether the Z and M keys were assigned to "yes" and "no" responses was counterbalanced across participants.

RESULTS

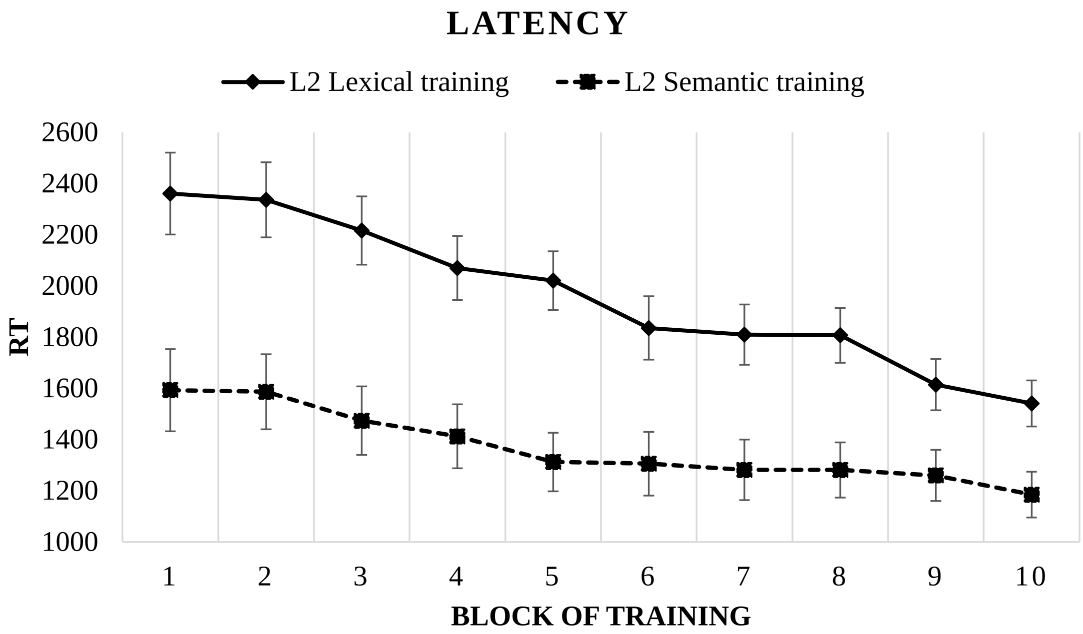
Firstly, we report the results obtained from the L2 lexical training and L2 semantic training tasks, followed by the results found in the translation recognition task.

Performance on the L2 lexical and semantic training tasks

Correct responses on the L2 lexical training and L2 semantic training tasks were 65.41% (*SD* = 19.61) and 79.03% (*SD* = 20.34), respectively. The reaction times (RTs) associated with correct responses on the training tasks were trimmed following the procedure described by Tabachnick, Fidell, and Osterlind (2001) to eliminate univariate outliers. Raw scores were converted to standard scores (z-scores). Data points that after standardization were 3 *SD* outside the normal distribution were considered outliers. After removing outliers from the distribution, z-scores were calculated again. The filter was applied in recursive cycles until no observations fell outside 3 *SD*. The data excluded from the L2 lexical and semantic training tasks was 4.64% and 6.43%, respectively. We conducted an analysis of variance (ANOVA) on RTs and accuracy rates with type of L2 training (semantic training, lexical training) as a between-participant factor and training block (10 levels, from the first block to the last block of training) as a within-participants variable.

The latency analyses revealed a main effect of type of training, $F(1, 28) = 17.03$, $p < .001$, $\eta^2 = .38$. RTs were slower on the semantic training trials (1,961 ms, *SE* = 101) relative to the lexical training trials (1,369 ms, *SE* = 101). The main effect of training block was also significant, $F(9, 252) = 16.31$, $p < .001$, $\eta^2 = .37$. Response times were faster at the end of the training (1,362 ms, *SE* = 63) relative to the beginning of the training (1,977 ms, *SE* = 113). In addition, the interaction between Type of training x Training block was significant, $F(9, 252) = 2.22$, $p = .02$, $\eta^2 = .07$. There were differences across blocks for the lexical training, $F(9, 126) = 9.42$, $p < .001$, $\eta^2 = .40$, with faster responses in Block 10 (1,541 ms, *SE* = 90) relative to Block 1 (2,361 ms, *SE* = 160). The differences between blocks were also significant for the semantic training, $F(9, 126) = 8.66$, $p < .001$, $\eta^2 = .38$. RTs were faster in Block 10 (1,184, *SE* = 90) relative to Block 1 (1,593 ms, *SE* = 160). Thus, the lexical and semantic training speeded up the participants' responses; however, the difference on RTs between Block 1 and Block 10 was larger for the lexical training (820 ms) relative to the semantic training (408 ms) (see Figure 11).

The accuracy analyses revealed a significant main effect of type of training, $F(1, 28) = 3.97$, $p = .05$, $\eta^2 = .12$. Participants committed more correct responses during L2 semantic training (79.06%, *SE* = 4.73), relative to the L2 lexical training (65.74%, *SE* = 4.73). The main effect of training block was also significant, $F(9, 252) = 12.21$, $p < .001$, $\eta^2 = .30$. The percentage of correct responses was higher at the end of the training (76.94%, *SE* = 3.75) relative to the beginning of the training (63.44, *SE* = 2.36). The Type of training x Training block interaction failed to reach significance, $F < 1$. The training block effect was significant for the lexical training, $F(9, 126) = 5.60$, $p < .001$, $\eta^2 = .28$, and the semantic training, $F(9, 126) = 7.34$, $p < .001$, $\eta^2 = .34$. For the lexical training, the difference in the percentage of correct responses between Block 10 (70.13%, *SE* = 5.31) and Block 1 (57.89%, *SE* = 3.33) was 12.24%. For the semantic training, the difference in percentage of correct responses between Block 10 (83.75%, *SE* = 5.31) and Block 1 (69.00%, *SE* = 3.33) was 14.75% (see Figure 11).



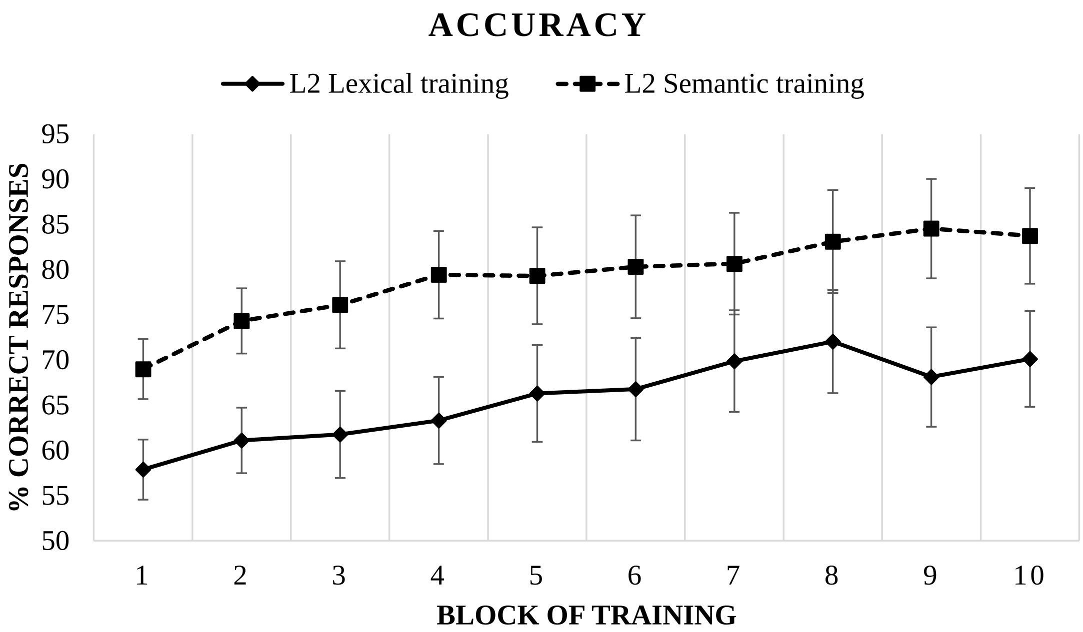


Figure 11. Response times (RT in milliseconds) (upper graph) and percentage of correct responses (lower graph) obtained in the L2 learning method (lexical and semantic) across blocks of training. Error bars represent standard errors.

Translation recognition task

Incorrect responses made by participants given the L2 lexical training (24.11%) and L2 semantic training (17.11%) were excluded from the latency analyses and submitted to the accuracy analyses. RTs associated with correct responses were trimmed as described in the performance on the L2 training. The percentage of outliers was 1.89% for the L2 lexical training and 2.33% for the L2 semantic training.

Firstly, we examined possible differences between participants due to the type of training (lexical and semantic) when they responded to prime-target pairs that were translation equivalents in Spanish and Vimmi ("yes" responses). The latency associated with correct prime-target translations was similar in the lexical training (960 ms, *SE* = 48) and the semantic training (981 ms, *SE* = 48), $F < 1$. In addition, error rates to prime-target translation equivalents did not differ in the lexical training (33.67%, *SE* = 4.53) and the semantic training (22.00%, *SE* = 4.53), $F(1, 28) = 3.31$, $p = .08$, $\eta^2 = .10$.

Next, we considered responses associated with prime-target word pairs that were not translation equivalents ("no" responses). The RTs and error rates to these responses were subjected to an ANOVA with type of L2 training (lexical training or semantic training) and prime-target relation (related or unrelated) as within-participant factors. The type of training was not significant in the latency analysis, $F < 1$, or in the accuracy analyses, $F(1, 28) = 1.11$, $p = .30$, $\eta^2 = .04$. The prime-target relation effect was marginal in the latency analyses, $F(1, 28) = 3.30$, $p = .08$, $\eta^2 = .10$, and significant in the accuracy analyses, $F(1, 28) = 8.82$, $p = .006$, $\eta^2 = .24$. Finally, the Type of training x Prime-target relation was not significant in the RT analyses, $F(1, 28) = 1.82$, $p = .19$, $\eta^2 = .06$, or in the error rates analyses, $F(1, 28) = 1.17$, $p = .29$, $\eta^2 = .04$. Since we were interested in exploring possible differences in the translation recognition task due to the type of training the participants received, we reported the data obtained in each L2 training task separately.

The students in the L2 lexical training condition did not show a prime-target relation effect in the latency analyses, $F < 1$. The RTs were similar in the related condition (946 ms, *SE* = 47) and the unrelated condition (940 ms, *SE* = 49). Similarly, participants given the lexical training did not show differences between error percentages committed in the related condition (21.67%, *SE* = 4.23), and the unrelated condition (17.00%, *SE* = 2.81), $F(1, 28) = 1.78$, $p = .19$, $\eta^2 = .06$. When students in the L2 semantic training were considered, the prime-target relation effect was significant, $F(1, 28) = 5.01$, $p = .03$, $\eta^2 = .15$. Slower responses were found in the related condition (1,021 ms, *SE* = 47) relative to the unrelated condition (979 ms, *SE* = 49). In the accuracy analyses, participants given the semantic training showed a higher percentage of errors in the related condition (19.67%, *SE* = 4.23), relative to the unrelated condition (9.67%, SE = 2.81), $F(1, 28)$ = 8.20, p = .008, η^2 = .22 (see Figure 12).

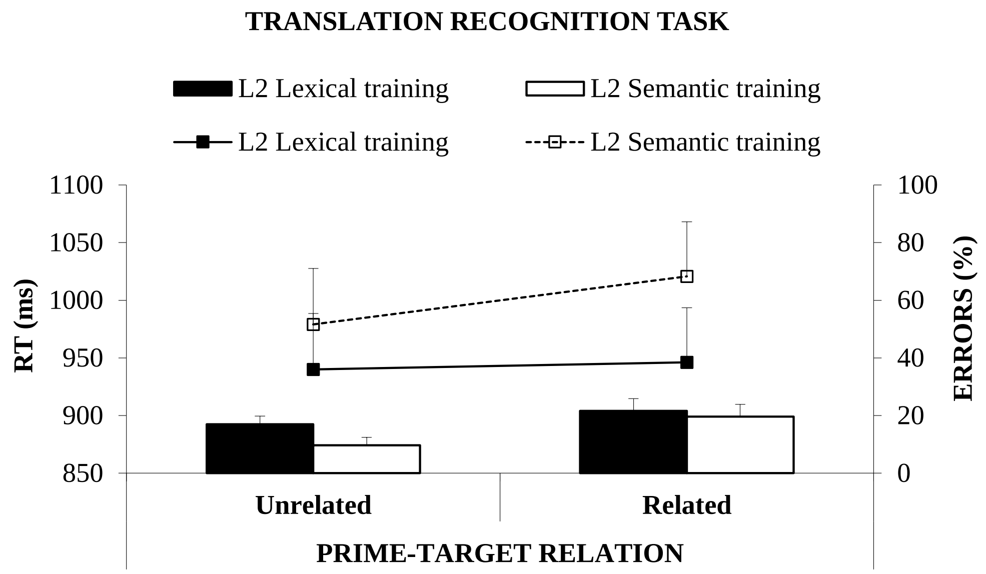


Figure 12. Reaction times (RT in milliseconds, ms) (lines) and error percentages (bars) obtained in the translation recognition task as a function of prime-target relation (related, unrelated) and L2 training method (lexical, semantic). Error bars represent standard errors.

DISCUSSION

The goal of the current study was to evaluate whether adult L2 learners are able to take advantage of a semantic training protocol to foster the connections between concepts and L2 words. Previous research appears to indicate that this might not occur because adult learners make use of lexical connections between their L1 and the new language they intend to acquire (Chen, 1990; Chen and Leung, 1989). However, the results found in the current study revealed that adult individuals are sensitive to a learning method that promotes semantic analyses of the words to be learned in a FL.

Two learning methods were compared in this report: lexical training and semantic training. Both of these methods favored the acquisition of new words in L2. When we considered the performance on the training tasks across blocks of learning, we observed a continuous improvement in L2 acquisition with faster responses and higher accuracy at the end of the training relative to the beginning of the learning task. At first glance, the type of learning would be irrelevant because this learning curve was found regardless of the method used to acquire L2 words. Moreover, after the training phase, the performance on the translation recognition task when the word pairs were translation equivalents was similar in the two learning groups. However, we observed a main effect of the learning method in the training task, which suggested that L2 learning was easier for individuals given the semantic training protocol in comparison with those that received a lexical learning protocol. Thus, overall better performance (faster responses, higher recall of words) was obtained with the use of a semantic learning in comparison with lexical learning.

The critical pattern of results found in this study lies in the outcomes of the translation recognition task after finishing the learning of L2 words. When adult L2 learners had to decide whether Vimmi-Spanish word pairs were correct translations, we found slower responses and reduced accuracy when these word pairs were incorrect translations but were semantically related (e.g., *boreda-caballo*, *boreda* meaning donkey in Vimmi and *caballo* meaning horse in Spanish). However, this semantic interference effect was observed only in learners that received the semantic training protocol.

The semantic interference effect clearly suggests that individuals have formed connections between the semantic systems and the newly acquired words in L2. Thus, when Vimmi words were shown (*boreda* = donkey), L2 learners activated their meaning (e.g., donkey) as well as related concepts (e.g., horse). Therefore, conflict emerged when the subsequent L1 word was incorrect (*caballo* = horse) but semantically related to the correct translation.

As we have described in the introduction section, the semantic interference effect in the translation recognition task is usually observed in expert bilinguals (De Groot & Poot, 1997; Frenck-Mestre & Prince, 1997; Potter et al., 1984; Talamas et al., 1999), so this effect can be regarded as a good marker of what fluent bilinguals do, and in particular it suggests that they make use of direct connections between L2 words and the conceptual system. Therefore, learning protocols that favor the establishment of these connections used by fluent bilinguals would be of interest in early stages of L2 learning. The sematic protocol used in the current study appears to be a good candidate.

Previous studies, conducted primarily with children, have also revealed the benefits associated with semantic mediated methods of L2 learning (Barcroft, 2002; Wimer & Lambert, 1959); such as the use of pictures denoting the meaning of L2 words (e.g., the picture association method, Finkbeiner & Nicol, 2003) or methods based on imagining the meaning of words to be learned in a FL (Ellis & Beaton, 1993; Wang & Thomas, 1995). However, adult learners of L2 seem to be less sensitive to semantic processing when acquiring L2 vocabulary later in life (Chen, 1990; Chen & Leung, 1989). To illustrate, Silverberg and Samuel (2004) evaluated the semantic priming effect across languages of Spanish (L1) – English (L2) bilinguals who acquired their L2 before or after the age of 7 (early and late bilinguals, respectively). In a lexical decision task, early bilinguals showed a semantic priming effect when the L1 word (e.g., *tornillo*, meaning screw in Spanish) was preceded by a semantically related L2 word (e.g., nail). However, late bilinguals did not show this semantic priming effect even when they were as proficient as early bilinguals in L2. Thus, late L2 learners were not affected by semantic variables during the learning of L2 words. However, the current study demonstrates that the lack of sensitivity to semantic processing in late L2 learners can be overcome with the use of a training method that promotes conceptual analyses of the words to be acquired (e.g., the semantic categorization task with new acquired L2 words used in the present study).

One factor that appears to promote adult L2 acquisition is the immersion environment. There is limited but encouraging evidence supporting the idea that studying abroad may enhance the learning process (e.g., Freed, 1995, Genesee, 1987). For instance, Freed, Dewey, Segalowitz, and Halter (2004) compared the acquisition of L2 (French) in a different learning context including a formal language classroom and an intensive summer immersion program. The results revealed that the immersion group made significant gains in oral performance, rate of speech, and speech fluency. Moreover, the learning context seems to determine the links that learners form between L2 words, L1 words, and semantic information. To be more specific, it has been suggested that a formal instruction setting might favor the attachment of new L2 words to previously acquired L1 words (lexical route). In contrast, an immersion environment would foster the establishment of direct connections between L2 words and concepts (semantic route) (e.g., Kroll et al., 1998). However, not all adult learners can attain L2 vocabulary acquisition through an immersion program nor can they enjoy a stay in an L2 speaking country. Nevertheless, the results found with the semantic training paradigm used in this study seems to suggests that even adults learning L2 words in a formal instruction setting can establish connections between L2 lexical forms and the conceptual system.

The results found in our report have implications for evaluating teaching methods used to acquire vocabulary in a L2. Research on effective L2 teaching (e.g., Myles, Hooper, & Mitchel, 1998) has usually focused on process-product studies where the efficiency of instructional methods is measured with quantitative data (e.g., overall scores obtained by the learners in L2 vocabulary tests). However, these general measures might not capture cognitive development associated with L2 vocabulary acquisition. In fact, in our study, individuals given the lexical learning protocol obtained the same overall percentage of responses to correct translations to those learners given the semantic method. However, only the semantic group showed the profile found in expert bilinguals i.e. the semantic interference effect that is indicative of the formation of strong connections between the semantic system and the new words acquired in their L2. Therefore, more fine-grained analyses are required to determine whether the method used to teach L2 vocabulary learning favors the acquisition of the cognitive functioning used by fluent bilinguals.

Experiment 5. The Way in which Foreign Words are Learned Determines Their Use in Isolation and within Sentences

We compared two learning methods for the acquisition of vocabulary in a L2. In addition, the use of the new L2 words was evaluated both in isolation and within sentences. In the semantic method, L2 words and pictures denoting their meanings were presented and participants learned by practicing a semantic categorization task (to indicate whether new words were exemplars of semantic categories). In the lexical methods, individuals received L1-L2 word pairs and they performed a letter-monitoring task (to indicate whether L1-L2 words contained a grapheme). After training, compared with the lexical learning group, the semantic learning group showed: (a) no cost between training sessions, (b) better performance in out-of-context tasks (e.g., picture naming), and (c) an efficient processing of L2 words in sentences. The pattern of outcomes shows the advantage of semantic learning for the acquisition and the use of L2 words in isolation and within sentences.

This study has been accepted for publication as García-Gámez, A. B., & Macizo, P. (2019). The way in which foreign words are learned determines their use in isolation and within sentences. *Journal of Experimental Psychology: Learning, Memory and Cognition.* Advance online publication. doi: 10.1037/xlm0000721

INTRODUCTION

Bilingualism is very common and much more the rule than the exception in most places. It is estimated that around two-thirds of the people in the world are bilinguals (Crystal, 2003). However, it is also true that not everyone speaks a L2. Studies on the acquisition of L2 in childhood and adulthood are, therefore, of particular interest. It would be desirable to implement learning methods that are theoretically based and supported by scientific studies to demonstrate their effectiveness in L2 acquisition.

Early methods of L2 vocabulary acquisition used a word association approach that fostered the establishment of connections between newly acquired L2 words and their translation equivalents in the L1. For example, in the keyword method (Atkinson & Raugh, 1975; Raugh & Atkinson, 1975), a L1 word that sounds like some part of a L2 word is used in the learning process (the keyword). Firstly, learners have to associate the spoken foreign word with the keyword and, afterwards, to associate the keyword with the L1 translation of the word that has to be learned in L2. Previous research has proven the efficacy of these methods for early stages of L2 learning due to the establishment of lexical connections between L1 and L2 (Atkinson, 1975). However, when fluent bilinguals want to express themselves in L2, the best route of processing would be the direct access to L2 words from their concepts. The use of between-language lexical connections and the retrieval of L1 words would be an unnecessary step when bilinguals communicate in L2.

One of the most influential models of L2 vocabulary acquisition reflects these thoughts, the RHM (Kroll & Stewart, 1994). In this model, L2 words are linked to L1 words and concepts. At the beginning of L2 processing, L1 translation equivalents mediate the access to meaning. In contrast, at later stages of L2 development, direct connections between L2 words and concepts become possible. Evidence in favor of increased semantic processing in fluent bilinguals relative to L2 learners comes from studies concerning the processing of words out of context. For instance, in a translation recognition task, L2 learners take more time to reject pairs of L2-L1 words as a correct translation when the L1 word and the correct L1 translation are similar in form (e.g., *man-hambre*; *hambre* means hungry and *hombre* means man in Spanish). On the contrary, fluent bilinguals experience a semantic interference effect with worse performance on L2-L1 word pairs that are incorrect translations but related in meaning (e.g., *man-mujer*; *mujer* means woman in Spanish) than on unrelated word pairs (Talamas et al., 1999). However, empirical evidence demonstrates that L2 word-to-concept links might be available for L2 learners at an earlier point in their L2 acquisition than previously assumed (Sunderman & Kroll, 2006). Thus, it would be relevant to search for learning protocols that favour the early development of connections between concepts and L2 words.

Previous research has confirmed that the acquisition of connections between concepts and L2 words is fostered by the use of training protocols that involve semantic processing (Barcroft, 2002; Comesaña et al., 2012; De Groot & Poot, 1997; Finkbeiner & Nicol, 2003; Kroll et al., 1998; Poarch et al., 2014; Wimer & Lambert, 1959). To illustrate, the presentation of L2 words with pictures denoting their meanings (picture association method, Finkbeiner & Nicol, 2003) favours the learning process relative to the learning of L2 words presented with their translations in L1 (Altarriba & Mathis, 1997; Lotto & De Groot, 1998; Van Hell & Candia-Mahn, 1997).

Evidence about the superiority of semantic vs. lexical L2 learning methods for establishing connections between concepts and L2 words has been gathered in studies concerned with the processing of isolated words (e.g., translation recognition task, translation task, etc.) (Comesaña et al., 2009; Comesaña et al., 2012; Poarch et al., 2014; Tonzar et al., 2009). To illustrate, Poarch et al. (2014, Experiment 1) evaluated a group of Dutch (L1) participants who were L2 learners of English in contexts enriched by pictures and listening/speaking exercises. The learners showed a semantic interference effect in the translation recognition task similar to that found in fluent bilinguals (Talamas et al., 1999), suggesting that they were able to exploit conceptual information during L2 processing and map L2 words to concepts.

To the best of our knowledge, studies comparing semantic vs. lexical learning methods have evaluated L2 acquisition through out-of-context word processing tasks. However, studies conducted with both monolinguals and bilinguals have shown that sentence context modulates single word processing. For example, in electrophysiological studies conducted with monolingual speakers, it is observed that sentence comprehension depends on the lexical properties of the words in the sentence (e.g., lexical frequency, word length, etc.). However, the expectation of an upcoming word within a sentence (cloze probability or semantic constraint) overrides the influence of these lexical factors when the words are highly predicted by the sentence context (Khachatryan et al., 2014; Van Petten et al., 1999).

Studies with bilingual individuals have also found that word processing within sentences may differ from processing words in isolation. Specifically, a large number of studies have shown that bilinguals activate their L1 when they process L2 words. This non-selective activation of words across the bilinguals' languages depends on several factors (such as the type of words to be processed or the fluency of bilinguals in their two languages). However, generally speaking, between-language co-activation has been found across a wide variety of tasks when individuals process out-of-context words (e.g., picture naming, lexical decision, and translation tasks) (Cristoffanini et al., 1986; De Groot & Nas, 1991; Dijkstra et al., 1999; Dijkstra et al., 1998; Dijkstra & Van Heuven, 2002; Lalor & Kirsner, 2001). Moreover, this non-selective activation reflects the use of L1-L2 lexical connections. For instance, words that share lexical overlap across languages (e.g., cognate words such as *tren,* meaning train in Spanish) relative to control words without L1-L2 lexical overlap facilitate performance on single word processing tasks. However, non-selective activation is attenuated when bilinguals process words in a sentence context (see Van Assche et al., 2016; Lauro & Schwartz, 2017, for reviews). For instance, Schwartz and Kroll (2006) considered the processing of cognate words as an index of the use of L2-L1 lexical connections. The authors observed cross-language facilitation when Spanish (L1) – English (L2) bilinguals processed cognate words (e.g., the word *piano*, meaning the same in Spanish and English) in low semantic constraint sentences (e.g., *when we entered the dining hall we saw the piano*…) but not in high semantic constraint sentences (e.g., *before playing, the composer first wiped the key of the piano*…). Thus, these studies suggest that the use of L2-L1 connections is affected by the semantic constraint of sentences because an effect associated to between-language lexical connections (cognate facilitation) is observed in low predictability sentences but not in high predictability sentences (although see, Van Assche et al., 2011; Libben & Titone, 2009; Pivneva et al., 2014; Titone et al., 2011, for cognate facilitation in both low- and high-constraint sentences).

In our study, to examine the use of L2-concepts vs. L2-L1 lexical connections in sentence reading, we evaluated the extent to which L2 vocabulary learners are able to use semantic context to anticipate upcoming L2 words during language comprehension. Previous studies have shown that bilinguals are able to use the context to predict the arrival of new words within a sentence. To illustrate, Foucart et al., (2014) evaluated monolingual and bilingual individuals during the reading of highly constrained sentences that might end with an expected word or not. All participants were sensitive to word expectation, indicating that bilinguals could use semantic context to anticipate upcoming words during reading.

The current study

Previous studies have shown that in learners of a FL, the type of L2 training modulates the establishment of connections between L1 words, L2 words, and the semantic system. On the one hand, training protocols based on lexical processing (e.g., word association method) favour lexical connections between L1 and L2 words. On the other hand, training protocols based on semantic processing (e.g., picture association method) promote the connections between the words in L2 and the semantic system. This dissociation has been confirmed following L2 training, when learners are evaluated in out-of-context tasks (Comesaña et al., 2012; De Groot & Poot, 1997; Finkbeiner & Nicol, 2003).

The aim of the current study was to examine if L2 training based on semantic vs. lexical processing of the material fosters the use of concept-L2 word connections when participants process L2 words in isolation and in linguistically rich contexts (sentences). Although previous studies have shown that fluent bilinguals primarily use the L2-concept connections (Talamas et al., 1999), it may not necessarily be true that adults learning a new language should "skip over" the L2-L1 links. In fact, in fluent bilinguals, L2-L1 connections are functional (e.g., cross-language activation persists for highly proficient bilinguals, Hoshino & Kroll, 2008). Furthermore, recent studies show that fluent bilinguals can learn new vocabulary via the L1 or dominant language by efficiently using regulatory skills that would benefit learning (Bogulski et al., 2018). The objective of this study was to demonstrate that semantic vs. lexical training would maximize the use of direct L2-concept connections.

As far as we know, studies comparing semantic vs. lexical learning methods have only been conducted using out-of-context tasks (Comesaña, Soares, & Lima, 2010; Comesaña et al., 2012; Poarch et al., 2014). Therefore, the current study explored for the first time the consequences of learning L2 words with lexical and semantic training methodologies in sentences. To address this issue, speakers of Spanish (L1) learned words in Vimmi (L2), an artificial language that has the advantage of controlling for several linguistic variables such as word length, familiarity of L2 words, and phonotactic factors (Jalbert et al., 2011). Two learning methods were compared. Participants undergoing semantic training were presented with L2 words and pictures denoting their meaning. In addition, they practiced with a categorization task in which they had to decide whether L2 words were exemplars of a semantic category. On the other hand, participants undergoing the lexical training were presented with L2 words and their L1 translations. Furthermore, they practiced with a grapheme-monitoring task in which they were asked to decide whether a grapheme present in a L2 word was part of the L1 translation (see Figure 13).

In our study, participants learned 60 words on two consecutive days. This two-day design was selected for two reasons. Firstly, to avoid the negative effect that cognitive fatigue has on the evaluation of performance after a single training session (Schwid et al., 2003). Secondly, to evaluate the possible effect that the interval between training and testing can produce in the learning process (e.g., a time delay between training sessions can produce both costs and benefits in the acquisition process, Rickard, 2007).

After training, participants performed several evaluation tasks in which the new words learned in L2 were presented in isolation and in context. The out-of-context tasks were a lexical decision task, a picture- naming task in L1 and L2, and a translation task from L1 to L2 and vice versa. These tasks have been previously used to evaluate single word processing in L2 (e.g., Caramazza & Brones, 1979; Lemhöfer & Dijkstra, 2004). The task that involved L2 words in context was a sentence-processing task in which participants read sentences in L1 that ended with a L2 word that was highly or poorly predicted by the sentence context (see Beatty-Martínez & Dussias, 2017; Moreno et al., 2002; Proverbio, Leoni, & Zani, 2004, for a similar between-language procedure).

Predictions about the processing of L2 words in sentences were grounded in studies on the modulatory role of the context in the use of L1-L2 lexical connections (e.g., Foucart et al., 2014; Schwartz & Kroll, 2006). As mentioned above, according to these studies, fluent bilinguals are able to make use of semantic context to anticipate upcoming words within sentences (e.g., Foucart et al., 2014). In our study, we hypothesized that L2 learners undergoing semantic training would more easily process L2 words in highly predictable sentences than in sentences of low predictability because highly predictable sentences would maximize the use of connections already practiced by these learners during training, that is, the connections between semantic information and L2 words. On the contrary, the predictability effect would be attenuated in the case of L2 learners undergoing lexical training because they would preferably use L1-L2 connections already practiced during the training phase.

METHOD

Participants

Forty students from the University of Granada participated in the study for course credits. They gave informed consent before performing the experiment. The study was approved by the ethical committee at the University of Granada. The participants reported no history of language disabilities and they had normal or corrected-to-normal visual acuity. All students were native speakers of Spanish. In order to control the possible effect that the linguistic experience of the participants in other languages could have on the learning of vocabulary, the inclusion criteria for participating in the study was that the students were monolingual speakers of Spanish, with no or reduced contact and/or formal education in other languages.

To avoid between-groups baseline differences, participants were randomly assigned to the lexical or semantic training. Twenty participants were assigned to the L2 lexical training (13 women, 7 men) and twenty students performed the L2 semantic training (17 women, 3 men). The mean age of the participants was equated in the L2 lexical training (*M* = 21.32 years, *SD* = 2.02), and the L2 semantic training (*M* = 21.10 years, *SD* = 1.84). The number of left/right-handed participants was also equated in the L2 lexical training and the L2 semantic training (2/20 and 2/20, respectively).

No previous measures were collected regarding linguistic ability of participants in L1 (e.g., lexical access). However, in order to ensure that the learners in the semantic and lexical training were matched in L1 skills, different post-test analyses were conducted. The result of these analyses revealed a similar performance of the two groups in word reading, lexical decision and picture naming in L1.

To ensure that the learners in the semantic and lexical training were matched in L1 skills, different post-test analyses were conducted. First, with regard to comprehension processes, we considered the performance of the participants in the lexical decision task with the set of new words in Spanish that were not practiced during the training. There was no difference between the lexical group (*M* = 628 ms, *SD* = 50) and the semantic group (*M* = 611 ms, *SD* = 57) when they processed these L1 words, $t(38) = 1.00$, $p = .32$. Second, the reading times of the first four words (L1 Spanish words) of the sentences used in the sentence-processing task were analyzed. The lexical and semantic group revealed similar reading times for the first word (*M* = 389 ms, *SD* = 82, and *M* = 410 ms, *SD* = 97, respectively), $t(38) = 0.74$, $p = .47$, the second word (*M* = 295 ms, *SD* = 50, and *M* = 295 ms, *SD* = 67, respectively), $t(38) = 0.00$, $p = .99$, the third word (*M* = 276 ms, *SD* = 51, and *M* = 290 ms, *SD* = 66, respectively), $t(38) = 0.75$, $p = .46$, and the fourth word (*M* = 288 ms, *SD* = 51, and *M* = 310 ms, *SD* = 71, respectively), $t(38) = 1.14$, $p = .26$. Third, regarding language production in L1, the picture naming task did not reveal between-groups differences when participants named the pictures in L1 ($F < 1$). Taken together, these analyses suggest that participants were matched in L1 linguistic skills.

Design and Materials

In the current study, participants learned a set of 60 Spanish (L1) - Vimmi (L2) words. The Vimmi words were Spanish pseudowords with legal orthography and phonology in accordance with the constraints of the Spanish language. The L2 vocabulary learning involved two training sessions conducted on two consecutive days.

In addition, a set of tasks was selected to evaluate the learning of L2 words, a lexical decision task, a picture-naming task, a translation task and sentence processing task.

L2 vocabulary learning tasks. Two L2 training methods were used. Half of the participants were subjected to a lexical training schedule and the remaining half of participants performed the semantic training. When the L2 training phase was considered, a 2 x 20 mixed design was used with type of L2 training (semantic training, lexical training) as a between-participants factor and block of training (20 levels, from the first block to the last block of training) as a within-participants variable. In order to control several linguistic variables, Spanish words from six semantic categories were chosen (Battig & Montague, 1969; Moreno-Martínez et al., 2014) three categories were from the living domain (four-footed animals, body-parts, fruits) and three from the non-living domain (kitchen utensils, musical instruments, vehicles). Within each semantic category, ten words were selected. Hence, the total amount of words to be learned was 60. In addition, for each word, a picture denoting its concept was selected (Pérez & Navalón, 2003; Snodgrass & Vanderwart, 1980). Each Spanish word and its corresponding picture were randomly paired with a Vimmi word. Table 7 lists the lexical characteristics of Spanish words taken from Cuetos et al. (2011), and Pérez and Navalón (2003) (length, word frequency, orthographic neighbourhood, age of acquisition, familiarity, manipulability, typicality, imageability, and concreteness of words), lexical properties of Vimmi words (length, orthographic neighbourhood, and shared graphemes with the Spanish words) and visual properties of pictures taken from Pérez and Navalón (visual complexity of pictures, image agreement, image variability, and picture-name agreement). The complete set of stimuli is presented in Appendix 2.

Table 7. *Characteristics of the Stimuli Used in the Study*

	Mean	SD
Spanish words		
Length	5.68	1.71
Frequency	38.90	61.80
Orthographic neighbourhood	3.36	5.00
AoA	2.47	0.74

Word familiarity	4.16	2.89
Word manipulability	3.36	0.97
Word typicality	4.50	0.47
Word imageability	4.00	2.98
Word concreteness	4.10	2.86
Vimmi words		
Length	5.66	1.45
Orthographic neighbourhood	1.11	2.24
Shared graphemes	2.11	0.97
Shared graphemes (positional)	0.50	0.67
Pictures		
Visual complexity	2.59	0.82
Image agreement	3.94	0.51
Image variability	2.38	0.67
Picture-name agreement	88.2	15.0

Note. Lexical characteristics of words and visual properties of pictures used in the study. Length: Number of graphemes of the word. Frequency: Spanish word frequency per one million words. Orthographic neighbourhood: Number of words that can be formed by substituting a single letter at any of the letter positions within the string. Age of Acquisition: the estimated age at which a word is acquired on a 7-point rating scale (0–2, 3–4, 5–6, 7–8, 9–10, 11–12, and 13 years or more). Word familiarity: the degree to which the concept denoted by the word is encountered in real life on a 5-point scale (1 = very unfamiliar, 5 = very familiar). Word manipulability: degree to which use of the human hand is necessary for the object denoted by the word to perform its function on a 5-point scale (1 = low manipulability, 5 = high manipulability). Word typicality: degree to which a concept denoted by a word is a representative exemplar of its category on a 5-point scale (1 = least typical, 5 = most typical). Word imageability: how easy it is for a word to arouse mental images on a 7-point scale (1 = low imageability, 7 = high imageability). Word concreteness: the degree to which the concept denoted by a word refers to a perceptible entity on a 7-point scale (1 = low concreteness, 1 = high concreteness). Visual complexity: of the pictures on a 5-point scale (1 = drawing very simple, 5 = drawing very complex). Image agreement: the degree to which the picture is similar to the real object depicted in the drawing on a 5-point scale (1 = low agreement, 5 = high agreement). Image variability: whether the name of the object evokes few or many different images on a 5-point scale (1 = few images, 5 = many images). Picture-name agreement: how close the picture matches the name given to the object in percentages (higher values, higher match).

The 60 Spanish-Vimmi words were randomly grouped into 10 learning sets of 6 words. On each trial of semantic training, a picture denoting the meaning of the Spanish word (e.g., *plátano*, banana in Spanish) and its Vimmi translation were presented on the screen (e.g., *raone*). After presenting a set of 6 picture-Vimmi words, participants performed a practice task with these Vimmi words. A category name was presented (e.g., *fruit*) followed by a Vimmi word (e.g., *raone*) and participants had to indicate whether or not the word denoted an exemplar of the category previously presented ("yes" in the *plátano-raone* pair). Within each set, participants received six trials; on half of the trials the Vimmi word was an exemplar of the semantic category and on the remaining trials it was not. Across the entire task, there were the same number of "yes" and "no" responses in each of the six semantic categories. In addition, across participants, all Vimmi words were assigned to the "yes" and "no" responses.

During lexical training, for each trial, a Spanish word and its Vimmi translation were presented on the screen (e.g., *plátano* and *raone*). After presenting 6 Spanish-Vimmi word pairs, participants completed a practice task with these Vimmi words. A grapheme was presented (e.g., *n*) followed by a Vimmi word that contained it (e.g., *raone*) and participants had to indicate whether this grapheme was also present in the Spanish translation of the Vimmi word ("yes" in the plátano-raone pair). Each grapheme was one letter randomly selected from each Vimmi word. On half of the trials, this grapheme was presented in the Spanish translation and on the remaining half of trials it was not. Across the task, half of the letters were vowels and the remaining graphemes were consonants. Across participants, all graphemes were assigned to "yes" and "no" responses.

Participants were randomly divided into two training groups so they only learned with the semantic or the lexical training. All participants performed twenty blocks with the learning and practice tasks described above. In each block, the 60 Vimmi words to be learned were presented once. A previous study conducted in our lab (García-Gámez & Macizo, 2019) revealed that 12 exposures to 40 Spanish-Vimmi word pairs produced 64% mean accuracy rate at the end of the L2 vocabulary learning phase. Although in this study we used a different methodology, the number of words to learn was higher (60 L2 words). Thus, with the aim of increasing the final learning rate, in the current study, the training involved a total of 20 exposures to each L2 word and the total amount of trials was 1200. The L2 learning task took place over two consecutive days: Blocks from 1 to 10 were administered on the first learning session (day 1) and blocks from 11 to 20 on the second learning session (day 2).

In the L2 vocabulary learning phase, the task started with the message "word learning". Afterwards, participants pressed the space bar and 500 ms later, the stimuli to be learned were presented. In the semantic training, a picture appeared in the middle of the screen (7 x 7 cm average size) and the Vimmi word below (Courier New, bold font, 18 point size). In the lexical training, the Spanish and the Vimmi words were presented one above the other (Courier New, bold font, and 18 point size). The stimulus to be learned remained on the screen for 4000 ms and 500 ms later, the next stimulus appeared. After the presentation of 6 stimuli, the message "practice new learned words" appeared on the screen, participants pressed the space bar to continue and 500 ms later, they performed the practice task. During semantic training, a category name was presented in capital letters for 750 ms. After 250 ms, a Vimmi word then appeared in the middle of the screen until the participant responded. Participants pressed the Z and M key to indicate whether or not the Vimmi word denoted an exemplar of the category previously presented. The practice task in lexical learning consisted of the presentation of a grapheme in the middle of the screen for 750 ms. Following a 250 ms delay, a Vimmi word was presented until the participant responded. Participants had to indicate by pressing the Z and M keys whether the Spanish translation of the Vimmi word contained the grapheme previously presented. In the training tasks (semantic learning and lexical learning), the assignment of Z and M keys to "yes" and "no" responses was counterbalanced across participants. However, a participant always received the same assignment of response keys for all tasks requiring response decision (training task and lexical decision task). In the L2 vocabulary learning phase, the participants did not receive feedback on their responses (see Figure 13).

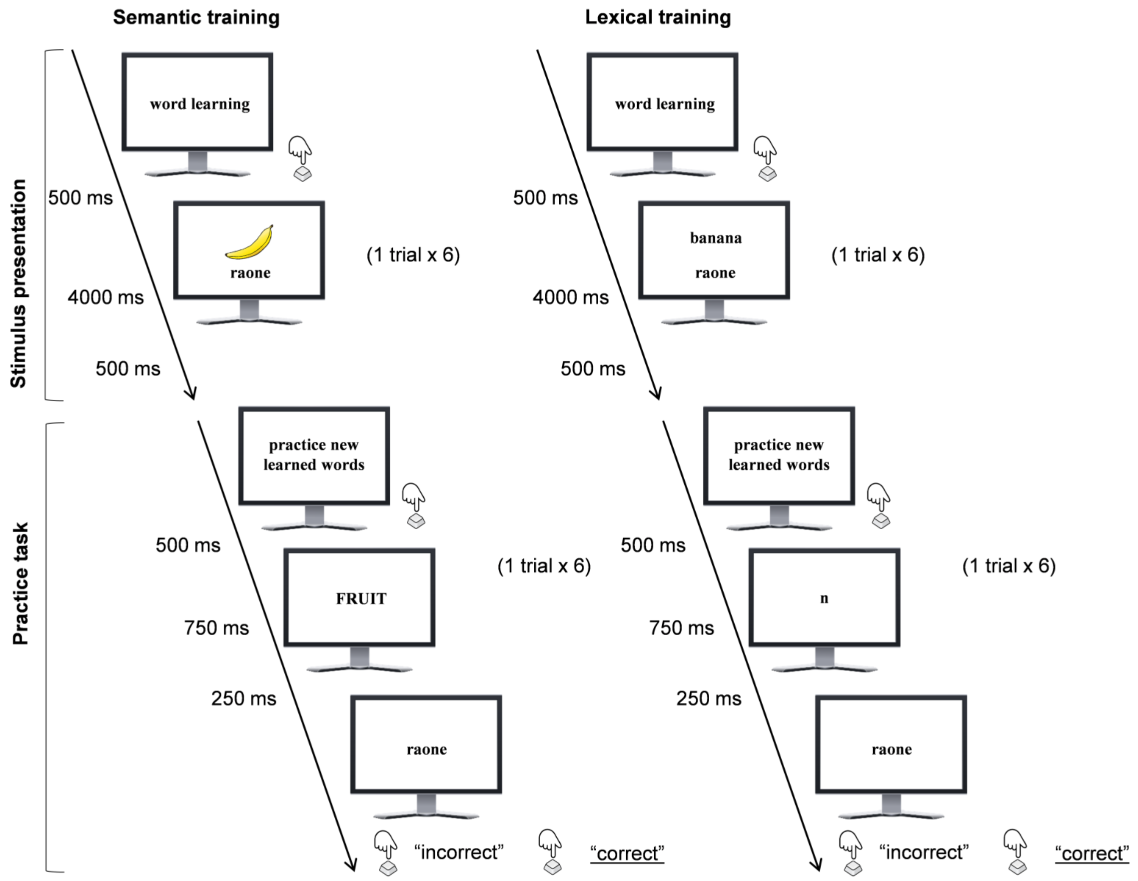


Figure 13. Description of the L2 (second language) vocabulary learning methods used in the current study.

Lexical decision task. In this task participants received stimuli (sequence of graphemes) and they had to indicate whether they were legal words or pseudo-words regardless of the language (Spanish or Vimmi). On each trial, a fixation point was presented for 300 ms, followed by a stimulus (word or pseudoword, Courier New, 18 point size, black font, white background), which remained on the screen until the participant's response. The interval between trials was 1000 ms. Response keys were M and Z on the keyboard. The assignment of the M/Z key to word/pseudoword was counterbalanced across participants. The main goal of the lexical decision task was to study how participants processed the trained words. Thus, the set of stimuli was composed of the 60 trained Spanish words (e.g., *cuchara*) and 60 trained Vimmi words (e.g., *lofuse*). In addition, other conditions were introduced to control different factors that could influence the performance of the task. With the aim that the "yes" responses (word trials) did not always correspond to previously practiced words, a new set of 60 untrained Spanish words (e.g., *muñeca)* was selected from the same semantic categories used to select the words in the training task. The set of trained and untrained Spanish words were equated in terms of lexical variables (Davis & Perea, 2005), frequency (trained words, *M* = 33.22, *SD* = 69.09; untrained words, *M* = 21.29, *SD*= 41.99, *t*(118) = 1.14, *p* = .26), length (trained words, *M* = 5.68, *SD* = 1.71, untrained words, *M* = 5.95, *SD*= 1.56, *t*(118) = .89, *p* = .37), orthographic neighbourhood (trained words, *M* = 4.07, *SD* = 5.38, untrained words, *M* = 4.02, *SD* = 5.16, *t*(118) = .05, *p* = .96), familiarity (trained words, *M* = 4.58, *SD* = 2.71, untrained words, *M* = 3.69, *SD* = 2.89, *t*(118) = 1.74, *p* = .08), imageability (trained words, *M* = 4.43, *SD* = 2.83, untrained words, *M* = 3.79, *SD* = 2.94, *t*(118) = 1.21, *p* = .23), and concreteness (trained words, *M* = 4.51, *SD* = 2.67, untrained words, *M* = 3.79, *SD* = 2.94, *t*(118) = 1.40, *p* = .16). A set of 60 untrained Vimmi words was also selected (e.g., *wefino*) and equated with the trained Vimmi words in terms of the following lexical factors: length (trained words, *M* = 5.67, *SD* =1.46, untrained words, *M* = 5.57, *SD* = .98, *t*(118) = .44, *p* = .66) and orthographic neighbourhood (trained words, *M* = 1.12, *SD* = 2.24, untrained words, *M* = 1.02, *SD* = 2.70, *t*(118) = .22, *p* = .82). These new Vimmi words functioned as pseudowords for the participants as they were untrained words coming from an artificial language. Furthermore, with the aim that not all the pseudowords came from the Vimmi language, an additional set of 60 pseudowords derived from Spanish was included. Therefore, the lexical decision task included 60% of "yes" trials and 40% of "no" trials. Participants performed a total of 300 trials and the stimuli were presented in random order.

Translation task. In this task the 60 Spanish-Vimmi word pairs learned in the L2 vocabulary-training phase were used. The 60 Spanish words were presented for translation into Vimmi (forward translation) and the 60 Vimmi words were presented for translation into Spanish (backward translation). The participants were not required to pronounce the Vimmi words before the translation task. The order in which the two translation tasks were presented was counterbalanced across participants. In addition, the words within each translation task were presented in random order. On each trial, a fixation point appeared for 1000 ms in the middle of the screen followed by the word to be translated for 500 ms (Arial, 30 point size, black font, white background). A white screen then remained until the participant's response. Participants were required to say aloud the translation of each word. Response latencies were collected using a microphone ATR 20 with low impedance connected to a PST serial Response Box (Schneider, 1995) and tape-recorded to eliminate trials with incorrect responses.

In this task, only correct responses were included in the analyses of the reaction times (RTs). Data points were excluded from the RT analyses if: (a) the participants produced nonverbal sounds that triggered the voice key, (b) the participants stuttered or hesitated in producing the word, (c) the participants produced something different than the required word. Some minor errors were allowed and considered correct responses depending on the length of the correct L2 word to be produced: (1) For monosyllabic words, the replacement of a vowel, 2) for disyllabic words, the replacement of a vowel or a consonant but not both, 3) for words with three or more syllables, the inversion of a vowel and a consonant or the replacement of a vowel or a consonant.

Naming task. Participants were presented with 60 pictures denoting the meaning of the words learned in L2. These pictures were selected from the Snodgrass and Vanderwart (1980) database. Participants performed the naming task in Spanish and in Vimmi. The order in which they performed these naming tasks was counterbalanced across participants. Within each naming task, the pictures were presented in random order. On each trial, a fixation point was presented for 1000 ms, after which the picture to be named (14 x 14 cm average size) appeared in the middle of the screen, followed by a white screen that remained until the participant's response. Oral responses were tape-recorded to eliminate trials with incorrect responses. The criteria for coding the correctness of the participants' oral responses were the same as those described in the translation task.

Sentence-processing task. In this task, participants received a set of sentences and they had to indicate whether or not they were semantically plausible. The sentences were presented with a mixed-language procedure, in which they appeared in Spanish (the participants L1), except for the last word of the sentence, which appeared in Vimmi (Beatty-Martínez & Dussias, 2017; Moreno et al., 2002; Proverbio et al., 2004, for a similar between-language procedure).

On each trial, a fixation point appeared until the participants pressed the space bar to receive the sentence. The sentences were presented word-by-word (Courier New, 18 point size, black font, white background) in the middle of the computer screen (Just, Carpenter, & Woolley, 1982). Participants read the sentences at their own pace by pressing the space bar every time they wanted to see the next word. After the participants had finished reading a sentence, the message "*¿Tiene sentido?*" (Does it make sense?) appeared on the screen and they had to indicate whether or not the sentence was semantically plausible by pressing the M and Z keys on the keyboard. The assignment of the M/Z keys to yes/no responses was counterbalanced across participants. The next sentence appeared after 1000 ms. The reading time to the last word of the sentence (Vimmi word) was considered in the latency analysis, while the response to the prompt "tiene sentido" was used to collect accuracy data.

The 60 Vimmi words learned in the L2 vocabulary-training phase were used as the last word of the sentences. For each Vimmi word, two sentence contexts were created. These contexts varied depending on the cloze probability (high or low) of the last word (the degree to which the last word could be predicted by the sentence context) (Altarriba, Kroll, Sholl, & Rayner, 1996; Green & Wei, 2014, 2016; Gunter, Friederici, & Schriefers, 2000; Moreno et al., 2002; Staub, Grant, Astheimer, & Cohen, 2015). For example, for the Vimmi word *ruzanego* (knife in English) the highly predictable sentence context was "A Pedro le costó mucho cortar el filete porque no estaba bien afilado el…" ("Pedro had a hard time cutting the steak because it was not sharp enough the…", note that the syntactic structure of all sentences were correct in Spanish); and the sentence context of low predictability was "Cuando llegaron al río, Pedro miró en su mochila y no encontró el…" ("When they arrived at the river, Pedro looked in his backpack and did not find the…").

The predictability of the sentences was determined in a pilot study in which a cloze probability task was administered to 40 participants that did not participate in the main experiment. Predictability is defined as the probability that a word is selected as a sentence continuation in a cloze task (Taylor, 1953) (see Staub et al., 2015 for a review). Thus, the participants performed a sentence completion task in which they received a sentence context in Spanish without the last word of the sentence and they were required to complete the sentence with the first word that came to their mind. Afterwards, we computed the percentage of participants that chose the trained word to complete the sentence. The cloze probability of final words of low predictability was 8.41% (*SE* = 12.77) whilst this was 90.67% (*SE* = 12.84) for highly predictable final words, $F(1, 59) = 1158.53$, $p < .001$, $\eta^2 = .95$. Thus, following several criteria, highly predictable sentence contexts were associated with high cloze probability (67%-100%, Block & Baldwin, 2010; 90% or above, Bloom & Fischler, 1980; 40% or higher, Coulson, Urbach, & Kutas, 2006).

Once the sentence contexts of high and low predictability were created for plausible sentences, these sentence contexts were recombined with the 60 Vimmi words to create sentences of high and low predictability, and implausible sentences. The implausible sentences were syntactically correct. In addition, the last word of the sentence was grammatically correct but it was incongruent with the meaning of the sentence context. For example, a high/low predictability implausible sentence was "Pedro had a hard time cutting the steak because it was not sharp enough the …/"When they arrived at the river, Pedro looked in his backpack and did not find the…" finished with the word "*boreda*" meaning donkey in Vimmi.

In the sentence-processing task, each participant received 60 sentences, 15 in the highly predictable plausible condition, 15 in the low predictability plausible condition, 15 in the highly predictable implausible condition, and 15 in the low predictability implausible condition. No participant ever received a Vimmi word or a sentence context twice but across participants the 60 Vimmi words were assigned to the four experimental conditions. Participants received the set of 60 sentences in random order. The complete set of stimuli used in this task is provided in Appendix 3.

Procedure

The study was conducted over two consecutive days, with participants being tested individually. E-prime experimental software was used for stimulus presentation and data acquisition (Schneider et al., 2002). Firstly, participants performed a familiarization task where all the pictures used in the experiment were presented along with their Spanish names. In the familiarization phase, participants received 60 trials randomly presented with a picture and its Spanish name. On each trial, a picture was presented in the middle of the screen with its Spanish name below. The stimulus remained on the screen until the participants pressed the space bar to continue with the next picture and Spanish word. The participants were instructed to see each picture and its more common name in Spanish and to press the space bar any time they wanted to see another picture.

The familiarization task was presented at the beginning of the first training session. After finishing this phase, the L2 vocabulary-learning task was introduced (semantic or lexical training). After 24 hours, participants received a second training session. After completing the L2 acquisition phase, participants took a 15-minute break and they continued with the evaluation tasks in the following order: sentence processing, lexical decision, translation, and naming). The first session lasted around 90 minutes and the second session approximately 150 minutes.

RESULTS

Firstly, we report the results obtained from the L2 lexical and semantic training phases. Then, we continue with the results obtained in the four evaluation tasks. Participants with less than 10% correct answers in any of the experimental tasks were excluded from the analysis. Following this exclusion criteria, three participants were removed in the naming task (one participant in the semantic training and two participants in the lexical training) and four participants were removed in the translation task (two participants in the semantic training and two participants in the lexical training).

Signal detection measures comparing hit rates and false alarm rates (discriminability, d') were also used to evaluate the type of training (lexical training, semantic training) in all two-alternative force choice tasks. In the training task, there was a tendency towards greater discriminability in the semantic group (d' = 2.61, SD = 1.18) compared to the lexical group (d' = 1.91, SD = 1.10), $t(38) = 1.97$, $p = .06$. This result was in line with that obtained in the accuracy analyses with more correct responses during L2 semantic training relative to L2 lexical training. In the lexical decision task, discriminability values were similar in the semantic group (d' = 3.15, SD =

.79) and the lexical group (d' = 3.31, *SD* = .80), $t(38) = 0.65$, $p = .52$. This result was also in line with that found in the accuracy analysis where the type of training effect was not significant. Finally, in the sentence processing task, discriminability values did not differ in the semantic training (d' = 2.99, *SD* = 1.25) and the lexical training (d' = 2.45, *SD* = 1.49), $t(38) = 1.23$, $p = .22$. Now, the results obtained in each task are presented separately.

Performance on the L2 lexical and semantic training phases. Correct responses in the L2 lexical training and L2 semantic training phases were 80.43% (*SE* = 3.09) and 86.52% (*SE* = 2.94), respectively. The reaction times (RTs) associated with correct responses in the training tasks were trimmed following the procedure described by Tabachnick and Fidell (2001) to eliminate univariate outliers. Raw scores were converted to standard scores (*z*-scores). Following standardization, any data points that were 3 *SD* outside the normal distribution were considered outliers. After removing outliers from the distribution, *z*-scores were calculated again. The filter was applied in recursive cycles until no observations were outside 3 *SD*. The percentages of data excluded from the L2 lexical and semantic training phases were 6.58% and 10.57%, respectively. We conducted an analysis of variance (ANOVA) on RTs and accuracy rates with type of L2 training (semantic training, lexical training) as a between-participants factor and block of training (20 levels, from the first block to the last block of training) as a within-participants variable. When Greenhouse-Geisser correction (Greenhouse & Geisser, 1959) for nonsphericity of variance was used in the analyses, the pattern of results was the same as that reported in text.

The latency analyses revealed a significant effect of the type of training, $F(1, 36) = 6.69$, $p = .013$, $\eta^2 = .16$. RTs were faster during semantic training (1064 ms, *SE* = 58.15) in comparison with lexical training (1282 ms, *SE* = 61.29). The main effect of block of training was significant, $F(19, 684) = 58.45$, $p < .001$, $\eta^2 = .61$. In addition, the Type of training x Block of training interaction was significant, $F(19, 684) = 10.24$, $p < .001$, $\eta^2 = .22$.

In the lexical training phase, linear trend analysis was significant, $F(19, 323) = 32.51$, $p < .001$, $\eta^2 = .65$, with faster responses at the end of training (1016 ms, *SE* = 90.04) relative to the beginning of training (2134 ms, *SE* = 92.08). Linear trend analysis was also significant during semantic training, $F(19, 361) = 29.71$, $p < .001$, $\eta^2 = .61$, with RTs being faster at the end of training (889 ms, *SE* = 36.32) relative to the beginning of the training (1412 ms, *SE* = 31.48). Thus, the lexical and semantic training phases speeded up the participants' responses; however, the difference in RT between the first and the last block was larger in the lexical training phase (1118 ms) compared with semantic training (523 ms) (see Figure 14). Furthermore, there were differences between the semantic training and the lexical training phases on the first block of the learning phase (Block 1), $F(1, 38) = 66.78$, $p < .001$, $\eta^2 = .63$; however, no differences were found between them at the end of training (block 20), $F(1, 38) = 3.57$, $p = .07$, $\eta^2 = .09$.

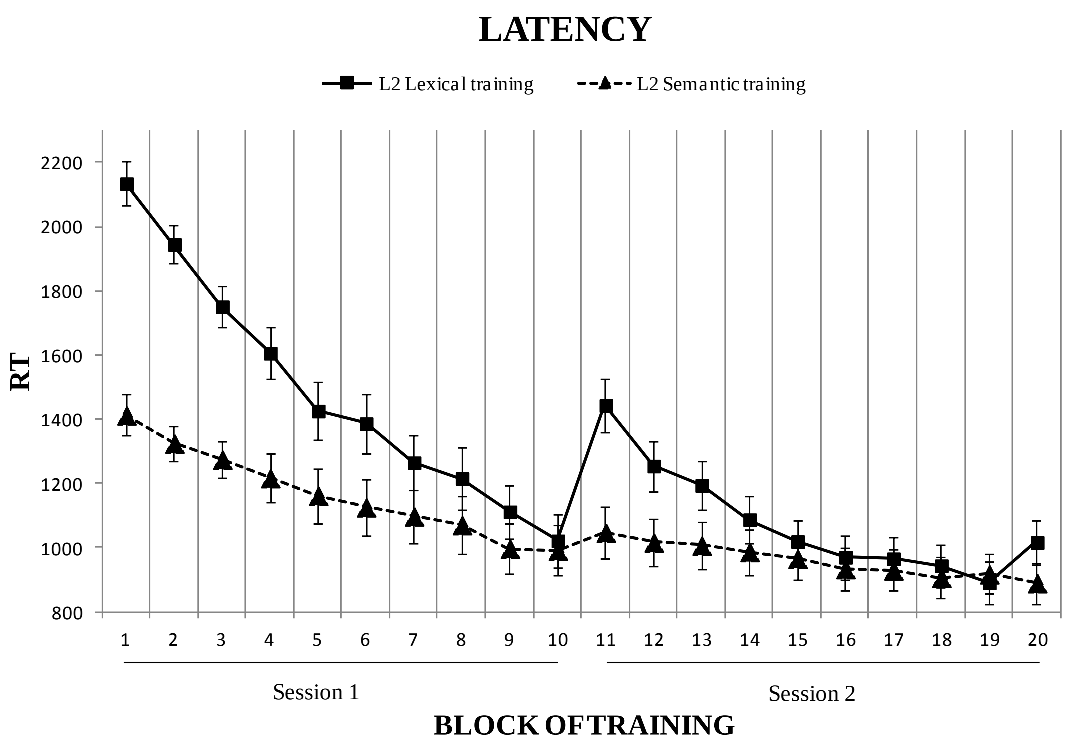


Figure 14. Response times (in milliseconds) and accuracy rates (in percentages) obtained in the L2 (second language) training (lexical training, semantic training) across blocks of training. Error bars represents standard errors.

The accuracy analyses revealed a significant effect of type of training[2], $F(1, 37) = 4.24$, $p = .046$, $\eta^2 = .10$, but only during the first 6 blocks of training. Both types of training revealed a ceiling effect from Block 11 thereafter, with correct scores of around 90% and 85% in the semantic and lexical types of training, respectively. Participants produced more correct responses during L2 semantic training (86.51%, *SE*

= 2.94), relative to L2 lexical training (80.42%, *SE* = 3.10). The main effect of block of training was significant, $F(19, 684) = 17.24$, $p < .001$, $\eta^2 = .32$. The Type of training x Block of training interaction was not significant, $F < 1$. Linear trend analysis was significant for lexical training, $F(19, 323) = 6.93$, $p < .001$, $\eta^2 = .29$, and semantic training, $F(19, 361) = 12.47$, $p < .001$, $\eta^2 = .40$. Thus, there was a practice effect with more correct responses at the end of the training relative to the beginning of the learning process (see Figure 14). The semantic training and the lexical training groups differed at the beginning of the learning phase (Block 1), $F(1, 38) = 8.94$, $p = .005$, $\eta^2 = .19$; however, the effect of type of training was not significant at the end of learning (Block 20), $F(1, 38) = 1.33$, $p = .26$, $\eta^2 = .03$.

Importantly, we examined the possible effect of cost associated to the change between learning sessions by comparing the performance of participants at the end of the first day of training (Block 10) and the beginning of the second day of training (Block 11). An ANOVA was conducted with type of training (semantic training, lexical training) as a between-participants factor and training session (last block of first day, first block of second day) as a within-participants variable. The interaction between Type of training x Training session was significant $F(1, 36) = 9.71$, $p = .003$, $\eta^2 = .21$. There were significant differences between training sessions for lexical training, $F(1, 36) = 24.53$, $p < .001$, $\eta^2 = .41$, but not for semantic training, $F < 1$. The difference in response times between sessions was 421.8 ms for lexical training whilst this was 55.97 ms for semantic training. On the other hand, there were no differences between the semantic and lexical training at the end of the first training session, $F < 1$; however, at the beginning of the second training session, response times were slower in lexical training (1443 ms, *SE* = 83.48) than semantic training (1047 ms, *SE* = 79.20), $F(1, 36) = 11.84$, $p = .001$, $\eta^2 = .25$.

In the accuracy analysis, the main effect of the training session was significant, $F(1, 36) = 5.61$, $p = .02$, $\eta^2 = .13$. Participants remembered more words at the beginning of the second session (87.24%, *SE* = .92), than at the end of the first session (83.12%, *SE* = 2.89). The interaction between Training session x Type of training was not significant, $F < 1$.

Lexical decision task. The RTs associated with correct responses were trimmed following the procedure described in the training task (Tabachnick & Fidell, 2001). The percentage of outliers was 8.12% in lexical training and 8.27% in semantic training. An ANOVA was conducted with type of training (semantic training, lexical training) as a between-participants factor, and language (L1-Spanish, L2-Vimmi) as a within-participants variable. In the latency analyses, the main effect of language was significant, $F(1, 38) = 338.68$, $p < .001$, $\eta^2 = .90$. Participants were faster when they responded to L1 words (596 ms, *SE* = 9.14) relative to L2 words (734 ms, *SE* = 9.55). The main effect of type of training was not significant, nor did this variable interact with language (*F*s < 1). In the accuracy analyses, the main effect of language was significant, $F(1, 34) = 25.46$, $p < .001$, $\eta^2 = .40$. Participants were more accurate in response to the L1 words (98.67%, *SE* = 0.27) relative to the L2 words (86.17, *SE* = 2.59). The main effect of type of training was not significant and this variable did not interact with language (*F*s < 1) (Table 8).

Possible switching cost effects due to language alternation in the lexical decision task were analysed considering the words trained in L1 (Spanish) and L2 (Vimmi) and the language switch (switch trial and non-switch trial). Bilinguals usually take longer to respond on switch trials than on non-switch trials (a switching cost effect), reflecting inhibition of the non-target language. Furthermore, switching costs are usually larger when switching to L1 suggesting the stronger suppression of the dominant language (Meuter & Allport, 1999). The results revealed a language switching effect, $F(1, 38) = 16.33$, $p < .001$, $\eta^2 = .31$. Participants responded more slowly on switch trials (664 ms, *SE* = 9.84) than on non-switch trials (635 ms, *SE* = 9.00). However, the Language x Switch interaction and the Type of Training x Language x Switch interactions were not significant, *F*s < 1. Thus, a switching effect was observed, regardless of the language and the type of training.

Table 8. *Response Times and Accuracy Rates Obtained in the Lexical Decision Task*

	Spanish (L1)		Vimmi (L2)	
	RT	AC	RT	AC
Lexical Training	600 (13)	99.17 (0.39)	739 (14)	86.50 (3.67)
Semantic Training	592 (13)	98.17 (0.39)	729 (14)	85.83 (3.67)

Note. Mean reaction times (RT, in milliseconds), accuracy rates (AC, in percentages) and standard errors (in brackets) obtained in the lexical decision task as a function of the type of training (lexical training, semantic training) and the language (L1-Spanish, L2-Vimmi).

Translation task. In this task, ANOVAs were conducted on RTs and accuracy data with type of training (semantic training, lexical training) as a between-participants variable and the translation direction (forward translation from L1 to L2, backward translation from L2 to L1) as a within-participants variable. The percentage of outliers in this task was 4.70% in lexical training and 6.33% in semantic training. In the latency and accuracy analyses, the order in which participants performed the forward translation and backward translation was not significant (*F*s < 1), nor did this variable interact with the type of training (*p*s > .05).

The latency analysis revealed a main effect of type of training, $F(1, 34) = 5.13$, $p = .03$, $\eta^2 = .13$. Participants were faster to respond during lexical training (1321 ms, *SE* = 108.85) compared with semantic training (1670 ms, *SE* = 108.85). The main effect of translation direction was also significant, $F(1, 34) = 118.97$, $p < .001$, $\eta^2 = .78$. Participants were faster in the backward translation task (L2-L1, Vimmi-Spanish) (1252 ms, *SE* = 67.24) relative to the forward translation task (1738 ms, *SE* = 91.21). Finally, the interaction between Type of Training and Translation direction was also significant, $F(1, 34) = 10.27$, $p = .003$, $\eta^2 = .23$. In the backward translation task, there were no differences between lexical and semantic training, $F < 1$. However, in the forward translation participants were slower in the semantic training (1983 ms, *SE* = 128.99) compared with those in the lexical training group (1492 ms, *SE* = 128.99), $F(1, 34) = 7.25$, $p = .011$, $\eta^2 = .18$. The difference in RT between forward and backward translation was larger for semantic training (628 ms) relative to lexical training (343 ms) (see Table 9).

In the accuracy analysis, the type of training was not significant, $F < 1$. The effect of translation direction was significant, $F(1, 34) = 67.27$, $p < 001$, $\eta^2 = .66$. The accuracy rates were higher in the backward translation direction (72.96%, $SE = 3.99$) relative to the forward translation direction (55.04%, $SE = 4.23$). The interaction between Translation direction and Type of training failed to reach significance, $F < 1$.

Table 9. *Response Times and Accuracy Rates Obtained in the Translation Task*

	Forward Translation (L1-L2)		Backward translation L2-L1	
	RT	AC	RT	AC
Lexical Training	1492 (129)	51.94 (5.99)	1149 (95)	72.41 (5.65)
Semantic Training	1983 (129)	58.15 (5.99)	1355 (95)	73.52 (5.65)

Note. Mean reaction times (RT, in milliseconds), accuracy rates (AC, in percentages) and standard errors (in brackets) obtained in the translation task as a function of the type of training (lexical training, semantic training) and the translation direction (forward direction L1-L2, backward direction L2-L1).

Naming task. The percentage of outliers in this task was 8.59% for lexical training and 12.77% for semantic training (Tabachnick & Fidell, 2001). ANOVAs were conducted on the RT data with the type of training (semantic training, lexical training) as a between-participants variable and language (L1-Spanish naming, L2-Vimmi naming) as a within participant variable. In the accuracy analyses, the language in which participants performed the naming task was not considered as a variable because there was no variability in any level (the accuracy of all participants was 100% when they named the pictures in Spanish). In the latency and accuracy analyses, the order in which participants performed the L1 naming and L2 naming was not significant (*p*s > .05), nor did this variable interact with the type of training (*p*s > .05).

In the latency analyses, significant main effects were found for training, $F(1, 35) = 15.16$, $p < .001$, $\eta^2 = .30$; and language, $F(1, 35) = 140.99$, $p < .001$, $\eta^2 = .80$; whilst there was also a significant interaction between Type of Training and Language, $F(1, 35) = 19.73$, $p < .001$, $\eta^2 = .36$. When participants performed the picture-naming task in Spanish, there were no differences between lexical and semantic training, $F < 1$. However, in the L2 naming task, participants were faster during semantic training (842 ms, *SE* = 60.30) relative to lexical training (1311 ms, *SE* = 61.95), $F(1, 35) = 29.51$, $p < .001$, $\eta^2 = .46$ (see Table 10).

In the accuracy analyses, when participants named the pictures in Vimmi, correct responses were 46.81% (*SE* = 5.81) and 59.37% (*SE* = 5.65) in the lexical and semantic training respectively. However, the accuracy analysis revealed no significant differences between types of training, $F(1, 35) = 2.40$, $p = .13$, $\eta^2 = .06$.

Table 10. *Response Times and Accuracy Rates Obtained in the Naming Task*

	Spanish (L1)		Vimmi (L2)	
	RT	AC	RT	AC
Lexical Training	497 (54)	100 (-)	1311 (62)	46.81 (5.81)
Semantic Training	471 (52)	100 (-)	842 (60)	59.37 (5.65)

Note. Mean reaction times (RT, in milliseconds), accuracy rates (AC, in percentages) and standard errors (in brackets) obtained in the naming task as a function of the type of training (lexical training, semantic training) and the language (L1-Spanish, L2-Vimmi).

Sentence-processing task. We conducted ANOVAs on RTs and accuracy rates with the training method (semantic training, lexical training) as a between-participants variable and the sentence predictability (high predictability, low predictability) and plausibility of sentences (plausible sentences, implausible sentences) as within-participants factors. The RTs associated with correct responses were trimmed (Tabachnick & Fidell, 2001) and the percentage of outliers in the L2 lexical and semantic groups was 8.10% and 6.90%, respectively.

In the latency analysis, the Predictability x Type of training interaction was significant, $F(1, 38) = 12.29$, $p = .001$, $\eta^2 = .24$, whilst the Plausibility x Predictability x Type of Training interaction was marginally significant, $F(1, 38) = 3.11$, $p = .08$, $\eta^2 = .08$. No other main effects or interactions were significant (all *p*s > .05). When plausible sentences were taken into account, the Predictability x Type of training interaction was significant, $F(1, 38) = 13.05$, $p < .001$, $\eta^2 = .26$. The semantic training group showed faster RTs in highly predictable sentences (423 ms, *SE* = 28.64) relative to sentences of low predictability (460 ms, *SE* = 25.45), $F(1, 38) = 5.11$, $p = .03$, $\eta^2 = .12$. In contrast, the lexical training group showed the opposite pattern of results, with slower RTs in highly predictable sentences (451 ms, *SE* = 28.64) than in sentences of low predictability (405 ms, *SE* = 25.45), $F(1, 38) = 8.11$, $p = .007$, $\eta^2 = .18$. When participants processed implausible sentences, no main effects or interactions between variables were significant (all *p*s > .05) (see Table 11).

Table 11. *Response Times and Accuracy Rates Obtained in the Sentence Processing Task*

	High predictability		Low Predictability	
	RT	AC	RT	AC
	Plausible Sentences			
Lexical Training	451 (28.64)	75.37 (3.74)	405 (25.45)	75.29 (4.15)
Semantic Training	423 (28.64)	76.62 (3.73)	460 (25.45)	76.75 (4.15)
	Implausible Sentences			
Lexical Training	452 (26.93)	71.64 (3.63)	427 (28.70)	78.29 (4.03)
Semantic Training	438 (28.70)	80.67 (3.63)	438 (24.70)	78.06 (4.03)

Note. Mean reaction times (RT, in milliseconds), accuracy rates (AC, in percentages) and standard errors (in brackets) obtained in the sentence processing task as a function of the type of training (semantic training, lexical training), sentence predictability (high predictability, low predictability) and plausibility of sentences (plausible sentences, implausible sentences). Error bars represent standard errors.

In the accuracy analysis, the Plausibility x Predictability x Type of training interaction was marginal, $F(1, 38) = 3.25$, $p = .08$, $\eta^2 = .08$. No other main effects or interactions were significant (all *p*s > .05). When participants processed plausible sentences, none of the effects were significant (all $F < 1$). When participants processed implausible sentences, the interaction between Type of training and Predictability was significant, $F(1, 38) = 5.93$, $p = .02$, $\eta^2 = .13$. In semantic training, the predictability effect was not significant, $F < 1$, whilst lexical learners were more accurate in low predictable sentences (78.29%, *SE* = 4.03), than in highly predictable sentences (71.64%, *SE* = 3.63), $F(1, 38) = 6.12$, $p = .01$, $\eta^2 = .14$) (see Figure 15).

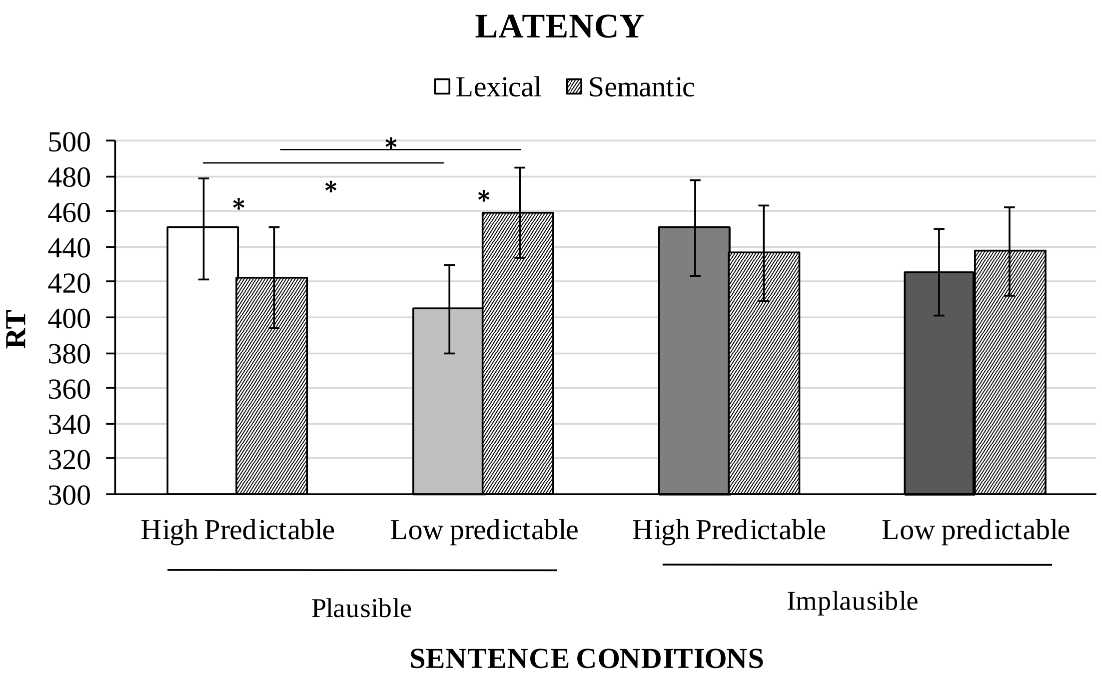


Figure 15. Response times (in milliseconds) obtained in the sentence-processing task as a function of the type of training (semantic training, lexical training), sentence predictability (high predictability, low predictability) and plausibility of sentences (plausible sentences, implausible sentences). Error bars represent standard errors.

DISCUSSION

The aim of this study was to compare the efficiency of a semantic method and a lexical method for the acquisition of L2 vocabulary. Previous studies have compared semantically mediated training with lexical training in out-of-context tasks (Chen, 1990; Comesaña et al., 2009, 2010; Poarch et al., 2014; Tonzar et al., 2009). Our study differed from previous research in two critical respects: (a) the inclusion of lexical and semantic learning tasks to reinforce the acquisition of L2 during training, and (b) the evaluation of the learning outcomes when participants processed new words in isolation and within sentences.

The pattern of results obtained in this study is organized as follows. Firstly, we discuss the course of learning in lexical and semantic training, and we evaluate the use of the new words acquired in L2 in out-of-context tasks. Finally, we examine the possible effect of the type of learning on the processing of L2 words within sentences.

Course of L2 vocabulary learning under semantic and lexical training

The usual way to study the efficiency of different L2 learning methods is through the use of evaluation tasks that are administered after finishing the learning phase (Hostetter, 2011; Kita et al., 2017; Macedonia, 2014; Macedonia & Klimesch, 2014; Macedonia & Knösche, 2011). However, this type of testing does not reveal progressive changes that occur throughout sessions of vocabulary acquisition. In our study, we evaluated the course of learning with two tasks of continuous training across two days of L2 acquisition. The two learning groups (semantic training and lexical training) showed continuous improvement during the 20 training blocks. In particular, through learning, the training tasks were associated with lower response latency and higher accuracy rates.

In addition, at the beginning of learning, the lexical training group showed worse performance than the semantic group. These differences cannot be due to L2 knowledge since it was the first training block and the participants had no previous knowledge of L2 (Vimmi, an artificial language). Two possible explanations could account for the pattern of results: (a) potential between-group baseline differences in L1 linguistic processing, (b) a greater difficulty associated with the lexical vs. semantic training task. With regard to the first alternative, participants were randomly assigned to training groups, which would minimize possible baseline differences. Furthermore, in order to ensure that the learners in the semantic and lexical training were matched in L1 skills, different post-test analyses were conducted. None of these analyses revealed differences between the training groups in the measures of L1 comprehension and L1 production. Thus, the initial differences observed between the two learning groups in the training task suggest that the lexical task was more difficult than the semantic task. In fact, the slope of learning (faster and more accurate responses through the training blocks) was more pronounced in the lexical group than in the semantic group, a pattern of results that is consistent with previous studies in which a steeper slope is observed in difficult vs. easy learning tasks (Bennett & Desforges, 1988; Duckworth & Quinn, 2009). Nevertheless, the performance of the two types of training was similar at the end of the practice blocks indicating that the two learning methods allowed the acquisition of L2 words in a similar manner.

As far as we know, there are no previous studies that have compared possible differences in the difficulty of semantic categorization and grapheme monitoring. However, Dufour and Kroll (1995) showed that in a categorization task, when category names were presented in L1 and target words in L2, relatively fluent bilinguals needed less time to respond (838 ms) than in a study conducted by Hauk, Rockstroh, and Eulitz (2001) with the grapheme monitoring task (between 950-1100 ms). However, direct comparison between studies is limited due to methodological differences. Hence, in order to further explore the differences between the semantic and lexical training tasks a new experiment was conducted. A group of 16 Spanish (L1) - English (L2) bilinguals performed the grapheme monitoring task and the semantic categorization task. In each task, participants received 3 blocks of trials (60 trials in each block). In the grapheme monitoring task, a grapheme was presented (e.g., *n*) followed by an English word that contained it (e.g., *hand*) and participants had to indicate whether this grapheme was also present in the Spanish translation of the English word (i.e., *mano*). In the semantic categorization task, a category name was presented (e.g., *fruit*) followed by an English word (e.g., *banana*) and participants had to indicate whether or not the word denoted an exemplar of the category previously presented. The two tasks were counterbalanced across participants. The latency analyses revealed a significant effect of task, $F(1, 15) = 108.45$, $p < .001$, $\eta^2 = .88$. RT were slower in the grapheme monitoring task (1495 ms, $SE = 70.74$) compared to the categorization task (848 ms, $SE = 20.96$). Furthermore, the task effect was significant in the accuracy analyses, $F(1, 15) = 40.89$, $p < .001$, $\eta^2 = .73$. Participants produced less correct responses in the grapheme monitoring task (74.79%, $SE = 2.18$) than in the semantic categorization task (89.87%, $SE = 0.89$). Thus, the results obtained in this new experiment with a new group of participants replicates the pattern of data obtained in the study, revealing a greater difficulty in the lexical vs. semantic practice task.

On the other hand, most studies comparing different methods of L2 vocabulary acquisition are conducted within a single day, during which participants receive the L2 training tasks followed by the tasks to evaluate the learning achievements (Macedonia et al., 2011; Poarch et al., 2014). In contrast, in our study, the training was administered over two consecutive days, which enabled us to evaluate the effect of lexical and semantic training through two learning sessions separated by a time interval of 24 hours. The results obtained in our study revealed a cost associated to the change between learning sessions in the lexical training but not in the semantic training group. In particular, there were no differences between the two learning groups at the end of the first session of learning; however, at the beginning of the second day of training, the lexical group performed the learning task more slowly than the semantic group. Previous studies have shown that the addition of a temporal delay between training sessions can produce both improvements in learning (i.e., a memory consolidation effect; Bakker, Takashima, Van Hell, Janzen, & McQueen, 2015; Tamminen & Gaskell, 2013) and a decline in the acquisition process (Adams, 1952; Anderson, Fincham, & Douglass, 1999; Digman, 1959). Currently, there is no consensus as to why an interval between training and testing can produce both benefits and costs in the learning process (Rickard, 2007). Bouton (1993) proposed that a retention interval could have differential effects on learning according to the type of knowledge acquired and represented in memory, whilst Rickard (2007) argued that the effect of forgetting over time depends on the type of task used for learning. Finally, Anderson et al. (1999) proposed that at the beginning of each learning session there is impairment in performance, which results from the forgetting of associative connections that are established in long-term memory during the learning process. Thus, the cost associated to the change between learning sessions in the lexical training suggests that the associative connections established between L1 and L2 words (lexical connections) are more susceptible to forgetting than the L2-concept connections formed in the semantic training group. However, it is important to note that this cost due to the change between learning sessions did not determine the final acquisition of vocabulary since the accuracy rates at the end of the training were similar in the lexical and semantic training group.

Processing of new L2 words in out-of-context tasks

In order to evaluate the processing of the new L2 words in isolation, we used three evaluation tasks: lexical decision, translation, and picture naming. The performance of the learners undergoing semantic and lexical training was similar in the lexical decision task. Thus, previous studies indicate that L2 learning based on lexical paradigms (e.g., word association method) makes learners particularly efficient in tasks involving lexical processing (e.g., Chen, 1990; Lotto & De Groot, 1998). However, the results found here show that people undergoing L2 semantic learning can achieve the same level of efficiency in lexical processing.

In the translation task, the semantic training group responded slower in the forward direction than the lexical training group. According to the RHM model (Kroll & Stewart, 1994) described in the introduction, forward translation is more conceptually mediated than backward translation. It is good to know the meaning of the words that are translated, and in fact, this type of processing is characteristic of fluent bilinguals compared with L2 learners (Kroll et al., 2010). However, the access to semantic information slows down the translation process relative to the use of a direct route between the words in L1 and L2. The data obtained in the forward translation task suggest that the semantic training group made use of a route of processing that is conceptually mediated while the lexical group made the translation through direct connections between their lexicons. However, this pattern of results does not imply that the semantic learning group could not make use of the lexical processing route. In fact, when we considered the backward translation task, performance was independent of the type of training used for word acquisition in L2.

In the naming task, the semantic training group was faster than the lexical training group when the pictures were named in L2. Semantic processing is mandatory in picture naming regardless of the output language (e.g., Glaser, 1992). However, after semantic access, L2 naming can be made by a direct link between the concepts and the words in L2 or by retrieving the equivalent translation in L1 (Potter & Faulconer, 1975, Potter et al., 1984). The faster L2 naming times found in the learners of the semantic training group suggests that they efficiently used a direct connection between concepts and L2 words. It could be argued that the advantage of the semantic group in the L2 naming task was due to the familiarity of the participants with the pictures since they were used in the L2 learning task. Although all participants underwent a familiarization phase in which they processed all the pictures at the beginning of the study; the semantic training group practiced more often with these pictures during the learning phase. Thus, the advantage of the semantic group in the L2 naming task could be driven by episode memory. Although this explanation cannot be discarded in our study, when participants were required to name the same pictures in L1, no differences were found between training groups. In addition, the episodic explanation would predict that the retrieval of episodic traces depends on the similarity between the context of learning and retrieval (the stimuli presented in the episode and the responses required to the participant) (Neill, 1997). In the semantic training phase, the pictures were paired with an L2 word and the participants did not have to provide any response, while in the naming task, the pictures were presented in isolation and participants had to name them. Thus, the pictures presented in the naming task would be a weak cue to retrieve the learning episode. Thus, whilst episodic memory may certainly play a role in the picture naming task, it is unlikely that this explanation alone is responsible for the observed data pattern.

Overall, the pattern of results obtained when learners performed out-of-context tasks suggest that semantic training favours the conceptual processing of L2 words (e.g., forward translation, picture naming in L2). Furthermore, this type of L2 vocabulary acquisition does not prevent efficient lexical processing when required by the task (e.g., backward translation, lexical decision task).

Processing of new L2 words in sentences

In our study, we evaluated the processing of new L2 words in a sentence-reading task. The learners read sentences in a mixing language paradigm (Altarriba et al., 1996), in which the sentences were presented word-by-word in Spanish (L1) and they ended with a final word in Vimmi (L2) (Kutas & Federmeier, 2014; Potter, Nieuwenstein, & Strohminger, 2008). The critical pattern of results was observed when participants read plausible sentences. The semantic training group benefited from the semantic constraint of sentences, so they processed more efficiently L2 words that were predicted by the sentence (highly predictable sentences) relative to the processing of words that were unlikely in the context of the sentence (sentences of low predictability). This pattern of data suggests that individuals undergoing semantic training were able to use the linguistic context to anticipate the arrival of the L2 word.

Previous studies show that bilinguals make use of semantic context (predictability of words within a sentence) to anticipate upcoming words during reading (Foucart et al., 2014). That is, sentence reading in bilinguals would be an incremental process involving predictions about upcoming words and consequently facilitating their integration once they appear. The results obtained in the semantic training group suggest that this type of training would favor the use of the semantic context during sentence reading in a way similar to that performed by fluent bilinguals. In contrast, learners undergoing lexical training showed that word processing in L2 was more efficient in sentences of low predictability compared with those of high predictability. Previous studies have shown that contexts of low predictability favour language co-activation and, therefore, the use of lexical connections between L1 and L2 (e.g., Schwartz & Kroll, 2006). Thus, the lexical training group processes L2 words more easily in contexts that allow the use of L1-L2 lexical links already acquired during L2 vocabulary training (sentences of low predictability).

A potential limitation of the semantic training method used in the current study is that only concrete familiar nouns were used as learning material. The semantic paradigm would be especially useful in early stages of L2 vocabulary acquisition as concrete words activate the semantic system more robustly than abstract words (Van Hell & De Groot, 1998), with the former being more readily acquired by L2 learners (Kaushanskaya & Rechtzigel, 2012). In future research, the semantic learning method could be extended to the acquisition of abstract words through the use of training tasks involving the use of gestures. In fact, previous studies show that gestures enhance semantic processing of words and gestures abolish the difficulty associated with the learning of abstract words (e.g., verbs) relative to nouns (e.g., concrete words) (García-Gámez & Macizo, 2019).

Conclusions

Speaking several languages is becoming increasingly necessary in today's world. Learning a L2 through immersion programs appears to be one of the best ways to learn a L2 (Cohen, 2014; Coyle, 2001; Genesse, 2014; Wesely, 2009). However, it is not always possible to study a language abroad. Thus, it is particularly important to investigate efficient L2 learning methods that can be applied in L1 speaking contexts.

In the current study, we compared a semantic vs. a lexical method for the acquisition of vocabulary in L2. The results indicate that both lexical and semantic training are useful for L2 vocabulary acquisition. Thus, the final learning rate and the latency in the retrieval of L2 words were similar in both types of learning at the end of training (the last block of learning). However, the results obtained in the evaluation tasks reveal differences in the cognitive architecture of the learners depending on the type of training. The group of learners that underwent lexical learning showed a strengthening of L1-L2 lexical links and they used these connections to process L2 words in isolation and in a sentence context. In contrast, the semantic training group used the connections between the conceptual system and the lexicon in L2 during out-of-context tasks (e.g., picture naming in L2) and when processing L2 words within sentences (e.g., the use of the context to process L2 words at the end of sentences). Thus, if we consider that a fluent bilingual is characterized by the semantic processing of the information when she/he communicates in L2 (Kroll et al., 2010; Kroll & Stewart, 1994), the semantic training used in our study appears to be suitable for L2 vocabulary acquisition.

Experiment 6. Lexical and Semantic Training to Acquire Words in a Foreign Language: An Electrophysiological Study

An ERP study was conducted to compare two learning methods for the acquisition of vocabulary in a FL. In the semantic method, the FL words were presented with a picture denoting their meaning and the learners practiced with a semantic categorization task (to indicate whether FL words were exemplars of a semantic category). In the lexical method, the FL words were paired with their translation in the L1 and the learners practiced with a letter-monitoring task (to indicate whether L1-FL words contained a grapheme). A translation task and a picture naming task were used to evaluate vocabulary acquisition. ERP modulations associated to semantic processing were more evident and they were more broadly distributed in the semantic versus lexical learning group. The pattern of results suggests that a single session of semantic learning favors the establishment of connections between semantics and the words learned in a new language.

This study has been submmited as García-Gámez, A. B., & Macizo, P (2019). Lexical and Semantic Training to Acquire Words in a Foreign Language: An Electrophysiological Study. *Brain and Language.*

INTRODUCTION

There are many questions about foreign FL acquisition that need to be addressed. What is the efficient way to learn a new language? Are there strategies able to enhance this learning process? Nowadays, we are involved in multicultural societies and speaking different languages is becoming necessary. Not only children, but also adults have to face with these new situations in their daily lives. It seems that immersion programs are the best way to learn a new language (Genesee, 2014) but this option is not always available. For this reason, researchers are looking for learning tools or strategies able to facilitate FL acquisition in L1 speaking contexts. It is necessary to implement efficient L2 learning methodologies based on scientific evidence not only at school for children, but also for adults.

To design effective FL learning methods it is first necessary to understand how novel and expert bilinguals manage linguistic processing across languages. According to the RHM (Kroll & Stewart, 1994), words in the first language (L1 lexicon), the second language (L2 lexicon) and the semantic system of L2 learners are interconnected. However, the weight and use of these connections depends on the fluency of participants in their L2 (De Groot & Poot, 1997). Thus, while novice learners make preferential use of L1-L2 connections, expert learners preferentially rely on the connections between semantics and L2 words (see Kroll et al., 2010, for a review). However, semantic processing is also possible and desirable in novice learners of a FL (Altarriba & Mathis, 1997; Dufour & Kroll, 1995; Ferré et al., 2006; Talamas et al., 1999).

The word translation task and the picture naming task have been used to evaluate semantic and L1-L2 lexical processing during the acquisition of a FL. With regard to the word translation, behavioral studies have shown that the performance on this task depends on the translation direction and the L2 fluency of the learners. In particular, poorly fluent bilinguals translate more rapidly in the backward direction (from L2 to L1) than in the forward direction (from L1 to L2) (Kroll & Stewart, 1994). The better performance in the backward translation would be explained by the use of direct connections between the L1 and L2 lexicons while the forward translation would involve additional semantic processing which would slow down the translation process. However, this asymmetry in the translation task is attenuated in fluent bilinguals which seems to indicate that they use a semantic route of processing regardless of the translation direction (Kroll & Linck, 2007).

On the other hand, regarding the picture naming task, many studies have confirmed that L2 naming is slower than L1 naming. In addition, these differences in the speed of processing depending on the language of the naming task are more evident in low vs. high fluency bilinguals (Kroll et al., 2002). Poorly fluent bilinguals would have weak connections between the semantic system and the L2 lexicon, so they would use a route of processing through their L1 which would slow down the retrieval of the picture names in L2. On the contrary, proficient bilinguals would employ a direct connection between meaning and L2 words which would produce faster naming times.

In general, the main feature of lexical and semantic processing during L2 vocabulary acquisition would be the strengthening of connections between meaning and L2 words as L2 fluency increases. Thus, it would be desirable to use learning strategies that favor the establishment of these connections between semantics and FL lexicon. In fact, the benefit of semantic vs. lexical learning programs has been confirmed in the past. To illustrate, Comesaña et al., (2009) compared the learning of L2 vocabulary in children through a word-word association method that would reinforce the L1-L2 lexical connections, and a learning method based on the association between pictures and L2 words that would favor semantic-L2 lexical processing. The results revealed that only children who received a picture-word association training showed effects derived from semantic processing (e.g., semantic interference effect in a translation recognition task); a pattern of results which is usually found in expert bilinguals (see also, Comesaña et al., 2010; Tonzar, et al., 2009).

The advantage associated to semantic learning strategies do not only appear in children but also in adult L2 learners. In a recent study, García-Gámez and Macizo (2019) compared the acquisition of L2 vocabulary in adults through a semantic training (based on picture-L2 word association and semantic categorization) and a lexical training (based on L1-L2 word association and grapheme monitoring). After learning, the authors found that the semantic training group responded more slowly than the lexical training group in a forward translation task. These results suggested that the semantic training group made use of a route of processing conceptually mediated while the lexical group made the translation through direct connections between their lexicons. In addition, in a L2 naming task, the semantic training group was faster than the lexical training group suggesting again that the semantic training group efficiently used a direct connection between concepts and L2 words.

Thus, there is abundant behavioral evidence showing that L2 vocabulary learning based on semantic training favors the establishment of connections between the semantic system and new L2 words. However, behavioral studies sometimes fail to capture subtle changes in language processing that occur with minimal training in L2 learning. For example, McLaughlin et al. (2004) showed that brain activity of adult L2 learners measured by event-related potentials (ERPs) discriminated the processing of new L2 words compared to the processing of L2 pseudowords with only one session of L2 vocabulary training. However, no differences were found between L2 words and pseudowords when behavioral measures were considered. The authors concluded that some aspects of a new language may be overlooked by current behavioral assessments. However, to the best of our knowledge, there are no previous electrophysiological studies comparing the effect of semantic vs. lexical training on L2 vocabulary acquisition. In our study, we address this point directly.

Two electrophysiological components have been used to index vocabulary acquisition in L2, the N400 component and the late positivity component (LPC) (Midgley et al., 2009; Yum et al., 2014). The N400 component is a negative-going waveform peaking at approximately 350-450 ms after stimulus onset, whose amplitude is sensitive to the processing of lexico-semantic information (Kutas & Hillyard, 1980). In bilingual studies, it has been shown that the processing of L2 words elicit smaller N400 than the processing of L1 words. Moreover, L2 words also produce larger N400 in highly proficient L2 bilinguals compared with L2 learners with low proficiency (Midgley et al., 2009). Thus, modulations of N400 amplitude when individuals process L2 words have been considered in previous studies as an index of L2 proficiency.

As indicated above, many behavioral studies have found that word retrieval is easier in backward translation than in forward translation due to the difficulty associated with semantic processing in L1-L2 translation (Kroll & Stewart, 1994). In electrophysiological terms, an easy retrieval of lexical information would be associated with an attenuation of the N400 component. For instance, word frequency is one of the main indicators of difficulty in lexical access (e.g., Hudson & Bergman, 1985; Monsell et al., 1989), and this lexical factor produce a N400 attenuation (Rugg, 1990, Van Petten & Kutas, 1990) with reduced brain-wave negativity during the processing of high vs. low frequency words. However, electrophysiological studies with the translation task are very limited and the results are mixed. Christoffels et al. (2013) observed a greater amplitude of the N400 component when Dutch (L1) - English (L2) bilinguals translated words in forward vs. backward direction. In contrast, Jost et al., (2018) did not find differences related to the translation direction in the N400 time-window.

The LPC component is a late-onset sustained positivity (about 500 ms) that has been related to reanalysis of information and response monitoring in the field of language processing (Kolk & Chwilla, 2007). In translation tasks, the LPC shows an inverse polarity between parietal and frontal regions and this component has been associated with the reprocessing of information between the input and the output language. For example, in a translation recognition task, a greater LPC amplitude appears in later regions when word pairs are not translations but are semantically related compared to unrelated word pairs (e.g., Guo et al., 2012). Other authors have linked the LPC found in translation tasks to the process of establishing connections between words across languages (Jackson et al., 2001).

On the other hand, the LPC has been obtained when bilinguals perform picture naming tasks (Martin et al., 2013). For instance, the LPC amplitude is greater when the difficulty in the retrieval of the picture names increases (e.g., naming tasks that involve language switching across trials). In general, studies of lexical processing in bilinguals seem to indicate that the more complex the processing of the stimuli (e.g., L2 naming vs. L1 naming) the greater the mean amplitude of the LPC component (Jackson et al., 2001; Khateb et al., 2007; Kieffaber et al., 2013; Moreno, et al., 2002).

The current study

The aim of our study was to evaluate the impact of a semantic vs. lexical training on the acquisition of vocabulary in a FL. To this end, two groups of Spanish speakers (L1) learned words in an artificial language (Vimmi) under two types of training. The semantic training group received L2 words accompanied by a picture representing their meaning and participants practiced during learning with a semantic categorization task. The lexical training group received the word in L2 along with its L1 translation and the practice task involved the identification of graphemes between languages (grapheme monitoring task). At the end of the training, electrophysiological indices were obtained when participants performed a translation and a naming task to evaluate the acquisition of L2 vocabulary. If we consider that the N400 and LPC components capture the difficulty of lexical-semantic processing and the retrieval and reprocessing of lexical information in translation and picture naming (Jost et al., 2018; Martin et al., 2013), these components would be modulated by the type of training (semantic vs. lexical learning) used for the acquisition of vocabulary in L2.

METHOD

Participants

Fifty-six Spanish monolingual students from the University of Granada participated in this study for course credits. None of the participants reported history of language disabilities and they had normal or corrected-to-normal visual acuity. Twenty eight participants were randomly assigned to the semantic training group (22 women, 6 men) and the rest of participants were assigned to the lexical training group (22 women, 6 men). The mean age of participants in the semantic (M = 21.85, SD = 4.34) and lexical group (M = 20.79, SD = 3.18) was equated, $t(54) = 1.05$, $p = .30$. Two participants in the semantic training and three participants in the lexical training were left-handed. The experiment was undertaken in accordance with the 1964 Helsinki declaration. The Ethic Committee at the University of Granada approved the experimental procedure used in the study (Number issued by the Ethical Committee: 86/CEIH/2015) and each participant provided written informed consent before taking part in the experiment. The required sample size was determined using the G*Power program 3.1.9.2 (Faul, Erdfelder, Lang, & Buchner, 2007). To achieve a 95% statistical power at $\alpha = .05$ and a small effect size (0.42) computed based on a $\eta_p^2 = .15$, in a 2 x 2 repeated measures analysis of variance (ANOVA), the required sample size was $N = 22$. Thus, the sample used in this study was sufficient to capture the effects evaluated in the experiment.

Design and Materials

In the current study, participants learned a set of 60 Spanish (L1) - Vimmi (L2) words. The L2 vocabulary learning was conducted in a single session. After finishing the training part, the experimenter mounted the electrode cap to the participants. Afterwards, participants carried out two tasks to evaluate the acquisition of L2 words in the following order: a translation task and a naming task.

L2 vocabulary learning tasks. Two L2 training methods were used. Half of the participants were subjected to a lexical training and the remaining half of participants performed the semantic training. When the L2 training phase was considered, a 2 x 10 mixed design was used with type of L2 training (semantic training, lexical training) as a between-participants factor and block of training (10 levels, from the first block to the last block of training) as a within-participants variable. In order to control several linguistic variables, Spanish words from six semantic categories were chosen (Battig & Montague, 1969; Moreno-Martínez et al., 2014) three categories were from the living domain (four-footed animals, body-parts, fruits) and three from the non-living domain (kitchen utensils, musical instruments, vehicles). Within each semantic category, ten words were selected. Hence, the total amount of words to be learned was 60. In addition, for each word, a picture denoting its concept was selected (Pérez & Navalón, 2003; Snodgrass & Vanderwart, 1980). Each Spanish word and its corresponding picture were randomly paired with a Vimmi word. Table 12 lists the lexical characteristics of Spanish words taken from Cuetos et al. (2011), and Pérez and Navalón (2003) (length, word frequency, orthographic neighbourhood, age of acquisition, familiarity, manipulability, typicality, imageability, and concreteness of words), lexical properties of Vimmi words (length, orthographic neighbourhood, and shared graphemes with the Spanish words) and visual properties of pictures taken from Pérez and Navalón (visual complexity of pictures, image agreement, image variability, and picture-name agreement). The complete set of stimuli is presented in the Appendix 2.

Table 12. *Characteristics of the Stimuli Used in the Study*

	Mean	SD
Spanish words		
Length	5.68	1.71
Frequency	38.90	61.80
Orthographic neighbourhood	3.36	5.00
AoA	2.47	0.74
Word familiarity	4.16	2.89
Word manipulability	3.36	0.97
Word typicality	4.50	0.47
Word imageability	4.00	2.98
Word concreteness	4.10	2.86
Vimmi words		
Length	5.66	1.45
Orthographic neighbourhood	1.11	2.24
Shared graphemes	2.11	0.97
Shared graphemes (positional)	0.50	0.67
Pictures		
Visual complexity	2.59	0.82
Image agreement	3.94	0.51

Image variability	2.38	0.67
Picture-name agreement	88.2	15.0

Note. Lexical characteristics of words and visual properties of pictures used in the study. *Length:* Number of graphemes of the word. *Frequency:* Spanish word frequency per one million words. *Orthographic neighbourhood:* Number of words that can be formed by substituting a single letter at any of the letter positions within the string. *Age of Acquisition:* the estimated age at which a word is acquired on a 7-point rating scale (0–2, 3–4, 5–6, 7–8, 9–10, 11–12, and 13 years or more). *Word familiarity:* the degree to which the concept denoted by the word is encountered in real life on a 5-point scale (1 = very unfamiliar, 5 = very familiar). *Word manipulability:* degree to which use of the human hand is necessary for the object denoted by the word to perform its function on a 5-point scale (1 = low manipulability, 5 = high manipulability). *Word typicality:* degree to which a concept denoted by a word is a representative exemplar of its category on a 5-point scale (1 = least typical, 5 = most typical). *Word imageability:* how easy it is for a word to arouse mental images on a 7-point scale (1 = low imageability, 7 = high imageability). *Word concreteness:* the degree to which the concept denoted by a word refers to a perceptible entity on a 7-point scale (1 = low concreteness, 1 = high concreteness). *Visual complexity*: of the pictures on a 5-point scale (1 = drawing very simple, 5 = drawing very complex). *Image agreement:* the degree to which the picture is similar to the real object depicted in the drawing on a 5-point scale (1 = low agreement, 5 = high agreement). *Image variability:* whether the name of the object evokes few or many different images on a 5-point scale (1 = few images, 5 = many images). *Picture-name agreement:* how close the picture matches the name given to the object in percentages (higher values, higher match).

The 60 Spanish-Vimmi words were randomly grouped into 10 learning sets of 6 words. All participants performed ten blocks with the learning and practice tasks described below. In each block, the 60 Vimmi words to be learned were presented once. Hence, the training involved a total of 10 exposures to each L2 word and the total amount of trials was 600.

In the L2 vocabulary learning phase, the task started with the message "word learning". Afterwards, participants pressed the space bar and 500 ms later the stimuli to be learned were presented (e.g.*, plátano*, banana in Spanish). In the semantic training, a picture appeared in the middle of the screen (e.g., a picture depicting a banana) (7 x 7 cm average size) and the Vimmi word below (e.g., *raone*) (Courier New, bold font, 18 point size). In the lexical training, the Spanish and the Vimmi words were presented one above the other (e.g., *plátano* and *raone*) (Courier New, bold font, and 18 point size). The stimulus to be learned remained on the screen for 4000 ms and 500 ms later, the next stimulus appeared. After the presentation of 6 stimuli, the message "practice new learned words" appeared on the screen, participants pressed the space bar to continue and 500 ms later, they performed the practice task. In the practice task of the semantic training, a category name was presented in capital letters for 750 ms (e.g., fruit). After 250 ms, a Vimmi word appeared in the middle of the screen until the participant responded (e.g., *raone*). Participants pressed the Z and M key to indicate whether or not the Vimmi word denoted an exemplar of the category previously presented (e.g., "yes" in the *plátano-raone* pair). Within each set, on half of the trials the Vimmi word was an exemplar of the semantic category and on the remaining trials it was not. Across the entire task, there were the same number of "yes" and "no" responses in each of the six semantic categories. In addition, across participants, all Vimmi words were assigned to the "yes" and "no" responses. The assignment of Z and M keys to "yes" and "no" responses was counterbalanced across participants.

The practice task in lexical learning consisted of the presentation of a grapheme in the middle of the screen for 750 ms (e.g., *n*). Following a 250 ms delay, a Vimmi word was presented until the participants' response (e.g., *raone*). Participants had to indicate by pressing the Z and M keys whether the Spanish translation of the Vimmi word contained the grapheme previously presented ("yes" in the *plátano-raone* pair). Each grapheme was one letter randomly selected from each Vimmi word. On half of the trials, this grapheme was present in the Spanish translation and on the remaining half of trials it was not. Across the task, half of the letters were vowels and the remaining graphemes were consonants. Across participants, all graphemes were assigned to "yes" and "no" responses. The assignment of Z and M keys to "yes" and "no" responses was counterbalanced across participants. The procedure used in the lexical and semantic training was taken from García-Gámez and Macizo (2019) (see Figure 16).

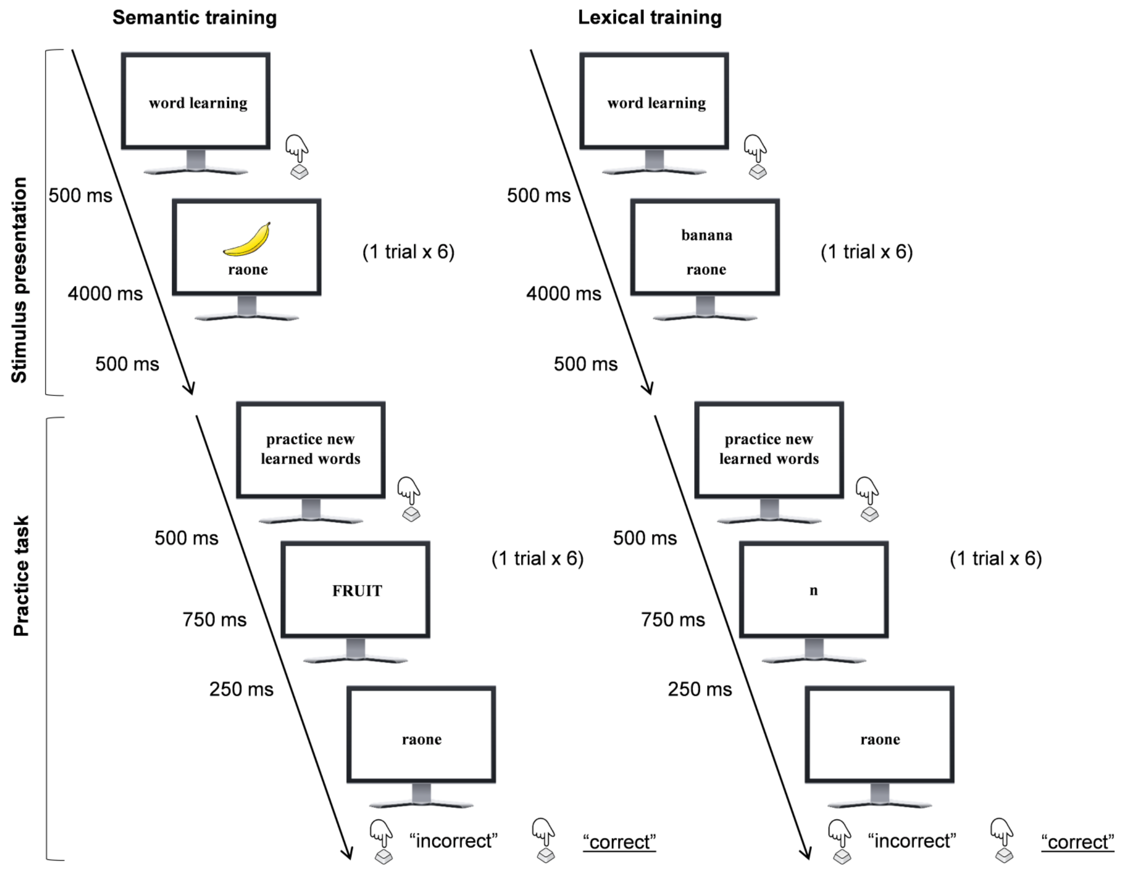


Figure 16. Description of the L2 vocabulary learning methods used in the current study.

Translation task. In this task, the 60 Spanish-Vimmi word pairs learned in the L2 vocabulary-training phase were used. The 60 Spanish words were presented for translation into Vimmi (forward translation) and the 60 Vimmi words were presented for translation into Spanish (backward translation). The order in which the two translation tasks were presented was counterbalanced across participants. In addition, the words within each translation task were presented in random order. On each trial, a fixation point appeared for 1000 ms in the middle of the screen followed by the word to be translated (e.g., raone) for 500 ms (Arial, 30 point size, black font, white background). A white screen then remained until the participants' response. Participants were required to say aloud the translation of each word.

Naming task. Participants were presented with 60 pictures denoting the meaning of the words learned in L2. Participants performed the naming task in Spanish and in Vimmi. The order in which they performed these naming tasks was counterbalanced across participants. Within each naming task, the pictures were presented in random order. On each trial, a fixation point was presented for 1000 ms, after which the picture to be named (e.g., the picture of a banana) (14 x 14 cm average size) appeared in the middle of the screen, followed by a white screen that remained until the participants' response.

Electrophysiological Recording

After the learning phase, the continuous Electroencephalogram (EEG) was recorded from 64 scalp electrodes mounted on an elastic cap (Quick-Cap, Neuroscan Inc.). The electrodes were arranged according to the extended 10-20 International System (Jasper, 1958). The EEG was initially recorded against an electrode placed in the midline of the cap (between Cz and CPz) and later off-line re-referenced to a common average reference. In order to control for blinks, a pair of electrodes was placed above and below the left eye. The horizontal and vertical eye movements were captured by another pair of electrodes located on the outer canthus in both eyes. The EEG signal was amplified by using the Neuroscan Synamps2 amplifiers (El Paso, TX) and filtered using a band pass of 0.01-100 Hz and digitalized at a 500 Hz sampling rate. The electrode impedance was kept below 5 kΩ. Digital tags were assigned to the stimuli of interest for each task. The EEG signal was analyzed by using the open-source toolbox ERPLab (López-Calderón & Luck, 2014). Eye blinks and other artifacts components were identified and corrected by means of independent component analysis (ICA) and careful visual inspection of the recordings. Epochs were baseline corrected using the mean activity during the -100 to 0 ms pre-stimuli period. Based on previous studies, a low pass filter of 30 Hz was applied (Verhoef, Roelofs, & Chwilla, 2009; Willems, Özyürek, & Hagoort, 2008).

For ERP data analysis, a representative sub-array of nine channels was used (Blackford, Holcomb, Grainger, & Kuperberg, 2012; Chauncey, Holcomb, & Grainger, 2008; Grainger, Kiyonaga, & Holcomb, 2006; Kuperberg, Delaney-Busch, Fanucci, & Blackford, 2018). As Blackford and colleagues (2012) mentioned, following this design, a single analysis of variance (ANOVA) can be used for each time window analysis and hence, this electrode selection is a good meeting point between the use of a simple design and the correct description of the overall distribution of the effects. The selected electrodes formed three columns in the central (Fz, Cz, Pz), left (F3, C3, P3) and right (F4, C4, P4) sides extending from the front to the back of the head (see Figure 17).

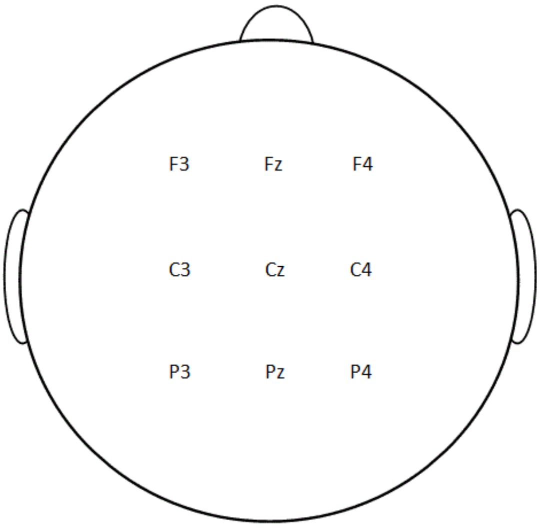


Figure 17. Scalp distribution of the electrodes selected for analyses. Electrodes formed three columns in the central (Fz, Cz, Pz), left (F3, C3, P3) and right (F4, C4, P4) sides extending from the front to the back of the head.

Data Analysis

The aim of the study was to examine electrophysiological data associated with the evaluation of learning tasks according to the type of L2 training. The behavioral data associated to the EEG recording session could not be analysed due to malfunctioning of the recording system. However, the pattern of behavioral data obtained with a different set of participants with exactly the same types of training (lexical training and semantic training) and evaluation tasks (translation task and naming task) are fully described by García-Gámez and Macizo (2019). Two participants, one lexical learner and one semantic learner, were removed from data analysis due to the high number of rejected epochs (more than 100 out of 120) obtained in both evaluation tasks.

Translation task. An ANOVA was conducted on ERP data with the Type of Training (semantic training, lexical training) as a between-participant factor and Translation Direction (forward, backward), Laterality (left, central, right) and Anterior-Posterior Electrode Distribution (frontal, central, parietal) as within participant variables. After eliminating trials in which the EEG signal was contaminated, the averages in each experimental condition (forward translation, backward translation) were comprised of a mean of 105 out of 120 trials in the lexical group of training trials and 97 trials out of 120 trials in the semantic group of training. The statistical analyses were conducted on three consecutive time windows time-locked to the onset of the words to be translated. These temporal windows were selected based on careful visual inspection and previous electrophysiological studies addressing translation tasks (Palmer, van Hooff, & Havelka, 2010; Phillips, Klein, Mercier, & de Boysson, 2006). The 150-300 ms window was used to index the P200 component associated to lexical access (Guo et al., 2012). The 300-500 ms time window was established to index the N400 component usually present in translation tasks (Phillips et al., 2006). Finally, the 500-700 ms time window was used to explore the P600 or LPC associated to the re-evaluation of information and semantic processing (Kolk & Chwilla, 2007).

Naming task. An ANOVA was conducted on ERP data with Type of Training (semantic training, lexical training) as a between participant factor and the Language in which participants named the pictures (L1 naming, L2 naming), Laterality (left, central, right) and Anterior-Posterior Electrode Distribution (frontal, central, parietal) as within participant variables. After eliminating trials in which the EEG signal was contaminated by eye movements or amplifier saturations, the averages in each experimental condition (L1 naming, L2 naming) were comprised of a mean of 106 out of 120 trials in the lexical group of training and 100 out of 120 trials in the semantic group of training. Three temporal windows were established time-locked to the presentation of the pictures to be named. These time-windows were selected based on visual inspection and previous studies on picture naming (Blackford et al., 2012; Verhoef et al., 2009). The 150-250 ms window was used to index the N200 component usually found in picture naming tasks (Christoffels, Firk, & Schiller, 2007). The 250-400 ms time window was selected to index the N400 component (Backford et al., 2012). Finally, the 400-700 ms time window was selected to evaluate late ERP modulations usually found when participants name pictures across languages (Christoffels et al., 2007).

In each time window of the translation and naming tasks, the main analyses are presented first. Afterwards, the results obtained in the lexical and semantic training group are presented separately. In all repeated-measure ANOVAs reported in the study, the Greenhouse-Geisser correction (Greenhouse & Geisser, 1959) for nonsphericity of variance was used for all *F*-ratios with more than one degree of freedom in the denominator; reported here are the original *df*, the corrected probability level, and the ε correction factor.

Procedure

The study was conducted in a single day, with participants being tested individually in the EEG recording room. E-prime experimental software was used for stimulus presentation and data acquisition (Schneider et al., 2002). Firstly, participants performed a familiarization task where all the 60 pictures used in the experiment were presented along with their Spanish names. The participants were instructed to see each picture and its more common name in Spanish and to press the space bar any time they wanted to see another picture.

After finishing the familiarization phase, the L2 vocabulary-learning task was introduced (semantic or lexical training). The training lasted 75 minutes approximately (depending on the response time of participants in the practice tasks). After completing the L2 acquisition phase, the experimenter proceeded with the mounting of the electrode cap. Finally, participants continued with the evaluation tasks (translation and naming tasks). The evaluation tasks order was not counterbalanced across participants. The complete session lasted approximately 120 minutes.

RESULTS

Firstly, we report the behavioral results obtained during the training task. Afterwards, we present the electrophysiological results of the two evaluation tasks (translation, naming) that participants performed after the training session.

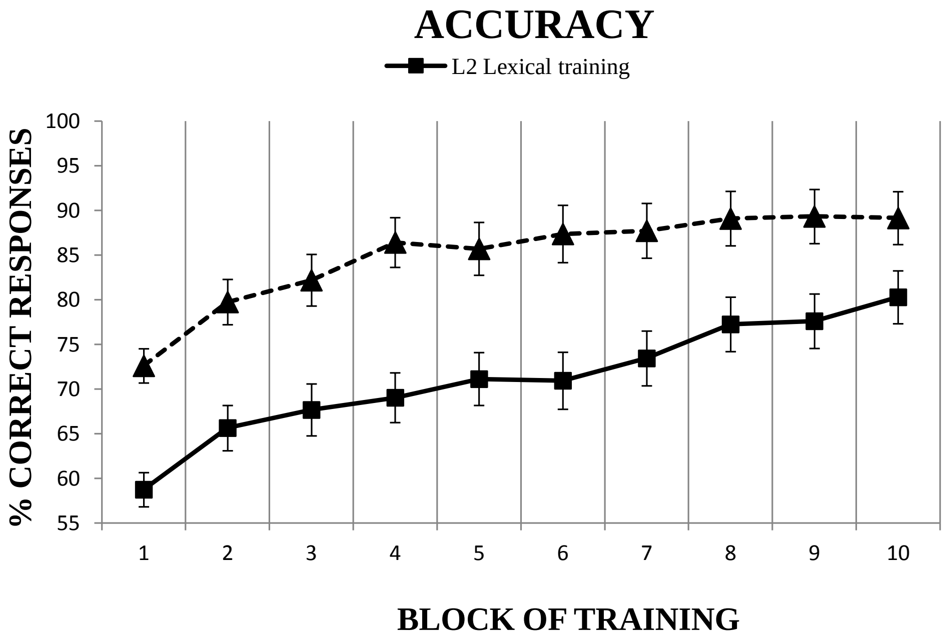
L2 vocabulary learning tasks

Correct responses in the L2 lexical training and L2 semantic training phases were 71.18% (*SE* = 6.40) and 84.94% (*SE* = 5.35), respectively. The reaction times (RTs) associated with correct responses in the training tasks were trimmed following the procedure described by Tabachnick and Fidell (2001) to eliminate univariate outliers. Raw scores were converted to standard scores (*z*-scores). Following standardization, any data points that were 3 *SD* outside the normal distribution were considered outliers. After removing outliers from the distribution, *z*-scores were calculated again. The filter was applied in recursive cycles until no observations were outside 3 *SD*. The percentages of data excluded from the L2 lexical and semantic training phases were 4.39% and 10.00%, respectively. We conducted an analysis of variance (ANOVA) on recall percentages and RTs with type of L2 training (semantic training, lexical training) as a between-participants factor and block of training (10 levels, from the first block to the last block of training) as a within-participants variable.

The accuracy analyses revealed a significant effect of type of training, $F(1, 54) = 13.70$, $p < .001$, $\eta^2 = .20$. Participants recalled more L2 words in the semantic training (84.95%, *SE* = 5.35), relative to the lexical training (71.18%, *SE* = 6.40). The main effect of block of training was significant, $F(9, 486) = 46.16$, $p < .001$, $\eta^2 = .46$. The Type of Training x Block of Training interaction was significant too, $F(9, 486) = 2.07$, $p = .03$, $\eta^2 = .04$. Linear trend analysis was significant in the lexical training, $F(1, 27) = 58.45$, $p < .001$, $\eta^2 = .68$, and the semantic training, $F(1, 27) = 59.72$, $p < .001$, $\eta^2 = .69$. In the lexical training, participants were more accurate at the end of the training (80.30%, *SE* = 2.97) compared to the beginning of the training (58.75%, *SE* = 1.90). The same pattern of results was found in the semantic training; participants recalled more L2 words at the end of the training (89.17%, *SE* = 3.00) compared to the beginning of the procedure (72.62%, *SE* = 1.90). Thus, there was a practice effect with more correct responses at the end of the training relative to the beginning of the learning process (see Figure 18). The differences in recall percentage between the first and the last block of training were higher in the lexical training (21.55%) compared to the semantic training (16.55%).

The latency analyses revealed a significant effect of the type of training, $F(1, 54) = 46.95$, $p < .001$, $\eta^2 = .47$. RTs were faster during semantic training (1292 ms, *SE* = 195) in comparison with lexical training (1964 ms, *SE* = 294). The main effect of block of training was significant, $F(9, 486) = 42.59$, $p < .001$, $\eta^2 = .44$. In addition, the Type of Training x Block of Training interaction was significant, $F(9, 486) = 3.10$, $p = .001$, $\eta^2 = .06$.

In the lexical training phase, linear trend analysis was significant, $F(1, 27) = 26.75$, $p < .001$, $\eta^2 = .50$, with faster responses at the end of training (1568 ms, *SE* = 71.43) relative to the beginning of training (2358 ms, *SE* = 108.14). Linear trend analysis was also significant during semantic training, $F(1, 27) = 107.02$, $p < .001$, $\eta^2 = .80$, with faster RTs at the end of training (1090 ms, *SE* = 71.43) relative to the beginning of the training (1695 ms, *SE* = 108.13). Thus, the more our participants trained, the faster they responded; however, the difference in RT between the first and the last block was larger in the lexical training phase (791 ms) compared with semantic training (604 ms) (see Figure 18).



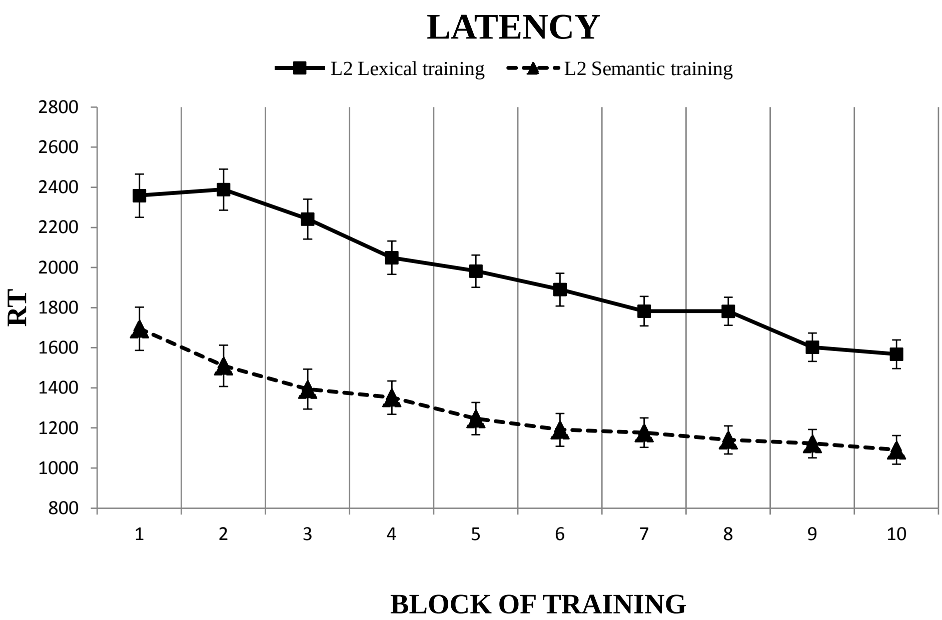


Figure 18. Percentage of correct responses and response times (in milliseconds) obtained in the L2 training task (lexical training, semantic training) across blocks of training. Error bars represent standard errors.

Translation task

The results obtained in the translation task are presented in Figure 19 and Figure 20 for the lexical and semantic training groups respectively. The complete pattern of statistical results obtained in the translation task for the lexical and semantic training groups is reported in Table 13.

150-300 ms time window. . The type of training effect was not significant, $F < 1$. No interactions including the type of training factor were significant (all *ps* > .05). No main effects or interactions were significant when these effects were evaluated separately in the lexical training and the semantic training group (all *p*s > .05).

300-500 ms time-window. The type of training effect was not significant, $F < 1$. The interaction between Type of Training x Laterality was marginal, $F(2, 104) = 2.89$, $p = .06$, $\varepsilon = .99$, $\eta^2 = .05$. There were significant differences between types of training in the left hemisphere, $F(1, 52) = 5.69$, $p = .02$, $\eta^2 = .10$. Brain waves were more negative in the semantic training (*M* = -.29 µV, *SE* = .14) compared with the lexical training (*M* = .20 µV, *SE* = .14). The type of training effect was not significant in the midline areas, $F < 1$, or in the right region, $F(1, 52) = 2.00$, $p = .16$, $\eta^2 = .04$.

Lexical training. The Translation Direction x Lateral axis interaction was significant, F(2, 52) = 4.77, p = .01, ε = .95, $\eta^2 = .16$. There were significant differences between translation directions in the right hemisphere, F(1, 26) = 10.17, p = .004, $\eta^2 = .28$. More negative brain waves were observed in the forward translation direction (M = -.49 µV, SE = .20) than in the backward translation direction (M = .01 µV, SE = .19). The differences between translation directions in the left hemisphere and the midline regions were not significant (all ps > .05).

Semantic training. In this group of participants, the Translation Direction x Anterior-Posterior axis x Lateral axis interaction was marginal, F(4, 104) = 2.27, p = .07, ε = .87, $\eta^2 = .08$. In frontal regions, the translation direction effect was not significant in the left area, t(26) = 0.25, p = .80, the midline area, t(26) = 0.26, p = .79, or the right area, t(26) = 0.72, p = .47. In central regions, the translation direction effect was not significant in the left area, t(26) = 0.37, p = .72, but brain waves were more negative in forward translation relative to backward translation in midline regions, t(26) = 2.13, p =

.04, and right regions, t(26) = 3.11, p = .004. Finally, in parietal regions, the translation direction was not significant in the left area, t(26) = 0.43, p = .67, it was significant in midline regions, t(26) = 2.47, p = .02, and it was marginal in the right region, t(26) = 1.95, p = .06.

Thus, the translation direction effect in the 300-500 ms time-window was more broadly distributed in the semantic training group than in the lexical training group.

500-700 ms time-window. The Type of Training x Anterior-Posterior axis x Lateral axis interaction was significant, $F(4, 208) = 2.50$, $p = .04$, $\varepsilon = .95$, $\eta^2 = .05$.

Lexical training. The Translation Direction x Anterior-Posterior axis interaction was significant, $F(2, 52) = 10.47$, $p = .001$, $\varepsilon = .62$, $\eta^2 = .29$. Additionally, the Translation Direction x Lateral axis interaction was significant, $F(2, 52) = 4.46$, $p = .03$, $\varepsilon = .66$, $\eta^2 = .15$. In frontal areas, the translation direction effect was not significant in the left region, $t(26) = 1.64$, $p = .11$, but it was significant in the midline region, $t(26) = 2.31$, $p = .02$, and in the right region, $t(26) = 3.60$, $p = .003$. In central areas, there were no significant differences between translation directions in the left region, $t(26) = 1.47$, $p = .15$, although they were significant in the midline region, $t(26) = 2.30$, $p = .03$, and in the right region, $t(26) = 4.03$, $p < .001$. Finally, in parietal areas, the differences between translation directions were significant in the left area, $t(26) = 3.56$, $p = .001$, and marginally significant in the midline, $t(26) = 1.76$, $p = .09$, and right areas, $t(26) = 1.73$, $p = .09$.

Semantic training. In the semantic training, the Translation Direction x Anterior-Posterior axis interaction was significant, $F(2, 52) = 16.71$, $p < .001$, $\varepsilon = 78$, $\eta^2 = .39$. There were differences between translation directions in frontal regions, $F(1,26) = 18.24$, $p < .001$, $\eta^2 = .41$, with more positive brain waves for the backward (M = -1.41 µV, SE = .63) than for the forward (M = -2.97 µV, SE = .58) translation direction. In central areas, the differences between translation directions were significant, $F(1, 26) = 16.46$, $p < .001$, $\eta^2 = .39$, with more positive amplitude for the backward (M = 1.52 µV, SE = .27) than for the forward translation direction (M = .52 µV, SE = .23). Finally, the difference between translation directions was significant in parietal regions, $F(1, 26) = 8.87$, $p = .006$, $\eta^2 = .25$. In this case, mean amplitude was more positive in the forward (M = 1.83 μV, SE = .52) than in the backward translation direction (M = 1.07 μV, SE = .57).

Table 13. *Statistical Analyses Performed on ERP Data. Translation Direction Effects and Interactions in the Translation task Performed by the Lexical Group of Training and the Semantic Group of Training.*

		Lexical Training		***Semantic Training***	
Time-window	**Effects**	***F***	***p***	***F***	***p***
150-300 ms	TD	.26	.61	2.23	.15
	TD x AP axis	1.99	.17	.62	.47
	TD x LM axis	1.51	.23	.46	.61
	TD x AP axis x LM axis	.84	.45	2.26	.09
300-500 ms	TD	1.27	.27	3.35	.07
	TD x AP axis	1.33	.26	1.57	.22
	TD x LM axis	4.77	.01*	2.72	.09
	TD x AP axis x LM axis	.92	.42	2.27	.07
500-700 ms	TD	12.05	.001**	12.82	.001**
	TD x AP axis	10.47	.001**	16.71	.001**
	TD x LM axis	4.46	.03*	.33	.63
	TD x AP axis x LM axis	.46	.72	1.78	.16

Note. TD: Translation Direction, AP: Anterior-Posterior, LM: Lateral-Medial. *$p \leq .05$, **$p \leq .001$

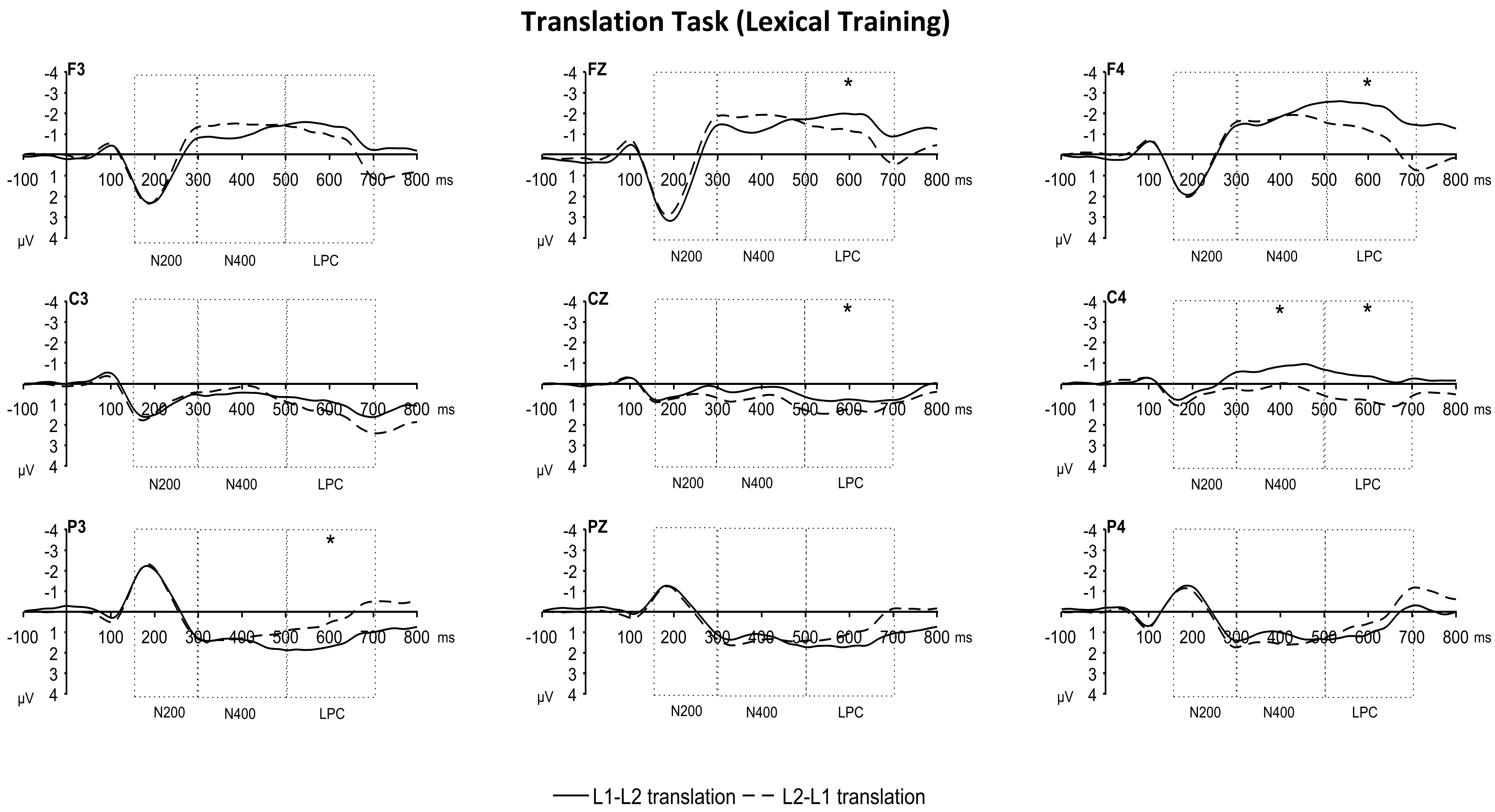


Figure 19. Grand Average ERPs for the forward (L1-L2) translation (solid lines) and backward (L2-L1) translation (dashed lines) obtained in the Lexical Group of Training. The time-windows analyzed in the study are framed by a dotted rectangle. * p < .05

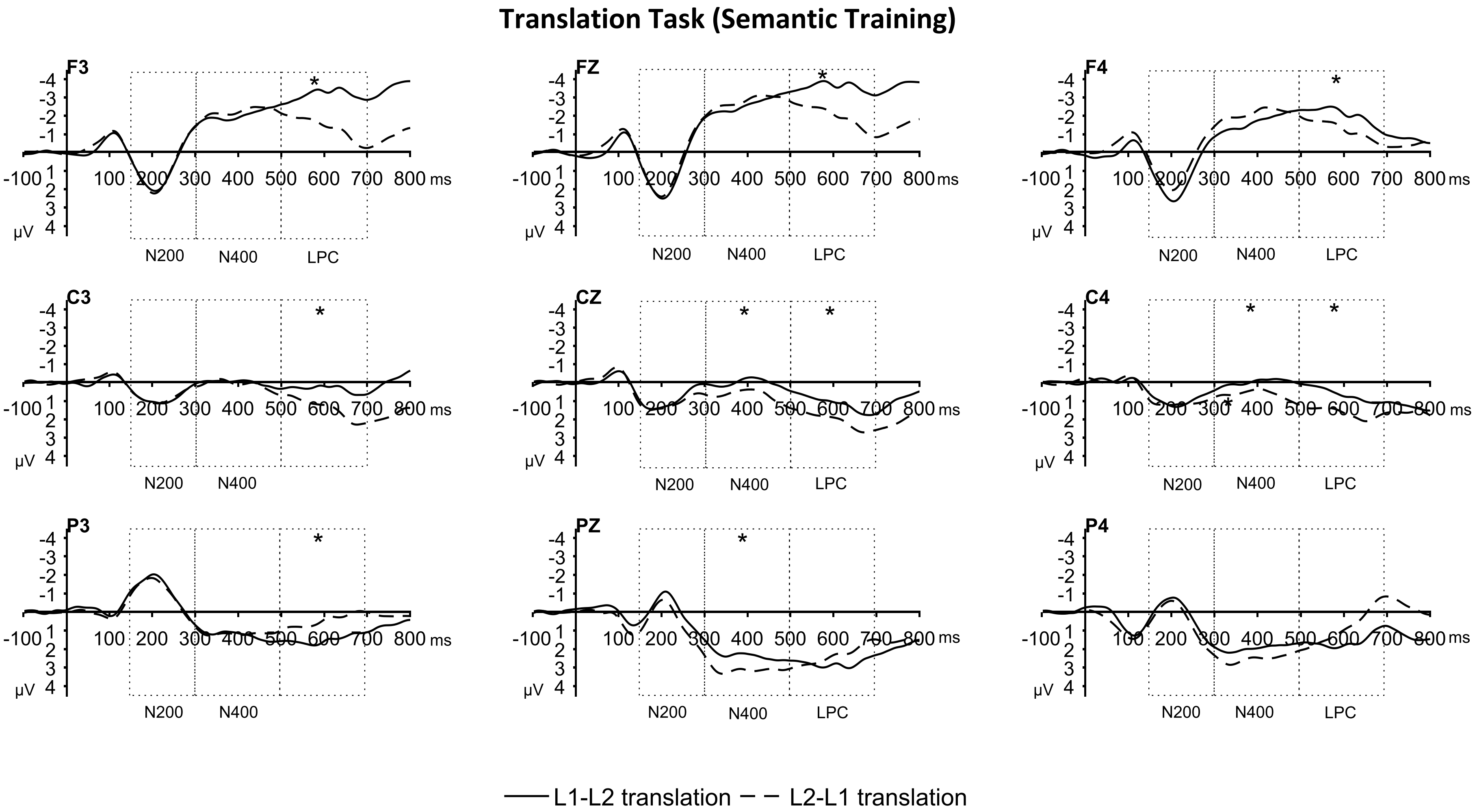


Figure 20. Grand Average ERPs for the forward (L1-L2) translation (solid lines) and backward (L2-L1) translation (dashed lines) obtained in the Semantic Group of Training. The time-windows analyzed in the study are framed by a dotted rectangle. * $p < .05$

Naming task

The results obtained by the lexical and semantic training groups are presented in Figure 21 and Figure 22 respectively. Additionally, summaries of statistical analyses conducted for both training groups are reported in Table 14.

150-250 ms time-window. The type of training was not significant, *F*(1, 52) = 1.51, *p* = .22, η^2 = .03, nor this variable interacted with any other (all *ps* > .05). No main effects or interactions were significant when brain waves of the lexical training group and the semantic training group were analyzed separately (all *p*s > .05).

250-400 ms time-window. The main effect of type of training was not significant, *F* < 1, and this variable did not interact with any other (all *p*s > .05).

Lexical training. In the lexical training, no main effects or interactions between variables were significant (all *p*s > .05).

Semantic training. In the semantic training, the Naming Language x Anterior-Posterior axis interaction was significant, *F*(2, 52) = 5.86, *p* = .007, *ε* = .86, η^2 = .18. In frontal areas, the naming language was not significant (*p* > .05). The effect was significant in central areas, *F*(1, 26) = 9.86, *p* = .004, η^2 = .28, with more negative amplitudes in L2 naming (*M* = -.36 Mv, *SE* = .39) than in L1 naming (*M* = .33 Mv, *SE* = .39). In parietal areas, the effect was not significant, *F* < 1.

400-700 ms time-window. In this time-window, the Type of Training x Lateral axis interaction was significant, *F*(2, 104) = 5.23, *p* = .007, *ε* = .96, η^2 = .09. Significant differences between types of training were obtained in the right hemisphere, *F*(1, 52) = 13.11, *p* < .001, η^2 = .20, with more positive amplitude in the semantic training (*M* = .71 Mv, *SE* = .18) than in the lexical training (*M* = -.24 Mv, *SE* = .18). The type of training was not significant in any other brain region (all *p*s > .05).

Lexical training. When the lexical training was considered, the Naming Language x Anterior-Posterior axis interaction was significant, *F*(2, 52) = 4.05, *p* = .03, *ε* = .76, η^2 = .13. The naming language effect was not significant in frontal and central regions (all *p*s > .05). However, in parietal areas, there were significant differences between naming languages, *F*(1, 26) = 5.40, *p* = .03, η^2 = .17, with more positive amplitude in L2 naming (*M* = 1.26 Mv, *SE* = .41) than in L1 naming (*M* = .60 Mv, *SE* = .33).

Semantic training. The Naming Language x Anterior-Posterior axis interaction was significant, $F(2, 52) = 6.02$, $p = .01$, $\varepsilon = .71$, $\eta^2 = .19$. There were differences between naming languages in frontal regions, $F(1,26) = 5.90$, $p = .02$, $\eta^2 = .23$, with more positive brain waves in L1 naming (*M* = -.95 Mv, *SE* = .50) than in L2 naming (*M* = -1.94 Mv, *SE* = .55). In central regions, the differences between naming languages were significant too, $F(1, 26) = 5.45$, $p = .03$, $\eta^2 = .21$, with more positive amplitude in L1 naming (*M* = 2.07 Mv, *SE* = .41) than in L2 naming (*M* = 1.34 Mv, *SE* = .28). In parietal areas, the differences between naming languages were also significant, $F(1, 26) = 4.88$, $p = .03$, $\eta^2 = .19$. However, in this case more positive amplitude was found in the L2 naming (*M* = 1.89 Mv, *SE* = .43) than in L1 naming (*M* = .97 Mv, *SE* = .49).

Table 14. *Statistical Analyses Performed on ERP Data. Naming Language Effects and Interactions in the Naming task Performed by the Lexical Group of Training and the Semantic Group of Training*

		Lexical Training		***Semantic Training***	
Time-window	**Effects**	***F***	***p***	***F***	***p***
150-300 ms	TD	.04	.83	.06	.80
	TD x AP axis	.35	.64	.64	.46
	TD x LM axis	.13	.78	.28	.73
	TD x AP axis x LM axis	.06	.86	.76	.52
300-500 ms	TD	.75	.39	.86	.36
	TD x AP axis	2.26	.13	5.86	.007*
	TD x LM axis	.58	.54	1.15	.30
	TD x AP axis x LM axis	2.20	.08	.56	.63
500-700 ms	TD	1.39	.25	2.65	.12
	TD x AP axis	4.05	.03*	6.02	.01*
	TD x LM axis	.81	.43	.52	.53
	TD x AP axis x LM axis	1.05	.37	.12	.93

Note. NL: Naming Language, AP: Anterior-Posterior, LM: Lateral-Medial. *$p \leq .05$

Naming Task (Lexical Training)

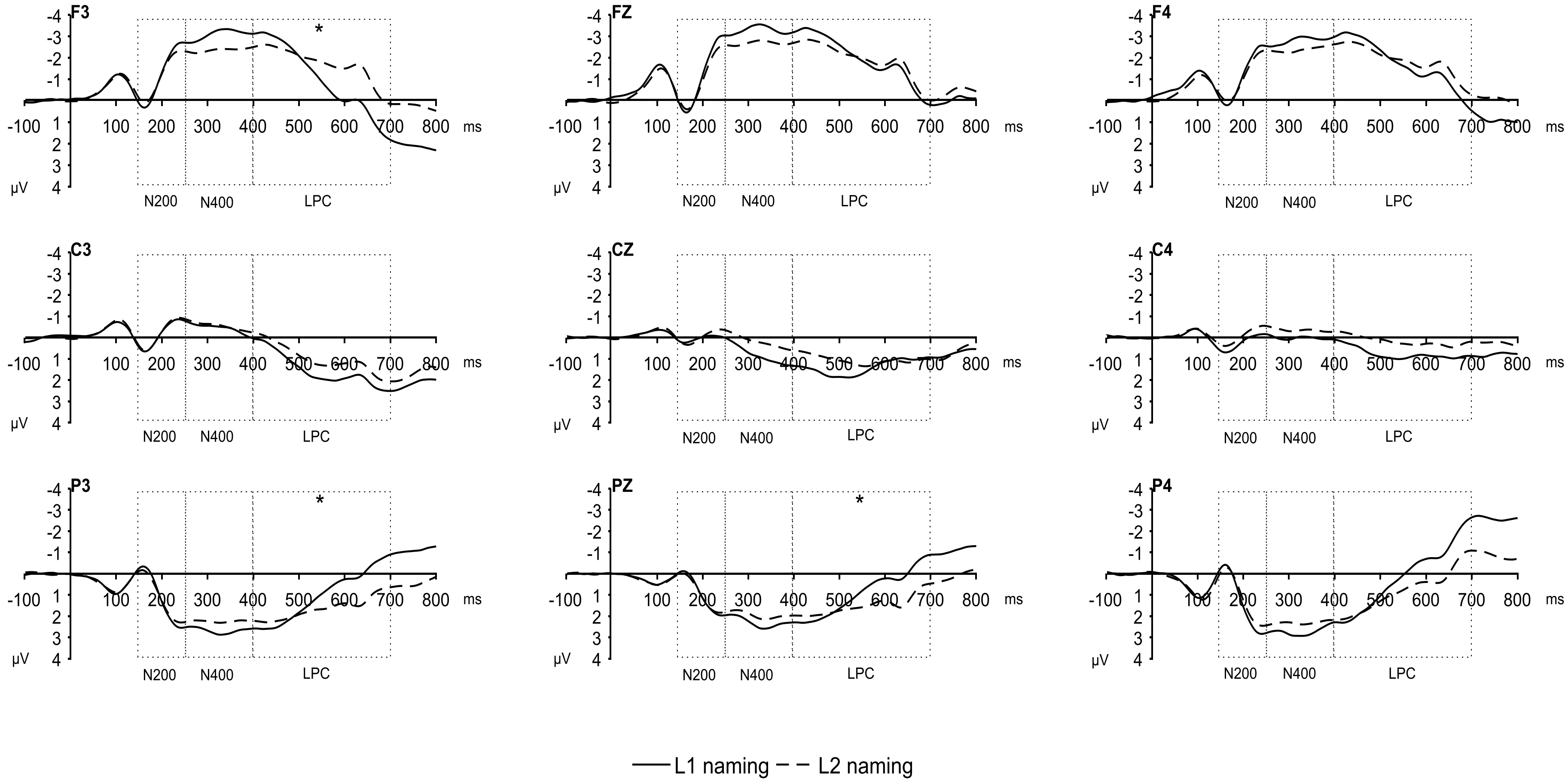


Figure 21. Grand Average ERPs for the Spanish naming (L1) (solid lines) and Vimmi naming (L2) (dashed lines) obtained in the Lexical Group of Training. The time-windows analyzed in the study are framed by a dotted rectangle. * $p < .05$

Naming Task (Semantic Training)

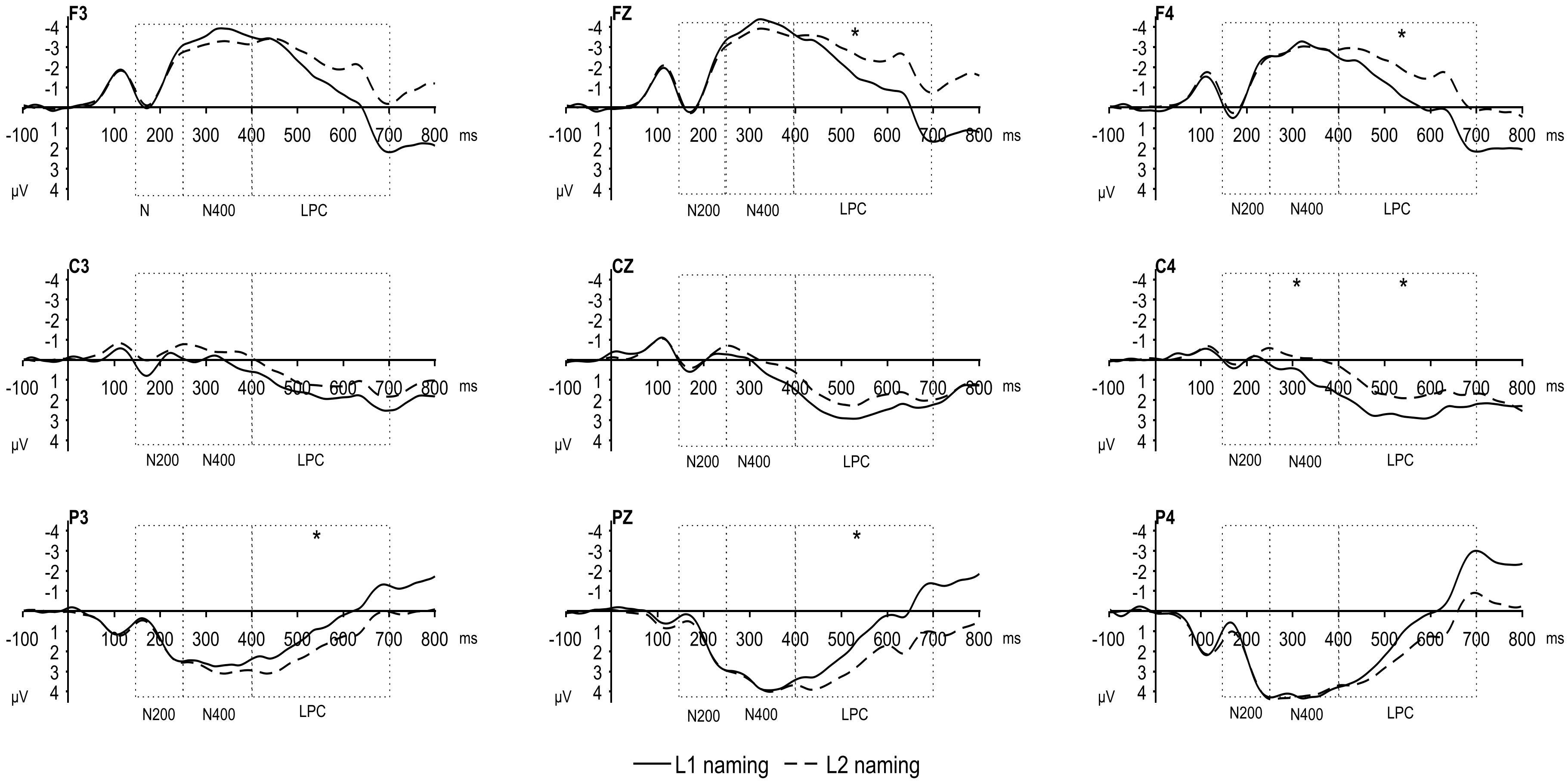


Figure 22. Grand Average ERPs for the Spanish naming (L1) (solid lines) and Vimmi naming (L2) (dashed lines) obtained in the Semantic Group of Training. The time-windows analyzed in the study are framed by a dotted rectangle. * $p < .05$

DISCUSSION

In the field of FL vocabulary acquisition, one of the main differences between novice and expert learners is the strength of the relationships between meaning and L2 words. Poorly fluent learners use their L1 lexicon when they process L2 words. In contrast, expert bilinguals rely on semantic processing during L2 processing (Frenck-Mestre & Prince, 1997; Kroll & Stewart, 1994; Potter et al., 1984; Talamas et al., 1999). Many behavioral studies show that this increased semantic analysis is possible in early stages of vocabulary acquisition if the learning tasks stimulate this type of processing (picture-word association training) as compared to other learning methods (word-word association training) (e.g., Comesaña et al., 2009). However, to our knowledge, there are no electrophysiological studies examining the consequences that the type of training (semantic vs. lexical) might have when individuals acquire vocabulary in a FL.

In our study, we compared the learning of L2 words from a lexical learning group that received pairs of L1-L2 words and practiced with a lexical task (grapheme monitoring) and a semantic learning group that received pictures-L2 words and practiced with a semantic task (semantic categorization). The results obtained at the end of 10 blocks of training revealed an overall learning effect with higher number of correct answers and faster response times at the end of training compared to the beginning of learning. Although the learning curve was independent of the type of learning, the performance of the learning task was lower in the lexical group than in the semantic group. This pattern of results has been previously obtained (e.g., García-Gámez & Macizo, 2019) with the same training tasks than those used in the current study and this task effect seems to be due to the greater difficulty of the grapheme monitoring than the semantic categorization (see García-Gámez & Macizo, 2019, for a detailed discussion on this issue and for experimental data showing the greater difficulty of the lexical vs. semantic task). However, regardless of the differences between the two training groups, word recall was higher at the end of training (above 80%) compared to the beginning of learning.

After the learning phase, we considered electrophysiological measures associated with the translation and naming tasks to evaluate the impact of lexical and semantic learning on L2 vocabulary acquisition.

In general, the results found in the translation task revealed greater brain-wave negativity in the 300-500 ms time-window in the semantic training group compared to the lexical training group. This pattern of results appears to indicate that the semantic training group was engaged in deeper semantic processing. In fact, previous studies show that the amplitude of the N400 component as an index of semantic processing depends on the fluency of participants in L2. Thus, L2 words produce larger N400 effects in highly proficient L2 bilinguals compared to L2 learners with low proficiency (Midgley et al., 2009; Rodríguez-Fornells, Cunillera, Mestres-Missé, & de Diego-Balaguer, 2009, for an overview). Therefore, the acquisition of L2 vocabulary based on semantic training would reflect the pattern of brain activity observed in fluent bilinguals during the performance of linguistic tasks.

Furthermore, the results revealed a translation direction effect with greater N400 amplitudes in the forward translation than in the backward translation. This pattern of data is consistent with other studies in which higher N400 is observed when the difficulty in retrieving lexical information increases (Rugg, 1990, Van Petten & Kutas, 1990). Moreover, this effect is consistent with the outcomes of many behavioral studies in which more difficulty (slower response time) is found when participants retrieve lexical forms in forward vs. backward translation (Kroll & De Groot, 2005; Kroll et al., 2010; Poarch et al., 2014; Sholl et al., 1995).

On the other hand, the N400 modulations due to the translation direction (greater negativity in forward than backward translation) were more widely distributed in the semantic training group (central and parietal regions) than the lexical training group (right central region) (Figures 19 and 20). The broader distribution of the N400 component when the semantic learning group performed the forward translation task is in line with previous research in which highly fluent bilinguals show an N400 component located in the same brain areas (Deacon et al., 2004; Holcomb, 1993; Kerhofs, Dijkstra, & de Bruijn, 2006; Palmer et al., 2010). Thus, electrophysiological response observed in the semantic vs. lexical training group closely resembles that of highly fluent bilinguals.

In the late time window (500-700 ms), the translation direction modulated the LPC component. However, the magnitude and distribution of this component was similar in the semantic and lexical training groups. In translation tasks, the LPC has been associated with reprocessing the information and checking the degree of correspondence between the source and target language (Guo et al., 2012; Kolk & Chwilla, 2007). Thus, the results obtained in our study suggest that, in translation tasks, the type of training (semantic vs. lexical learning) modulates the retrieval of words in the output language (modulation of the N400 component) but does not determine the evaluation of correspondences between words across languages.

With regards to the picture naming task, the results revealed N400 modulations depending on the output language. This effect was found in the semantic training group but not in the lexical training group. Specifically, the semantic training group showed greater negativity in central brain regions when the naming task was conducted in L2 compared to the L1 naming task. This data seems to indicate that the learners in the semantic group behaved as fluent bilinguals during the L2 naming task. In fact, in previous research it has been observed that expert vs. novice learners show greater sensitivity to semantic processing when performing L2 linguistic tasks which is reflected in a higher N400 (Midgley et al., 2009). On the other hand, in the semantic group, the greater N400 amplitude when naming in L2 vs. L1 would come from the weight of the connections between the semantic and the lexical systems in L1 and L2. It is widely demonstrated that in L2 learners, the weight of the connections between the meaning and the words in their native language are stronger than the connections between the meaning and the words in L2 (De Groot & Poot, 1997; Kroll & Stewart, 1994; Poarch et al., 2014; Van Hell & De Groot, 2008). Thus, the participants in the semantic group would show greater ability in recovering the name of the pictures in their native language than in their L2.

However, it may be asked why the participants in the lexical group did not show electrophysiological differences in the N400 time-window associated with the language used to name the pictures (L1 vs. L2). This outcome appears to suggest that, regardless of the naming language, these participants performed the task in the same manner by retrieving the name of the picture in their native language. In fact, several behavioral studies show that in poorly fluid bilinguals, L2 picture naming involves the retrieval of lexical information in the L1 lexicon (Kroll et al., 2002).

Thus, in the N400 time-window, the results obtained in the picture naming task would indicate that the learners in the semantic training group made an early differentiation between their languages and were able to use a processing route from semantics to L1 words or to L2 words according to task (L1 naming or L2 naming, respectively). In contrast, the lexical training group seemed to use a common processing route via L1 lexical activation regardless of the language needed to perform the naming task.

Finally, in the late temporal window, a LPC modulation was found due to the language in which the participants named the pictures. Specifically, in posterior regions, greater positivity was found when participants named the pictures in L1 than in L2. The polarity of the LPC found in our study was reversed in frontal regions; a pattern similar to that reported in previous work in which L2 learners show an opposite distribution of the LPC over the anterior scalp than the posterior scalp (e.g., Guo et al., 2012). The amplitude of the LPC component has been related to the difficulty in rechecking information. For instance, Guo et al. (2012) reported greater amplitude of the LPC component in a translation recognition task when word pairs were not translation equivalents but were semantically related compared to unrelated word pairs. In the related word pair condition, further reanalysis of the information would be necessary to corroborate that the words were not correct translations of each other. In our study, the greater amplitude of the LPC in the L2 vs. L1 naming task would imply the increased need to verify that the selected word actually designated the picture name in L2.

To conclude, the results obtained in this study seem to indicate that the type of training used for L2 vocabulary acquisition determines the strength and type of connections established between the semantic system and new L2 words. Electrophysiological evidence seems to indicate a deeper semantic processing in the picture-L2 word association training compared to the L1 word-L2 word association learning. Furthermore, the semantic training group appears to prefer a direct processing route between meaning and L2 lexicon (e.g., in L2 naming) while the lexical training group seems to be anchored in retrieving L1 lexical information when they perform linguistic tasks that involve the use of the new words learned in the FL.

-PART 3-

General Discussion, Conclusions and Future Research Questions

CHAPTER VI

General Discussion and Conclusions

We always thought that the title “How to learn vocabulary in a foreign language” was too pretentious and perhaps the reader had the same feeling at some point. There are lots of researchers working on this issue since long time ago. Currently, there is no consensus on the best strategy for learning a new language. Of course, we do not have a final conclusion on this issue. Learning a language is not as easy as rubbing a magic lamp. In our opinion, we will never find a perfect strategy for learning a new language. We are not robots; we are not programmed in the same way. Each person is unique and will probably have her best way of learning a new language. However, our research is a starting point to look for training programs that improve the process of learning a FL.

First of all, we would like to analyze what is common in our experiments and why these common factors are relevant. The age of the participants was one of the most important choices for our study. As mentioned in previous sections, children are still in the process of creating connections at brain level when they are exposed to a new language. However, when adults decide to learn a new language, their linguistic connections are already formed. It is therefore reasonable to expect that the two population groups will not behave in the same way when learning a new language. (Chen & Leung, 1989; Comesaña et al., 2009). This is the reason why we decided to select young adults as participants for our studies. Thus, the age of our participants was equated through the experiments (young adults) and this allows us to make direct comparisons across experiments with different vocabulary acquisition strategies in L2. On the other hand, all participants who took part in our experiments were native Spanish monolingual speakers. The monolingual condition was crucial because, as observed in previous works, people who already speak another language apart from their mother tongue, could have an advantage to learn a third language. Is not the objective of this work to analyze the positive influence of bilingualism in third language acquisition (TLA) and the linguistic consequences of multilingualism, but people who are used to learn languages develop special strategies and they have a larger intercultural and linguistic repertoire available during the learning process (Elorza & Muñoa, 2008; Jarvis & Pavlenko, 2008, see Cenoz, 2013 for a review). In addition, although our participants received different training methods across studies, they always learned Vimmi, an artificial language, as the FL (Jalbert et al., 2011). Using an artificial language allowed us to equate the participants in their knowledge of the language of learning. After reviewing different common aspects through our experimental series, we continue to integrate the results of our study within the RHM model (Kroll & Stewart, 1994).

Gestures and L2 learning

The results obtained in our first line of research seem to suggest that gestures favor semantic processing and the establishment of links between the semantic system and the lexicon in L2. Different data seem to support this conclusion. On the one hand, word learning was greater in the congruent gesture condition compared to the meaningless condition. What differentiates these two conditions was the degree of semantic relationship between the meaning of the word and the gesture. Thus, in the congruent condition both the gesture and the word had a common meaning while in the meaningless condition the gesture had no meaning. Therefore, gestures and words in the congruent condition would increase semantic processing which benefited the final learning outcome. On the other hand, processing gestures congruent with the meaning of the word to be learned mitigated the negative effect associated with performing a dual task (processing gestures and words). In fact, in our studies, we observed the negative consequences of performing a double task: In the incongruent condition and in the meaningless gesture condition, the performance became worse than in the no gesture condition. However, this negative effect due to the processing of gestures turned into facilitation when these gestures were semantically congruent with the meaning of the word.

On the other hand, in our studies, we found differences between the learning of words and verbs. In general, the learning rate was higher with nouns than with verbs. However, the inclusion of a congruent gesture condition eliminated the difference associated with the learning of these two types of words. This effect again demonstrates the efficiency of L2 learning based on semantic training. First, one of the main features that makes nouns and verbs different is the greater semantic content of nouns than of verbs. This fact would explain the superiority of nouns vs. verbs during learning. Second, verbs have an intrinsic motor component in their representation (e.g., action verbs involving movement). Thus, the greatest benefit of gestures during the learning of verbs was expected (since verbs and gestures share common aspects of movement processing). Third, the lack of differences in processing verbs and nouns occurred in the congruent condition, a condition that favoured the semantic processing of the learning material.

In the research line of gesture processing during L2 learning, we evaluated also the effect of self-performed movements and movement observation in the L2 vocabulary learning process**.** In the educational system, this topic has been in the spotlight for decades. The "learning-by-doing" theory posits that self-performed actions while learning enhance the acquisition process in many educational fields (Aleven & Koedinger, 2002; Bessen, 2015). Learning-by-doing can positively affect the formation of neural networks underlying the acquisition of knowledge and the performance of many cognitive skills (Goldin-Meadow et al., 2012). Several studies have examined the differences between self-performed tasks and experimenter-performed task (Cohen, 1981; Engelkamp & Zimmer, 1983). Some of them have confirmed the importance of self-generated movements in the learning of linguistic material. The performance of motor actions favors the learning of new words and the benefit associated with the performance of movement during learning seems to be due to the formation of a motor trace that would be activated during the subsequent retrieval of information. However, other studies have found similar pattern of results when participants produce actions and when they only see actions produced by others (e.g. Rizzolatti & Craighero, 2004). Thus, although movement performance appears to improve learning, it is not clear whether a learning protocol that involves self-generated actions would have an additional benefit to the mere observation of movements.

Implementing the same learning conditions addressed in Experiments 1 and 2, we examined the role of gestures imitation vs. gestures observation while learning L2 verbs. In general, the learning performance was better in participants who imitated the gestures while learning. However, when words and gestures were congruent in meaning participants in both learning strategies reached similar levels of recall. This pattern suggested that the mere exposure to gestures is sufficient to observe the learning advantages associated to gesture processing. In previous studies, watching movements produces a pattern of activation at brain level similar to that found during actions performance (Stefan et al., 2005). This encoding facilitation is related to the addition of sensorimotor networks to the word meaning. However, not all gestures conditions are able to improve L2 acquisition. The facilitation effect was not found when participants were exposed to incongruent and meaningless gestures in the "see" learning group. The semantic mismatch between words and gestures meaning produced an interference effect, reflecting integration difficulty in working memory (Bernardis et al., 2008). In this case, the results obtained in our study suggest that gesture imitation mitigates the negative effect of meaning mismatch between gestures and words. This positive effect of self-performed gestures might be due to a reduction in working memory load during the learning process (Cook & Goldin-Meadow, 2006). Finally, the interference effect found with incongruent and meaningless gestures was reduced in the forward translation direction, which is considered as more difficult than backward translation. Hence, the facilitative effect of performing gestures is greater when task demands increase (Marstaller & Burianová, 2014).

Thus, the results obtained in this line of research suggest that the mere processing of gestures would facilitate the establishment of connections between the semantic system and the lexicon in L2. On the other hand, the performance of gestures would reduce semantic competition within the conceptual system between the meaning of the gesture and the word in the incongruent gesture condition.

L2 learning with lexical and semantic instructions

As we have discussed in different sections of this doctoral thesis, many studies have compared semantic vs. lexical learning programs for L2 vocabulary acquisition. Therefore, at this point, we want to highlight the new insights found in our work. First, the usual way to implement semantic vs. lexical training programs is through the use of picture-L2 word pairs and L1-L2 word pairs, respectively. In addition to this methodology, we implemented a semantic task (semantic categorization) and a lexical task (grapheme monitoring) during training. These tasks are effective because they replicate the pattern of effects previously seen in word comprehension tasks. For example, we found a semantic interference effect in the translation recognition task in learners under semantic vs. lexical training. As in previous studies, this pattern of results suggests that semantic training stimulates the strengthening of connections between meaning and new words in L2. Also, through our studies, the recall percentage after learning was high in both groups of training (about 80%).

Importantly, taken together, the results of our studies reveal a consistent pattern of increased semantic processing in the group that received the picture-L2 word association learning, both in comprehension (translation recognition task) and production (word translation, picture naming). For example, the semantic vs. lexical training group performed the forward translation task more slowly, suggesting that they used a semantic processing path that slowed down the translation process. In contrast, these participants were faster in naming L2 pictures indicating that they used a processing path from semantics to L2 lexicon, while the lexical group used a processing route that involved the retrieval of L1 words.

Furthermore, in our research, electrophysiological indices (modulations of the N400 component) seem to confirm the greater conceptual processing in the group of learners under semantic training, both in translation tasks and in naming tasks. Thus, the semantic group showed greater N400 amplitude when translating in the forward direction than in the backward direction, and a greater scalp distribution of this effect which suggested deeper semantic processing than that of the lexical group. Moreover, greater N400 amplitudes were found when the naming task was conducted in L2 vs. L1 in the semantic group but not in the lexical group. This outcome shows the greater difficulty associated with semantic processing in the picture-L2 word association method compared to the similar processing of the lexical group when naming pictures in L1 and L2 (possibly by retrieving the name of the picture in L1). Finally, these differences in semantic analysis due to the type of training were not only observed during the processing of single words but also when processing words embedded in sentences. Thus, the semantic group was able to make use of the context to anticipate the occurrence of new words at the end of a sentence. In contrast, the lexical training group seemed to focus on the superficial analysis of words and did not display this cloze probability effect.

On the other hand, it is important to note that the increased semantic processing in the semantic processing group was not detrimental to the functioning of other linguistic processes. Thus, both learning group exhibited similar pattern of lexical processing (e.g., similar performance in the lexical decision task). In addition, the reanalysis of information process (e.g., to check the correction of words produced in word translation and picture naming tasks) was roughly similar in both training groups as revealed by the similar LPC modulations found in both groups of training.

Finally, it should be pointed out that although the learning rate obtained in the semantic and lexical training groups was similar across studies, the learning based on conceptual processing seemed to reduce the cost associated with switching between learning sessions. Thus, when the training method was conducted on two consecutive days, the lexical training group showed a cost between the end of the first session and the beginning of the second session. This cost was not found in the semantic training group.

Conclusion

In this doctoral thesis we have compared the efficiency of different strategies for learning a L2. In general, we have observed that with different methodologies (the use of congruent, incongruent, and meaningless gestures, the use of lexical and semantic tasks) the learners are able to achieve good levels of learning at the end of the training. However, the cognitive structure underlying these types of learning varies. A fluent bilingual is defined by the richness of semantic processing. When an expert bilingual wants to express in a L2, she would produce the communicative intention (meaning) directly with words in that language (L2 lexicon). In our study, we have seen that some learning strategies are better than others in promoting semantic processing and, consequently, fostering the type of cognitive processing present in fluent bilinguals. The inclusion of gestures congruent with the words to be learned, the performance of gestures during learning, the use of associations between pictures and words, the use of semantic categorization tasks; all these procedures converge on the same point: to increase semantic processing during the acquisition of vocabulary in a L2.

CHAPTER VII

Future Research Questions

A total of six different experiments have been presented in this thesis. However, there is still a lot of work to do. There were ideas since the beginning of the process which could not be carried out and, in addition, these experimental series made raise new questions.

Regarding the first of the experimental series about the role of gestures in L2 learning, one more study was developed and carried out. After reading an interesting paper about changes in time-frequency oscillations as a result of new words learning (Bakker et al., 2015) and seeing the opportunity of having my international research stay, we contacted the Max Plant University of Nijmegen, in Holland. We proposed a new experiment to perform during the stay. We wanted to see oscillatory consequences of learning words coupled with gestures. We made specific predictions taking into account the results obtained in our experiments and the information provided in this paper. This work has not been included in this dissertation because these data are not still fully analyzed and interpreted. As my thesis director used to say, “We can't include everything; we have to cut at some point”. However, I can already say that the preliminary analyses reveal promising results.

Following with this research line, during the writing of the paper addressing Experiments 1 and 2, some ideas appeared. In the congruent condition, the motor trace associated with the gesture was part of the meaning of the word. Hence, semantic processing was enhanced, which promoted the acquisition of L2 words. The facilitation effect observed with congruent gestures could also be explained by the motor-trace perspective (Engelkamp & Zimmer, 1984, 1985). According to this view, familiar gestures would have a strong motor trace representation in memory, which would favor L2 learning. A way of dissociating between the motor-trace and the motor-imagery account would be the use of beats and deictic gestures. Previous studies have evaluated the role of these gestures in L2 learning (Kushch et al., 2018; Morett, 2014; So et al., 2012). For example, Morett showed that the use of beat and deictic gestures facilitated the recall of words in a language not familiar to the participants. Beats and deictic gestures are typically produced with language and they have meaning associated with them. However, their meaning may not be part of a word's meaning (particularly for beats). Therefore, the comparison between iconic gestures versus beats would be of interest to dissociate between theoretical perspectives. A facilitation effect with beats would favor the motor-trace theory, and a greater facilitation effect in case of iconic gestures would be explained by the motor-imagery account.

In the second experimental line about the comparison of learning strategies based on lexical or semantic instructions, some ideas also emerged. Our study suggested that the use of gestures congruent in meaning with the word to be learned enhances the acquisition process. Other methods based on conceptual analysis of the material foster L2 learning. As in our semantic training, pairing new words with pictures denoting its meaning enhance the semantic processing of information. In the future, it would be interesting to evaluate whether these two supporting materials might have additive effects on L2 vocabulary acquisition. Or even it would be interesting to compare the effect of both learning methodologies to directly observe the magnitude of their effectiveness.

Another possible and interesting option could be to evaluate the effect of both training types on the learning of abstract words. In our studies, we implement concrete and familiar nouns as learning material. These types of words are semantically rich and methods as the picture-word association could easily affect then.

However, abstract words have proven to have a lesser conceptual information associated to their semantics. It would be interesting to observe the effect of the lexical-semantic training methods in the learning of abstract words. This work would be very relevant but the selection of experimental material would be challenging. It is not easy to find pictures clearly associated to abstract words.

Considering the results obtained in our EEG study, designing an experiment including different learning and evaluation sessions, would give us more information about the development of the learning process. Evaluating participants 24 hours after the first learning session could give us information about the consequences of the consolidation effect depending on the learning strategy. In addition, exposing participants to the learning material through more learning sessions and to establish different evaluation sessions after training would show which of the learning strategies create more perdurable memories.

After checking the effects of lexical-semantic methods with isolated words, we evaluated the electrophysiological correlates of these training procedures. After that, we observed that differences between learning procedures were also present when words were included in sentence contexts. Following a coherent succession of our studies, it would be interesting to evaluate the impact of lexical and semantic learning methods in sentence reading with electrophysiological measures.

Capítulo VIII

Resumen y Conclusiones

El campo de investigación sobre el bilingüismo ha despertado siempre gran interés, pero ha cobrado especial importancia en los últimos años (Kroll et al., 2014; Kroll & Tokowicz, 2005). Hoy en día, nos encontramos envueltos en sociedades multiculturales. Este fenómeno se encuentra en acuciante auge como consecuencia de diferentes factores sociales y demográficos. Por ejemplo, cada vez más personas que buscan sus primeros trabajos o condiciones laborales más favorables, se ven empujadas a buscar empleo fuera de las fronteras de sus países de origen o se les requiere hablar varios idiomas para acceder a puestos de trabajo. En cualquier caso, para desenvolvernos plenamente en estas sociedades cambiantes de las que somos partícipes, es necesario que adquiramos diferentes destrezas y el poder comunicarnos eficientemente cobra gran importancia. Por este motivo, los investigadores del campo del bilingüismo llevan varias décadas buscando estrategias que faciliten este proceso de aprendizaje de lenguas (Comesaña et al., 2009; Goldin-Meadow et al., 1992; Macedonia et al., 2011; Macedonia & Kriegstein, 2012; Tonzar et al., 2009).

En primer lugar, cabe destacar el importante papel de los profesores en general, y más concretamente en este caso, de los profesores de lenguas. Está en sus manos elegir las metodologías de enseñanza que mejor se adapten a sus alumnos y sus materias, y no hay mejor manera de reciclarse que mantenerse al día sobre nuevas estrategias de enseñanza basadas en investigación (Maclellan, 2015).

El tema de aprendizaje de lenguas no afecta únicamente a niños en edad escolar. Aunque a día de hoy en muchos países del mundo los niños son educados desde sus primeros años de escolarización en el aprendizaje de otras lenguas diferentes de su lengua materna, también los adultos, que no tuvieron esta oportunidad, se ven necesitados de adquirir estas habilidades.

En este ámbito se ha centrado nuestra investigación. Dentro de la gran variedad de temáticas que abarca el bilingüismo, nos hemos propuesto estudiar qué estrategias son capaces de facilitar y cuáles pueden perjudicar al aprendizaje de una segunda lengua en población adulta joven. Tomando como participantes monolingües hablantes de español, se han desarrollado diferentes experimentos en los que se ha empleado como material de aprendizaje una lengua artificial llamada Vimmi. Por tanto, en nuestros estudios, adultos monolingües de Español aprendieron Vimmi como segunda lengua. Esta lengua, a pesar de ser artificial, respeta las normas lingüísticas de las lenguas latinas. Teniendo clara la población a la que destinaríamos la investigación y el material que usaríamos como segunda lengua, se plantearon dos líneas de investigación paralelas, centradas en dos campos concretos de estrategias que afectan al aprendizaje de vocabulario en una nueva lengua. En primer lugar, nos plantemos investigar el rol del uso de gestos durante el aprendizaje de vocabulario, y en una segunda línea de trabajo, comparamos los resultados del aprendizaje basado en estrategias léxicas y semánticas.

Una base fundamental para entender esta investigación, es conocer uno de los modelos de procesamiento de segunda lengua más aceptados. El *Revised Hierarchical Model* fue propuesto en 1994 por Kroll y Stewart. Según sus directrices, cabría destacar tres nodos principales que tienen un papel fundamental en el aprendizaje de una segunda lengua. El léxico de la primera y de la segunda lengua y el sistema semántico, es decir, las palabras en la primera lengua, las palabras en la segunda lengua y los significados. Estos tres elementos principales se encontrarían interconectados y el peso de las conexiones entre estos nodos haría diferenciar entre el procesamiento de aprendices de una segunda lengua y bilingües expertos. Tanto en aprendices como en expertos, el léxico de la primera lengua estaría fuertemente conectado con la semántica. Por otro lado, en aprendices, las conexiones entre el léxico de la segunda lengua y la semántica serían débiles y emplearían preferiblemente conexiones léxicas entre la primera y la segunda lengua durante el procesamiento. A medida que avanza el proceso de aprendizaje de la segunda lengua, los aprendices dejan de depender de las conexiones léxicas entre lenguas y empiezan a fortalecer la ruta semántica directa que une las palabras en la segunda lengua con los conceptos. De esta forma, el peso de las conexiones léxicas de las dos lenguas con la información semántica en expertos sería similar (ver Figura 2, página 28).

Una vez descrito un modelo base para entender cómo las personas representan palabras en una segunda lengua, pasamos a describir las dos líneas de investigación desarrolladas en esta tesis doctoral. Aunque las dos líneas se han desarrollado paralelamente, se analizarán por separado para facilitar el trabajo del lector.

Aprendizaje de Vocabulario en Segunda Lengua Basado en Gestos

En las series experimentales realizadas en esta línea, los participantes aprendieron palabras en Vimmi en cuatro condiciones diferentes. En la condición sin gestos, palabras en Español eran presentadas junto con sus traducciones en Vimmi (por ejemplo, cuchara era presentada con su traducción en Vimmi, *deschoga*). En la condición congruente, las palabras en Español y su traducción en Vimmi aparecían junto con un gesto congruente (por ejemplo, el par cuchara-*deschoga* aparecía asociado al gesto de sujetar una cuchara invisible y llevarla hacia la boca). En la condición de gestos incongruentes, las palabras en Español y Vimmi aparecían asociadas a un gesto que, aunque tenía significado, no se correspondía con el de las palabras que debían aprender (por ejemplo, el par cuchara-*deschoga* aparecía junto al gesto de de encender una cerilla). Por último, en la condición de gestos sin significado, las palabras en Español y Vimmi se presentaban con un gesto que no tenía significado (por ejemplo, el par cuchara-*deschoga* aparecía junto al gesto de mover la mano desde la frente hacia la oreja) (ver Figura 3, página 84). Estas cuatro condiciones fueron incluidas con el fin de evaluar tres grandes teorías en relación al papel de los gestos en la adquisición de vocabulario en una segunda lengua. Estas teorías son: a) *the self-involvement explanation* (Helstrup, 1987), apoya que el mero hecho de hacer gestos durante el aprendizaje hace que los participantes presten más atención a la tarea y esto produciría una facilitación en el aprendizaje, b) *the motor trace account* (Engelkamp & Zimmer, 1984, 1985), sugiere que el trazo motor asociado a la realización de gestos favorecería el aprendizaje con independencia del significado de los gestos y de las palabras, y por último, c) *the motor imagery explanation* (Denis et al., 1991) indica que solo en el caso en el que los gestos y las palabras tengan significados coherentes, se facilitará el aprendizaje. Por último, se generaron predicciones específicas para cada condición experimental teniendo en cuenta las teorías.

Según *the self-involvement account*, si los gestos solo hacen que los participantes presten más atención a la tarea, todas las condiciones en las que se aprendía con gestos mejorarían el aprendizaje en comparación a la condición sin gestos. Si *the motor-trace theory* está en lo correcto, los trazos motores asociados a gestos familiares (condiciones congruente e incongruente) mejorarían el aprendizaje de nuevas palabras en relación a gestos que no son familiares para los participantes (condición de gestos sin significado). Por último, *the motor-imagery explanation* sugiere que los gestos podrían tener efectos beneficiosos o perjudicar el aprendizaje dependiendo de la convergencia entre el significado de los gestos y de las palabras. Los gestos congruentes facilitarían el aprendizaje mientras que los gestos incongruentes serían un obstáculo en el proceso de adquisición.

Por otro lado, hasta ahora la mayor parte de los estudios que abordan el papel de los gestos en la adquisición de vocabulario en una segunda lengua, emplean verbos como material de aprendizaje. Esta elección parece lógica si tenemos en cuenta que los verbos tienen un claro componente motor asociado a su semántica y, por tanto, el efecto de los gestos en el aprendizaje sería más fácilmente observable con este tipo de palabras. Tras plantear el experimento con verbos, nos preguntamos qué podría pasar durante el aprendizaje de nombres. Los nombres son las primeras palabras aprendidas cuando se sigue el desarrollo normal del lenguaje y además, en estudios previos se ha mostrado que son más fáciles de aprender en comparación con los verbos, cuando se comienza el contacto con una segunda lengua (Childers & Tomasello, 2002; Tardif et al., 1997). Por todo ello, nos pareció coherente estudiar el efecto de los gestos durante el aprendizaje de verbos y nombres. Con el fin de poder realizar comparaciones, el mismo procedimiento fue usado en ambos estudios. Dos grupos de participantes aprendieron 40 nombres o verbos en Vimmi durante tres días consecutivos. Las nuevas palabras eran presentadas siguiendo las cuatro condiciones experimentales presentadas anteriormente (no gestos, gestos congruentes, gestos incongruentes y gestos sin significado) para su aprendizaje. Cada día, tras finalizar la fase de aprendizaje, evaluamos a los participantes con una tarea de traducción directa e inversa. Las palabras se presentaban en Español (traducción directa) o en Vimmi (traducción inversa) para su traducción en voz alta.

Los resultados mostraron que el uso de gestos que son congruentes con el significado de la palabra que se ha de aprender, facilitan el aprendizaje de vocabulario en una lengua extranjera. Por otro lado, el ejecutar gestos incongruentes o sin significado parece generar una situación de tarea dual que interfiere en el aprendizaje. Aunque, en general, los verbos resultaron más difíciles de aprender que los nombres, en la condición congruente no se observaron diferencias en el aprendizaje de estos dos tipos de palabras. Lo anterior sugiere que el uso de gestos congruentes elimina la dificultad asociada al aprendizaje de verbos frente al de nombres. Si tenemos en cuenta el patrón de datos, parece que la *motor imagery account* podría ajustarse en mayor medida a nuestros resultados puesto que la congruencia entre el significado de los gestos y las palabras marcó la diferencia en el aprendizaje de palabras.

Además, quisimos investigar si el efecto beneficioso de los gestos requería que los participantes realizasen activamente los gestos o bastaba con la mera observación de los gestos realizados por otra persona. Para tal fin, dos grupos de participantes aprendieron verbos en Vimmi siguiendo cuatro condiciones experimentales (condición sin gestos, gestos congruentes, incongruentes y sin significado). El procedimiento fue similar al del estudio anteriormente descrito, pero uno de los grupos imitaba los gestos del experimentador durante el aprendizaje, mientras que otro grupo solo los observaba. Los resultados mostraron que tanto la realización como la observación de gestos tuvieron un efecto beneficioso similar. Sin embargo, la interferencia asociada a los gestos incongruentes y gestos sin significado, se veía reducida en el grupo de participantes que realizaba los gestos. Por tanto, aunque los efectos beneficiosos de los gestos congruentes se observan en similar medida en ambos tipos de aprendizaje (observación e imitación de gestos), la realización de los movimientos por parte de los participantes aporta algo adicional que disminuye la interferencia en el aprendizaje asociada a los gestos incongruentes y sin significado.

Estrategias de Aprendizaje de Segunda Lengua Basadas en Metodologías Léxicas y Semánticas

En esta serie experimental, planteamos tres experimentos. El fin era comparar cómo estrategias basadas en aprendizaje de tipo léxico o de tipo semántico afectaban al procesamiento de las nuevas palabras aprendidas en una lengua extranjera. Teniendo siempre como marco teórico el *Revised Hieralchcal Model* (Kroll & Stewart, 1994), cabría esperar que los sistemas de aprendizaje basados en métodos semánticos fortaleciesen las conexiones entre las palabras de la segunda lengua y su significado. Este patrón de procesamiento está presente en bilingües expertos y por tanto, una buena estrategia de aprendizaje podría estar basada en el establecimiento de estas conexiones. Por otro lado, métodos de aprendizaje léxico permiten establecer conexiones entre las palabras de la primera lengua y sus traducciones en la segunda lengua. En nuestro caso, seleccionamos como método de aprendizaje semántico un paradigma de asociación palabra-dibujo, en el que las nuevas palabras eran presentadas con imágenes que denotaban su significado. Además, como método de apoyo durante el proceso de aprendizaje, los participantes realizaban una tarea de categorización semántica, que requiere del procesamiento semántico de las palabras. Los participantes debían indicar si palabras que se presentaban en Vimmi pertenecían a un determinado campo semántico (por ejemplo, el nombre del campo semántico "fruta" aparecía seguido de la palabra "*raone",* plátano en Vimmi). Para la estrategia de aprendizaje léxica seleccionamos un método de asociación palabra-palabra, en que las palabras en la primera y la segunda lengua se presentaban juntas para su aprendizaje. También en este caso se incluyó una tarea de apoyo durante el aprendizaje, pero en este caso, favorecía el procesamiento léxico de las palabras. Los participantes realizaron una tarea de monitorización de grafemas, en la que debían indicar si un grafema estaba presente en tanto en la palabra en Español como en su traducción en Vimmi (por ejemplo, se presentaba el grafema "n" seguido de la palabra en Vimmi "*raone*") (ver Figura 10, página 143).

Durante el aprendizaje, el grupo de participantes sujetos al entrenamiento semántico llegó a mostrar un nivel de aprendizaje más alto y tiempos de respuesta más rápidos. El número total de exposiciones en ambos grupos fue idéntico. Tras la fase de aprendizaje, sólo los participantes que habían adquirido las nuevas palabras mediante el entrenamiento semántico mostraron un efecto de naturaleza semántica, en concreto, un efecto de interferencia. Este efecto se caracteriza por mayores tiempos de respuesta y mayor número de errores cometidos ante estímulos semánticamente relacionados. Cuando en una tarea de reconocimiento de traducción se presentan ejemplares que pertenecen a la misma categoría semántica (por ejemplo, si la traducción correcta de plátano es *raone* y se presenta la palabra pera) para que el participante decida si son traducciones correctas, los participantes tienen más problemas en comparación con una condición no relacionada, donde las palabras no pertenecen al mismo campo semántico (por ejemplo, si la palabra *roane*, plátano en Vimmi, se presenta junto a la palabra perro). Este efecto es habitual en bilingües expertos (Sunderman & Kroll, 2006; Talamas et al., 1999). Por tanto, podemos concluir que el método de aprendizaje semántico imita el patrón de procesamiento presente en bilingües expertos y potencia el establecimiento de conexiones entre las palabras de la segunda lengua y el sistema semántico.

Tras comprobar el efecto positivo de las metodologías semánticas en el aprendizaje de nuevo vocabulario, nos propusimos ir un paso más allá. El someterse a metodologías de aprendizaje léxicas o semánticas ha demostrado tener consecuencias sobre la adquisición de palabras aisladas, pero ésta no es una forma de comunicación habitual y normalmente, formamos oraciones con estas palabras para transmitir significados complejos. Por tanto, nos propusimos comprobar los efectos de ambos tipos de aprendizaje sobre el procesamiento de las nuevas palabras dentro de oraciones.

Siguiendo el mismo tipo de entrenamiento presentado anteriormente (asociación palabra-dibujo con tarea de categorización o asociación palabra-palabra con tarea de monitorización de grafemas), los participantes aprendieron nuevas palabras de forma léxica o semántica. Tras el entrenamiento, los participantes realizaron tareas de evaluación con palabras aisladas y en el contexto de oraciones. En concreto, para la evaluación de las palabras de forma aislada, se emplearon tareas de decisión léxica, traducción directa e inversa y nombrado de imágenes. En la tarea de decisión léxica, los participantes indicaban si palabras presentadas eran palabras existentes en Español o Vimmi o no palabras. En la tarea de traducción, palabras en Español o Vimmi eran presentadas para su traducción en voz alta (traducción directa o inversa respectivamente). Por último, la tarea de nombrado de imágenes consistía en la presentación de dibujos que representaban las palabras que los participantes habían aprendido en Vimmi para su nombrado en voz alta. Para evaluar el procesamiento de las nuevas palabras dentro del contexto de las oraciones, se diseñó una tarea de procesamiento semántico en oraciones. En esta tarea, se presentaban frases en español en las que la última palabra aparecía en Vimmi, y era una de las palabras aprendidas en el entrenamiento. Una vez leían la oración completa, los participantes debían indicar si era semánticamente correcta o no (por ejemplo, A Pedro le costó mucho cortar el filete porque no estaba bien afilado el *ruzanego*, cuchillo en Vimmi, sería semánticamente correcta y, A Pedro le costó mucho cortar el filete porque no estaba bien afilado el *boreda*, burro en Vimmi, sería semánticamente incorrecta). Además, se manipuló el *cloze probability*, que es la probabilidad de que una oración acabe con una palabra en concreto (Khachatryan et al., 2014; Staub et al., 2015). Por tanto, había oraciones predecibles (por ejemplo, A Pedro le costó mucho cortar el filete porque no estaba bien afilado el...) y no predecibles (por ejemplo, Cuando llegaron al río, Pedro miró en su mochila y no encontró el…).

Teniendo en cuenta estudios sobre el papel modulador del contexto en el procesamiento de una segunda lengua, se elaboraron diferentes predicciones (Foucart et al., 2014; Schwartz & Kroll, 2006). Según estos estudios, los bilingües expertos son capaces de hacer uso del contexto semántico para anticipar palabras dentro de la oración. En nuestro estudio, hipotetizamos que los participantes que aprendiesen con el método semántico serían capaces de imitar este patrón de anticipación, especialmente en oraciones predecibles porque el contexto predecible de la oración maximizaría el uso de las conexiones entre las palabras de la segunda lengua y el sistema semántico que habían fortalecido durante el aprendizaje. Por el contrario, este efecto se vería atenuado en el caso de los participantes que aprendían con métodos léxicos porque ellos usarían de forma preferente las conexiones entre la primera y la segunda lengua que habrían practicado durante el entrenamiento.

Además, con el fin observar los efectos de periodos de consolidación de aprendizaje, los participantes realizaron dos sesiones de entrenamiento en dos días consecutivos. La demora entre sesiones de entrenamiento puede generar tanto costes como beneficios en el proceso de adquisición (Rickard, 2007). Los resultados obtenidos durante el aprendizaje replicaron los datos del primer estudio de esta serie experimental, donde participantes sometidos al aprendizaje de tipo semántico (asociación dibujo-palabra) respondieron más rápidamente y cometieron menos errores que participantes que aprendían con la metodología léxica (asociación palabra-palabra). Además, mostraron un interesante patrón de procesamiento de las nuevas palabras dentro de las oraciones. Los resultados indicaron que el grupo de participantes sometido al entrenamiento léxico, fortaleció las conexiones entre las palabras de la primera y la segunda lengua y empleó estas conexiones tanto en tareas con palabras aisladas (tarea de decisión léxica, traducción directa e inversa y nombrado de imágenes), como en el contexto de oraciones (tarea de procesamiento semántico dentro de oraciones). Por el contrario, el grupo de entrenamiento semántico, mostró patrones de resultados acordes con el fortalecimiento de conexiones entre las palabras aprendidas y su semántica. Los resultados indicaron que mientras el grupo de entrenamiento semántico se veía favorecido por el contexto predecible de las oraciones, el grupo léxico presentaba una mejor ejecución en casos en los que no era predecible el contenido. Estudios previos han mostrado que contextos de baja predictibilidad favorecen el uso de las conexiones léxicas entre lenguas (Schwartz & Kroll, 2006). Por tanto, los aprendices léxicos procesaron más fácilmente las palabras en un contexto que favorecía el uso de las conexiones con las que ellos habían entrenado.

Por otro lado, investigaciones anteriores han mostrado que en algunas ocasiones, las medidas comportamentales no son lo suficientemente finas como para captar diferencias iniciales que tienen lugar al comienzo del proceso de aprendizaje de una segunda lengua (McLaughlin et al., 2004; Tokowicz & MacWhinney, 2005). En este contexto, tomar medidas electrofisiológicas puede aportar información adicional. La actividad eléctrica en la superficie del cuero cabelludo puede medirse y asociarse en el tiempo a la aparición de los estímulos de interés. Este registro se conoce como medidas de potenciales relacionados con eventos y puede informar sobre procesamiento neural subyacente durante la presentación de los estímulos. Gracias al gran número de estudios que emplean este tipo de medidas, se han identificado componentes concretos asociados a procesos cerebrales específicos.

En el último de nuestros estudios, se tomaron medidas electrofisiológicas mientras dos grupos de participantes, sometidos previamente al entrenamiento léxico o semántico anteriormente descritos (entrenamiento léxico o semántico), realizaban tareas de traducción entre lenguas (traducción directa e inversa) y de nombrado de imágenes. Nuestra hipótesis era que si efectivamente, el método de aprendizaje semántico generaba patrones de procesamiento más próximos al de los bilingües expertos, tras una sola sesión de entrenamiento, los participantes que aprendieron con la asociación de imágenes a palabras, mostrarían una mayor modulación de componentes relacionados con el procesamiento léxico-semántico de las palabras en la segunda lengua y el patrón de procesamiento sería más similar al de los bilingües. En concreto, centramos nuestra investigación en dos componentes cerebrales, el N400 y el *late positive component* (LPC).

El componente N400 se relaciona con una negatividad que aparece unos 400 milisegundos después de la presentación de los estímulos y que es sensible al procesamiento léxico-semántico de la información (Kutas & Hillyard, 1980). En estudios previos, la modulación de este componente se ha considerado como un índice que indica el nivel de dominio en la segunda lengua. En bilingües expertos, se observa mayor amplitud del componente N400 ante palabras presentadas en la segunda lengua que en aprendices de la segunda lengua (Midgley et al., 2009). Por su parte, el LPC se corresponde con una positividad duradera en el tiempo que aparece alrededor de los 500 ms después de la presentación del estímulo. Se asocia a procesos de monitorización y de reanálisis de información y su amplitud aumenta en función de la dificultad de procesamiento de los estímulos (Jackson et al., 2001; Kieffaber et al., 2013).

Los resultados mostraron que, efectivamente, el grupo de entrenamiento semántico, mostró un patrón de respuesta electrofisiológica más similar a los bilingües. Se dejaba ver un patrón de acceso semántico temprano incluso tras una única sesión de aprendizaje. Sin embargo, el grupo de participantes que aprendió con la asociación de palabras en la primera y la segunda lengua, mostró un patrón de procesamiento de las nuevas palabras menos distribuido y que además, se llevaba a cabo a través de la información léxica de la primera lengua.

Conclusión

Como se ha mostrado en los anteriores experimentos y como muestra la investigación previa en general, el modo en el que aprendemos determina el establecimiento de conexiones entre las palabras de la primera y la segunda lengua y el sistema semántico a nivel cerebral y juega un papel fundamental en la adquisición de una segunda lengua. Como hemos comprobado, las metodologías semánticas fortalecen el establecimiento de conexiones directas entre las palabras de la segunda lengua y el sistema conceptual, mientras que los entrenamientos léxicos fomentan el establecimiento de conexiones entre los léxicos de las dos lenguas.

Todos estos datos, nos llevan a pensar que metodologías de enseñanza basadas en la ampliación de la información semántica asociada a las nuevas destrezas, facilitan el aprendizaje y ocasionan diferentes ventajas a muchos niveles. En nuestras series experimentales, cuando monolingües de español deben aprender palabras en una nueva lengua, el proceso se ve favorecido por el uso de gestos congruentes con su significado o por la asociación de las nuevas palabras a imágenes que denotaban su significado. En ambos casos, el proceso de aprendizaje ha demostrado fortalecer las conexiones semánticas de estas nuevas palabras y los participantes han mostrado un patrón de procesamiento más próximo al de bilingües expertos.

Como conclusión, dadas las características de las sociedades multiculturales en las que nos encontramos hoy en día, hablar varias lenguas se hace especialmente necesario para poder desenvolvernos de manera efectiva. No solo los niños, si no también adultos, se ven necesitados de adquirir estas habilidades, y por tanto, ya sea dentro de la educación obligatoria, postobligatoria o en academias de idiomas, los docentes deben emplear metodologías de enseñanza basadas en evidencia científica que faciliten el proceso de adquisición de nuevas lenguas.

References

Adams, J. A. (1952). Warm-up decrement in performance on the pursuit-rotor. *The American Journal of Psychology, 65*(3), 404-414. doi: 10.2307/1418761

Aleven, V. A., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science, 26*(2), 147-179. doi: 10.1207/s15516709cog2602_1

Alibali, M. W., Flevares, L. M., & Goldin-Meadow, S. (1997). Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of Educational Psychology, 89*(1), 183-193. doi: 10.1037/0022-0663.89.1.183

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, *24*(4), 477-492. doi: 10.3758/BF03200936

Altarriba, J., & Mathis, K. M. (1997). Conceptual and lexical development in second acquisition. *Journal of Memory and Language, 36*(4), 550-568. doi:10.1006/jmla.1997.2493

Anderson, J., Fincham, J., & Douglass, S. (1999). Practice and retention: A unifying analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(5), 1120-1136. doi: 10.1037/0278-7393.25.5.1120

Asher, J. J., & Price, B. S. (1967). The learning strategy of the total physical response: Some age differences. *Child Development, 38*(4), 1219-1227. doi: 10.2307/1127119

Atkinson, R. C. (1975). Mnemotechnics in second-language learning. *American Psychologist, 30*(8), 821-826. doi: 10.1037/h0077029

Atkinson, R. C., & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory*, *104*, 126-133. doi: 10.1037/0278-7393.1.2.126

Bäckman, L., Nilsson, L. G., & Chalom, D. (1986). New evidence on the nature of the encoding of action events. *Memory & Cognition*, *14*(4), 339-346. doi: 10.3758/BF03202512

Bakker, I., Takashima, A., Van Hell, J. G., Janzen, G., & McQueen, J. M. (2015). Changes in theta and beta oscillations as signatures of novel word consolidation. *Journal of Cognitive Neuroscience*, *27*(7), 1286-1297. doi: /10.1162/jocn_a_00801

Barbieri, F., Buonocore, A., Dalla Volta, R., & Gentilucci, M. (2009). How symbolic gestures and words interact with each other. *Brain and Language, 110*(1), 1-11. doi:10.1016/j.bandl.2009.01.002

Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning, 52*(2), 323-363. doi: 10.1111/0023-8333.00186

Battig, W. F., & Montague, W. E. (1969). Category norms of verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology, 80*(3), 1-46. doi: 10.1037/h0027577

Beardsmore, H. B. (1986). *Bilingualism: basic principles*. Clevedon: Multilingual Matters.

Beattie, G., & Coughlan, J. (1999). An experimental investigation of the role of iconic gestures in lexical access using the tip-of-the-tongue phenomenon. *British Journal of Psychology, 90*(1), 35-56. doi: 10.1348/000712699161251

Beatty-Martínez, A. L., & Dussias, P. E. (2017). Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language*, *95*, 173-189. doi: 10.1016/j.jml.2017.04.002

Bennett, N., & Desforges, C. (1988). Matching classroom tasks to students' attainments. *The Elementary School Journal, 88*(3), 221-234. doi: 10.1086/461535

Bernardis, P., & Gentilucci, M. (2006). Speech and gesture share the same communication system. *Neuropsychologia, 44*(2), 178-190. doi: 10.1016/j.neuropsychologia.2005.05.007

Bernardis, P., Salillas, E., & Carameli, N. (2008). Behavioural and neuropsychological evidence of semantic interaction between iconic gestures and words. *Cognitive Neuropsychology, 25*(7)*,* 1114-1128. doi:10.1080/02643290801921707

Bessen, J. (2015). *Learning by doing: The real connection between innovation, wages, and wealth*. New Haven: Yale University Press.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417-444. doi: 10.1146/annurev-psych-113011-143823

Blackford, T., Holcomb, P. J., Grainger, J., & Kuperberg, G. R. (2012). A funny thing happened on the way to articulation: N400 attenuation despite behavioral interference in picture naming. *Cognition*, *123*(1), 84-99. doi: 10.1016/j.cognition.2011.12.007

Block, C. K., & Baldwin, C. L. (2010). Cloze probability and completion norms for 498 sentences: Behavioral and neural validation using event-related potentials. *Behavior Research Methods*, *42*(3), 665-670. doi: 10.3758/BRM.42.3.665

Bloom, P. A., & Fischler, I. (1980). Completion norms for 329 sentence contexts. *Memory & Cognition*, *8*(6), 631-642. doi:10.3758/BF03213783

Bloomfield, L. (1933). *Language.* New York: Holt, Rinehart & Winston.

Bogulski, C. A., Bice, K., & Kroll, J. F. (2018). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. *Bilingualism: Language and Cognition*, 1-16. doi: 10.1017/S1366728918000858

Botting, N., Riches, N., Gaynor, M., & Morgan, G. (2010). Gesture production and comprehension in children with specific language impairment. *British Journal of Developmental Psychology, 28*(1), 51-69. doi: 10.1348/026151009X482642

Boulenger, V., Hauk, O., & Pulvermüller, F. (2009). Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. *Cerebral Cortex, 19*(8), 1905-1914. doi: 10.1093/cercor/bhn217

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, *114*(1), 80-99. doi: 10.1037/0033-2909.114.1.80

Bransdorfer, R. L. (1991). *Communicative value and linguistic knowledge in second language oral input processing* (Doctoral dissertation). Retrieved from Illinois Digital Environment for Access to Learning and Scholarship. (UMIAAI9136551).

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism: Language and Cognition, 13*(3), 359-371. doi:10.1017/S1366728909990344

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., ... & Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, *13*(2), 400-404. doi: 10.1111/j.1460-9568.2001.01385.x

Butcher, C., & Goldin-Meadow, S. (2000). Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill (Ed.)*, Language and gesture* (pp. 235-258). Cambridge: Cambridge University Press.

Caramazza, A., & Brones, I. (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society*, *13*(4), 212-214. doi: 10.3758/BF03335062

Carels, P. (1981). Pantomime in the foreign language classroom. *Foreign Language Annals, 14*(5), 407-411. doi: 10.1111/j.1944-9720.1981.tb01660.x

Cassell, J., McNeill, D., & McCullogh, K. (1999). Speech-gesture mismatches: Evidence for one underlying representation of linguistic & nonlinguistic information. *Pragmatics & Cognition, 6*(2), 1-34. doi: 10.1075/pc.7.1.03cas

Cenoz, J. (2013). The influence of bilingualism on third language acquisition: Focus on multilingualism. *Language Teaching*, *46*(1), 71-86. doi: 10.1017/S0261444811000218

Chauncey, K., Grainger, J., & Holcomb, P. J. (2008). Code-switching effects in bilingual word recognition: A masked priming study with event-related potentials. *Brain and Language*, *105*(3), 161-174. doi: 10.1016/j.bandl.2007.11.006

Chen, H. C. (1990). Lexical processing in a non-native language: Effects of language proficiency and learning strategy. *Memory & Cognition*, *18*(3), 279-288. doi: 10.3758/BF03213881

Chen, H. C., & Leung, Y. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(2), 316-325. doi: 10.1037/0278-7393.15.2.316

Chieffi, S., Secchi, C., & Gentilucci, M., (2009). Deictic word and gesture production: Their interaction. *Behavioural Brain Research, 203*(2), 200-206. doi:10.1016/j.bbr.2009.05.003

Childers, J., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology, 38*(6), 967-978. doi: 0.1037//0012-1649.38.6.967

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research, 1147*, 192-208. doi: 10.1016/j.brainres.2007.01.137

Christoffels, I. K., Ganushchak, L., & Koester, D. (2013). Language conflict in translation: An ERP study of translation production. *Journal of Cognitive Psychology*, *25*(5), 646-664. doi: 10.1080/20445911.2013.821127

Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General, 137*(4), 706-723. doi: 10.1037/a0013157

Cohen, R. L. (1981). On the generality of some memory laws. *Scandinavian Journal of Psychology*, *22*(1), 267-281. doi: 10.1111/j.1467-9450.1981.tb00402.x

Cohen, A. D. (2014). *Strategies in learning and using a second language*. New York: Routledg.

Comesaña, M., Perea, M., Piñeiro, A., & Fraga, I. (2009). Vocabulary teaching strategies and conceptual representations of words in L2 in children: Evidence with novice learners. *Journal of Experimental Child Psychology, 104*(1), 22-33. doi: 10.1016/j.jecp.2008.10.004

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

Comesaña, M., Soares, A. P., Sánchez-Casas, R., & Lima, C. (2012). 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. doi: 10.1111/j.2044-8295.2011.02080.x

Cook, S. W., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development, 7*(2), 211-232. doi: 10.1207/s15327647jcd0702_4

Cook, S. W., Yip, T. K., & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and Cognitive Processes, 27*(4), 594-610. doi: 10.1080/01690965.2011.567074

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language, 50*(4), 491-511. doi: 10.1016/j.jml.2004.02.002

Coulson, S., Urbach, T. P., & Kutas, M. (2006). Looking back: Joke comprehension and the space structuring model. *Humor*, *19*(3), 229-250. doi: 10.1515/HUMOR.2006.013

Coyle, D. (2001). Content and language integrated learning motivating learners and teachers. *Scottish Languages Review*, *13*, 1-18. doi: 10.2167/beb459.0

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*, 268-294. doi: 10.1037/0096-3445.104.3.268

Cristoffanini, P., Kirsner, K., & Milech, D. (1986). Bilingual lexical representation: The status of Spanish-English cognates. *The Quarterly Journal of Experimental Psychology Section A*, *38*(3), 367-393. doi: 10.1080/14640748608401604

Crystal, D. (2003). *English as a Global Language*. Cambridge: Cambridge University Press.

Cuetos, F., Glez-Nosti, M., Barbon, A., & Brysbaert, M. (2011). Spanish word frequencies based on film subtitles. *Psicológica*, *32*(2), 133-144.

Davis, C., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods, 37*(4), 665-671. doi: 10.3758/BF03192738

Deacon, D., Dynowska, A., Ritter, W., & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology, 41*(1), 60-74. doi: 10.1111/1469-8986.00120

De Grauwe, S., Willems, R. W., Rueschermeyer, S., Lemhöfer, H. S., & Schriefers H. (2014). Embodied language in first- and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia, 56*, 334-349. doi:10.1016/j.neuropsychologia.2014.02.003

De Groot, A. M., & Nas, G. (1991). No lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, *30*(1), 90-123. doi: 10.1016/0749-596X(91)90012-9

De Groot, A. M., & Poot, R. (1997). Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language Learning, 47*(2), 215-264. doi:10.1111/0023-8333.71997007

Denis, M., Engelkamp, J., & Mohr, G. (1991). Memory of imagined actions: Imagining oneself or another person. *Psychological Research, 53*(3), 246-250. doi: 10.1007/bf00941394

de Ruiter, J. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and gesture* (pp. 284-311). Cambridge: Cambridge University Press.

Digman, J. M. (1959). Growth of a motor skill as a function of distribution of practice. *Journal of experimental psychology, 57*(5), 310-316. doi: 10.1037/h0045049

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*(4), 496-518. doi: 10.1006/jmla.1999.2654

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175-197. doi:10.1017/S1366728902003012

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition, 1*(1), 51-66. doi: 10.1017/S1366728998000121

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short Grit Scale (Grit-S). *Journal of Personality Assessment*, *91*(2), 166-174. doi: 10.1080/00223890802634290

Dufour, R., & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition, 23*(2), 166-180. doi: 10.3758/BF03197219

Duyck, W., & Brysbaert, M. (2004). Forward and backward number translation requires conceptual mediation in both balanced and unbalanced bilinguals. *Journal of Experimental Psychology: Human Perception and Performance, 30*(5), 889-906. doi: 10.1037/0096-1523.30.5.889

Duyck, W., & Brysbaert, M. (2008). Semantic access in number word translation: The role of crosslingual lexical similarity. *Experimental Psychology, 55*(2), 102-112. doi: 10.1027/1618-3169.55.2.102

Elgort, I., Perfetti, C. A., Rickles, B., & Stafura, J. Z. (2015). Contextual learning of L2 word meanings: Second language proficiency modulates behavioural and event-related brain potential (ERP) indicators of learning. *Language, Cognition and Neuroscience*, *30*(5), 506-528. doi: 10.1080/23273798.2014.942673

Ellis, N., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning, 43*(4), 559-617. doi: 10.1111/j.1467-1770.1993.tb00627.x

Elorza, I., & Muñoa, I. (2008). Promoting the minority language through integrated plurilingual language planning: The case of the Ikastolas. *Language, Culture and Curriculum*, *21*(1), 85-101. doi: 10.2167/lcc345.0

Engelkamp, J., & Zimmer, H. D. (1983). Zum Einflu–S von Wahrnehmen und Tun auf das Behalten von Verb-Objekt-Phrasen. *Sprache & Kognition, 2*(2), 117-127*.*

Engelkamp, J., & Zimmer, H. D. (1984). Motor program information as a separable memory unit. *Psychological Research, 46*(3), 283-299. doi: 10.1007/BF00308889

Engelkamp, J., & Zimmer, H. D. (1985). Motor programs and their relation to semantic memory. *German Journal of Psychology, 9*(3)*,* 239-254.

Engelkamp, J., Zimmer, H. D., Mohr, G., & Sellen, O. (1994). Memory of self-performed tasks: Self-performing during recognition. *Memory & Cognition*, *22*(1), 34-39. doi: 10.3758/BF03202759

Eurostat (2019). Foreign language skills statistics. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php/Foreign _language _skills_statistics

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191.doi: 10.3758/BF03193146

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development, 64*(3), 637-656. doi: 10.1111/j.1467-8624.1993.tb02933.x

Ferré, P., Sánchez-Casas, R., & Guasch, M. (2006). Can a horse be a donkey? Semantic and form interference effects in translation recognition in early and late proficient and nonproficient Spanish-Catalan bilinguals. *Language Learning, 56*(4), 571-608. doi: 10.1111/j.1467-9922.2006.00389.x

Feyereisen, P. (2006). Further investigation on the mnemonic effect of gestures: Their meaning matters. *European Journal of Cognitive Psychology, 18*(2), 185-205.doi: 10.1080/09541440540000158

Finkbeiner, M., & Nicol, J. (2003). Semantic category effects in second language word learning. *Applied Psycholinguistics, 24*(3)*,* 369-383. doi: 10.1017/S0142716403000195.

Foucart, A., Martin, C. D., Moreno, E. M., & Costa, A. (2014). Can bilinguals see it coming? Word anticipation in L2 sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(5), 1461-1469. doi: 10.1037/a0036756

Freed, B. F. (1995). *Second Language Acquisition in a Study Abroad Context*. Amsterdam/Philadelphia: John Benjamins Publishing.

Freed, B. F., Dewey, D. P., Segalowitz, N., & Halter, R. (2004). The language contact profile. *Studies in Second Language Acquisition, 26*(2), 349-356. doi: 10.1017/S027226310426209X

Frenck-Mestre, C., & Prince, P. (1997). Second language autonomy. *Journal of Memory and Language, 37*(4), 481-501. doi: 10.1006/jmla.1997.2526

Ganis, G., Keenan, J. P., Kosslyn, S. M., & Pascual-Leone, A. (2000). Transcranial magnetic stimulation of primary motor cortex affects mental rotation. *Cerebral Cortex*, *10*(2), 175-180. doi: 10.1093/cercor/10.2.175

García-Gámez, A. B., & Macizo, P. (2019). Learning nouns and verbs in a foreign language: The role of gestures. *Applied Psycholinguistics, 40*(2), 473-507. doi: 10.1017/S0142716418000656

García-Gámez, A. B., & Macizo, P. (2019). The way in which foreign words are learned determines their use in isolation and within sentences. *Journal of Experimental Psychology: Learning, Memory and Cognition.* Advance online publication. doi: 10.1037/xlm0000721

Genesee, F. (1987). *Learning through two Languages: Studies of Immersion and Bilingual Education*. New York: Newbury house publishers.

Genesee, F. (2014). Second language learning in school settings: Lessons from immersion. In A. G. Reynols (Ed.), *Bilingualism, multiculturalism, and second language learning* (pp. 203-222). New York: Psychology Press.

Gentner, D. (1981). Some interesting differences between verbs and nouns. *Cognition and Brain Theory, 4*(2), 161-178.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.), *Language development* (pp. 301-334). New York: Lawrence Erlbaum.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215-256). Cambridge: Cambridge University Press.

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

Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology*, *96*(3), 424-436. doi: 10.1037/0022-0663.96.3.424

Glenberg, A. M., Sato, M., & Cattaneo, L. (2008). Use-induced motor plasticity affects the processing of abstract and concrete language. *Current Biology*, *18*(7), R290-R291. doi: 10.1016/j.cub.2008.02.036

Goldfield, B. A. (1993). Noun bias in maternal speech to one-year-olds. *Journal of Child Language, 20*(1), 85-99. doi: 10.1017/S0305000900009132

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences*, *3*(11), 419-429. doi: 10.1016/S1364-6613(99)01397-2

Goldin-Meadow, S. (2000). Beyond words: The importance of gesture to researchers and learners. *Child Development, 71*(1), 231-239. doi: 10.1111/1467-8624.00138

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology, 64,* 257-283. doi: 10.1146/annurev-psych-113011-143802

Goldin-Meadow, S., Levine, S. C., Zinchenko, E., Yip, T. K., Hemani, N., & Factor, L. (2012). Doing gesture promotes learning a mental transformation task better than seeing gesture. *Developmental Science, 15*(6), 876-884. doi: 10.1111/j.1467-7687.2012.01185.x

Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing knowledge through gesture: Using children's hands to read their minds. *Cognition and Instruction, 9*(3), 201-219. doi: 10.1207/s1532690xci0903_2

Gopnik, A., & Choi, S. (1995). Names, relational words, and cognitive development in English and Korean speakers: Nouns are not always learned before verbs. In M. Tomasello & W. E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 63-80). Hillsdale, NJ: Lawrence Erlbaum Associates.

Grainger, J., Kiyonaga, K., & Holcomb, P. J. (2006). The time course of orthographic and phonological code activation. *Psychological Science, 17*(12), 1021-1026. doi: 10.1111/j.1467-9280.2006.01821.x

Green, D. W., & Wei, L. (2014). A control process model of code-switching. *Language, Cognition and Neuroscience, 29*(4), 499-511. doi: 10.1080/23273798.2014.882515

Green, D. W., & Wei, L. (2016). Code-switching and language control. *Bilingualism: Language and Cognition*, *19*(5), 883-884. doi: 10.1017/S1366728916000018

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95-112. doi: 10.1007/bf02289823

Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual & Multicultural Development*, *6*(6), 467-477. doi: 10.1080/01434632.1985.9994221

Gullberg, M. (2014). Gestures and second language acquisition. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill & S. Tessendorf (Eds.), *Body language communication: An international handbook on multimodality in human interaction* (pp. 1868-1875). Amsterdam: De Gruyter Mouton.

Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs Reveal Early Autonomy and Late Interaction. *Journal of Cognitive Neuroscience*, *12*(4), 556-568. doi: 10.1162/089892900562336

Guo, T., Misra, M., Tam, J. W., & Kroll, J. F. (2012). On the time course of accessing meaning in a second language: An electrophysiological and behavioral investigation of translation recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(5), 1165-1186. doi: 10.1037/a0028076

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research, 30*(3), 251-266. doi: 10.1023/A:101049091

Hamers, J. F., & Blanc, M. (1983). *Bilingualité et bilinguisme*. Bruxelles: Editions Mardaga.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*(2), 307-322. doi:10.1016/S0896-6273(03)00838-9

Hauk, O., Rockstroh, B., & Eulitz, C. (2001). Grapheme monitoring in picture naming: an electrophysiological study of language production. *Brain Topography, 14*(1), 3-13. doi: 10.1023/A:1012519104928

Helstrup, T. (1987). One, two, or three memories: A problem-solving approach to memory for performed acts. *Acta Psychologica, 66*(1), 37-68. doi:10.1016/0001-6918(87)90017-5

Hirata, Y., Kelly, S. D., Huang, J., & Manansala, M. (2014). Effects of hand gestures on auditory learning of second-language vowel length contrast. *Journal of Speech, Language and Hearing Research, 57*(6), 2090-2101. doi: 10.1044/2014_JSLHR-S-14-0049

Hogrefe, K., Ziegler, W., Wiesmayer, S., Weidinger, N., & Goldenberg, G. (2013). The actual and potential use of gestures for communication in aphasia. *Aphasiology, 27*(9), 1070-1089. doi: 10.1080/02687038.2013.803515

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology*, *30*(1), 47-61. doi: 10.1111/j.1469-8986.1993.tb03204.x

Holle, H., & Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *Journal of Cognitive Neuroscience, 19*(7), 1175-1192. doi: 10.1162/jocn.2007.19.7.1175

Holler, J., Shovelton, H., & Beattie, G. (2009). Do iconic gestures really contribute to the communication of semantic information in a face-to-face context? *Journal of Nonverbal Behavior, 33*(2), 73-88. doi: 10.1007/s10919-008-0063-9

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition, 106*(1), 501-511. doi: 10.1016/j.cognition.2007.02.001

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin, 137*(2), 297-315. doi: 10.1037/a0022128

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment. Gestures as simulated action. *Psychonomic Bulletin & Review*, *15*(3), 495-514.doi: 10.3758/PBR.15.3.495

Hudson, P. T., & Bergman, M. W. (1985). Lexical knowledge in word recognition: Word length and word frequency in naming and lexical decision tasks. *Journal of Memory and Language*, *24*(1), 46-58. doi: 10.1016/0749-596x(85)90015-4

Internet World Stats (2017). World Languages. Retrieved from https://www.internetworldstats.com/languages

Jackson, G. M., Swainson, R., Cunnington, R., & Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition, 4*(2), 169-178. doi: 10.1017/S1366728901000268

Jalbert, A., Neath, I., Bireta, T., & Surprenant, A. M. (2011).When does length cause the word length effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(2), 338-353. doi: 10.1037/a0021804

James, K. H., & Swain, S. N. (2011). Only self-generated actions create sensori-motor systems in the developing brain. *Developmental Science*, *14*(4), 673-678. doi: 10.1111/j.1467-7687.2010.01011.x

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic Influence in Language and Cognition*. New York: Routledge.

Jasper, H. H. (1958). The 10-20 electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371-375. doi: 10.1016/0013-4694(58)90053-1

Jost, L. B., Radman, N., Buetler, K. A., & Annoni, J. M. (2018). Behavioral and electrophysiological signatures of word translation processes. *Neuropsychologia*, *109*, 245-254. doi: 10.1016/j.neuropsychologia.2017.12.034

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*(2), 228-238. doi: 10.1037/0096-3445.111.2.228

Kaushanskaya, M., & Rechtzigel, K. (2012). Concreteness effects in bilingual and monolingual word learning. *Psychonomic Bulletin & Review, 19*(5), 935-941. doi: 10.3758/s13423-012-0271-5

Kelly, S. D., Creigh, P., & Bartolotti, J. (2010). Integrating speech and iconic gestures in a Stroop-like task: Evidence for automatic processing. *Journal of Cognitive Neuroscience, 22*(4), 683-694. doi:10.1162/jocn.2009.21254

Kelly, S. D., Healey, M., Özyürek, A., & Holler, J. (2015). The processing of speech, gesture and action during language comprehension. *Psychonomic Bulletin & Review, 22*(2), 517-523. doi:10.3758/s13423-014-0681-7

Kelly, S. D., Hirata, Y., Manansala, M., & Huang, J. (2014). Exploring the role of hand gestures in learning novel phoneme contrasts and vocabulary in a second language. *Frontiers in Psychology, 5*, 1-11. doi: 10.3389/fpsyg.2014.00673

Kelly, S. D., Manning, S. M., & Rodak, S. (2008). Gesture gives a hand to language and learning: Perspectives from cognitive neuroscience, developmental psychology and education. *Language and Linguistics Compass, 2*(4), 569-588. doi: 10.1111/j.1749-818X.2008.00067.x

Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes, 24*(2), 313-334. doi: 10.1080/01690960802365567

Kelly, S. D., Özyürek, A., & Maris, E. (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological Science, 21*(2), 260-267. doi: 10.1177/0956797609357327.

Kendon, A. (1981). The study of gesture: Some remarks on its history. In J. N. Deely et al. (Eds.), *Semiotics* (pp. 153-166). New York: Plenum Press.

Kerkhofs, R., Dijkstra, T., Chwilla, D. J., & De Bruijn, E. R. (2006). Testing a model for bilingual semantic priming with interlingual homographs: RT and N400 effects. *Brain Research, 1068*(1), 170-183. doi: 10.1016/j.brainres.2005.10.087

Khachatryan, E., Van Vliet, M., De Deyne, S., Storms, G., Manvelyan, H., & Van Hulle, M. M. (2014, May). Amplitude of N400 component unaffected by lexical priming for moderately constraining sentences. In *Cognitive Information Processing (CIP),* 2014 4th International Workshop on (pp. 1-6). IEEE.

Khateb, A., Abutalebi, J., Michel, C. M., Pegna, A. J., Lee-Jahnke, H., & Annoni, J. M. (2007). Language selection in bilinguals: A spatio-temporal analysis of electric brain activity. *International Journal of Psychophysiology*, *65*(3), 201-213. doi: 10.1016/j.ijpsycho.2007.04.008

Kieffaber, P. D., Kruschke, J. K., Cho, R. Y., Walker, P. M., & Hetrick, W. P. (2013). Dissociating stimulus-set and response-set in the context of task-set switching. *Journal of Experimental Psychology: Human Perception and Performance, 39*(3), 700-719. doi: 10.1037/a0029545

Kircher, T., Strube, B., Leube, D., Weibs, S., Sachs, O., Willmes, K. … Greena, A. (2009). Neural interaction of speech and gesture: Differential activations of metaphoric co-verbal gestures. *Neuropsychologia, 47*(1), 169-179. doi:10.1016/j.neuropsychologia.2008.08.009

Kita, S., Alibali, M., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review, 124*(3), 245-266. doi: 10.1037/rev0000059

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48*(1), 16-32. doi: 10.1016/S0749-596X(02)00505-3

Knopf, M. (1992). *Gedächtnis für Handlungen: Funktionsweise und Entwicklung* (Unpublished doctoral dissertation). Ruprecht-Karls-Universität, Heidelberg.

Knudsen, E. I. (2007). Fundamental components of attention. *Annual Review of Neuroscience*, *30*(57-58), 57-78. doi: 10.1146/annurev.neuro.30.051606. 094256

Kolk, H., & Chwilla, D. (2007). Late positivities in unusual situations. *Brain and Language*, *100*(3), 257-261.doi: 10.1016/j.bandl.2006.07.006

Koriat, A., & Pearlman-Avnion, S. (2003). Memory organization of action events and its relationship to memory performance. *Journal of Experimental Psychology: General*, *132*(3), 435-454. doi: 10.1037/0096-3445.132.3.435

Kormi-Nouri, R., & Nilsson, L. G. (2001). The motor component is not crucial! In H. D. Zimmer, R. L. Cohen, M. J. Guynn, J. Engelkamp, R. Kormi-Nouri, & M. A. Foley (Eds.), *Memory for action: A distinct form of episodic memory?* (pp. 97-111). New York: Oxford University Press.

Krashen, S. D., & Terrell, T. D. (1983). *The natural approach: Language acquisition in the classroom.* San Francisco: The Alemany Press.

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261-283). Cambridge: Cambridge University Press.

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica, 128*(3), 416-430. doi: 10.1016/j.actpsy.2008.02.001

Kroll, J. F., & De Groot, A. M. (2005). *Handbook of bilingualism: Psycholinguistic approaches.* New York: Oxford Press.

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind, and brain. *Annual Review of Linguistics, 1*, 377-394. doi: 10.1146/annurev-linguist-030514-124937

Kroll, J. F., & Linck, J. A. (2007). Representation and skill in second language learners and proficient bilinguals. In I. Kecskes & L. Albertazzi (Eds.), *Cognitive Aspects of Bilingualism* (pp. 237-269). Dordrecht: Springer.

Kroll, J. F., Michael, E., & Sankaranarayanan, A. (1998). A model of bilingual representation and its implications for second language acquisition. In A. F. Healy (Eds.), *Foreign Language Learning: Psycholinguistic Studies on Training and Retention* (pp. 365-395). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

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

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

Kroll, J. F., & Tokowicz, N. (2001). The Development of Conceptual Representation for Words in a Second Language. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 49-71). Malden, MA: Blackwell Publishers.

Kroll, J. F., & Tokowicz, N. (2005). Models of bilingual representation and processing: Looking back and to the future. In J. F. Kroll, & De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches*. New York: Oxford University Press.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism*: *Language and Cognition, 13*(3), 373-381. doi: 10.1017/S136672891000009X

Krönke, K., Mueller, K., Friederici, A. D., & Obrig, H. (2013). Learning by doing? The effect of gestures on implicit retrieval of newly acquired words. *Cortex, 49*(9), 2553-2568. doi: 10.1016/j.cortex.2012.11.016

Kuperberg, G. R., Delaney-Busch, N., Fanucci, K., & Blackford, T. (2018). Priming production: Neural evidence for enhanced automatic semantic activity preceding language production in schizophrenia. *NeuroImage: Clinical*, *18*, 74-85. doi: 10.1016/j.nicl.2017.12.026

Kushch, O., Igualada, A., & Prieto, P. (2018). Prominence in speech and gesture favour second language novel word learning. *Language, Cognition and Neuroscience, 33*(8), 1-13. doi: 10.1080/23273798.2018.1435894

Kutas, M., & Federmeier, K. D. (2014). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*, 621-647. doi: 10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 287-301. doi: 10.1126/science.7350657

Lalor, E., & Kirsner, K. (2001). The representation of “false cognates” in the bilingual lexicon. *Psychonomic Bulletin & Review*, *8*(3), 552-559. doi: 10.3758/BF03196191

Lambert, W. E., & Tucker, G. R. (1972). *Bilingual education of children: The St. Lambert experiment.*Mass: Newbury House Publishers.

Lauro, J., & Schwartz, A. I. (2017). Bilingual non-selective lexical access in sentence contexts: A meta-analytic review. *Journal of Memory and Language, 92*, 217-233. doi: 10.1016/j.jml.2016.06.010

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory and Cognition*, *32*(4), 533-550. doi: 10.3758/BF03195845

Lennon, P. (1996). Getting 'easy' verbs wrong at the advanced level. *IRAL-International Review of Applied Linguistics in Language Teaching, 34*(1), 23-36. doi: 10.1515/iral.1996.34.1.23

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning Memory and Cognition*, *35*(2), 381-390. doi: 10.1037/a0014875

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological science*, *20*(12), 1507-1515. doi: 10.1111/j.1467-9280.2009.02480.x

López-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. doi: 10.3389/fnhum.2014.00213

Lotto, L., & De Groot, A. M. (1998). Effects of learning method and word type on acquiring vocabulary in a new language. *Language Learning, 48*(1)*,* 31-69. doi: 10.1111/1467-9922.00032

Macedonia, M. (2014). Bringing back the body into the mind: gestures enhance word learning in foreign language. *Frontiers in Psychology, 5*, 1467. doi: 10.3389/fpsyg.2014.01467

Macedonia, M., Bergmann, K., & Roithmayr, F. (2014). Imitation of a pedagogical agent's gestures enhances memory for words in second language. *Science Journal of Education, 2*(5), 162-169. doi: 10.11648/j.sjedu.20140205.15

Macedonia, M., & Klimesch, W. (2014). Long-term effects of gestures on memory for foreign language words trained in the classroom. *Mind, Brain and Education, 8*(2), 74-88. doi: 10.1111/mbe.12047

Macedonia, M., & Knösche, T. (2011). Body in mind: How gestures empower foreign language learning. *Mind, Brain and Education, 5*(4), 196-211. doi: 10.1111/j.1751-28X.2011.01129.x

Macedonia, M., & Kriegstein, K. (2012). Gestures enhance foreign language learning. *Biolingüistic, 6*(3-4), 393-416.

Macedonia, M., & Mueller, K. (2016). Exploring the neural representation of novel words learned through enactment in a word recognition task. *Frontiers in Psychology, 7*, 953. doi: 10.3389/fpsyg.2016.00953

Macedonia, M., Müller, K., & Friederici, A. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping, 32*(6), 982-998. doi: 10.1002/hbm.21084

Maclellan, E. (2015). Updating understandings of 'teaching': taking account of learners' and teachers' beliefs. *Teaching in Higher Education, 20*(2), 171-182. doi: 10.1080/13562517.2014.966238

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163-203. doi: 10.1037/0033-2909.109.2.163

Macnamara, J. (1967). The bilingual's linguistic performance-a psychological overview. *Journal of Social Issues*, *23*(2), 58-77. doi: 10.1111/j.1540-4560.1967.tb00576.x

Marstaller, L., & Burianová, H. (2014). The multisensory of co-speech gestures- A review and meta-analysis of neuroimaging studies. *Journal of Neurolinguistics, 30*, 69-77. doi: 10.1016/j.jneuroling.2014.04.003

Martin, C. D., Strijkers, K., Santesteban, M., Escera, C., Hartsuiker, R. J., & Costa, A. (2013). The impact of early bilingualism on controlling a language learned late: an ERP study. *Frontiers in Psychology*, *4*, 815. doi: 10.3389/fpsyg.2013.00815

Masumoto, K., Yamaguchi, M., Sutani, K., Tsuneto, S., Fujita, A., & Tonoike, M. (2006). Reactivation of physical motor information in the memory of action events. *Brain Research, 1101*(1), 102-109. doi: 10.1016/j.brainres.2006.05.033

McCafferty, S., & Stam, G. (2008). *Gesture second language acquisition and classroom research.* New York: Taylor & Francis.

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, *7*(7), 703-704. doi: 10.1038/nn1264

McNeill, D. (1992). *Hand and brain.* Chicago: The University of Chicago Press.

McNeill, D. (2005). *Gesture and thought.* Chicago: The University of Chicago Press.

McNeill, D., Levy, E., & Duncan, S. (2015). Gesture in discourse. In D. Tannen, H. Hamilton & D. Schiffrin (Eds.), *Handbook of discourse analysis* (pp. 262-319). UK: John Wiley & Sons.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*(1), 25-40. doi: 10.1006/jmla.1998.2602

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009). Language effects in second language learners and proficient bilinguals investigated with event-related potentials. *Journal of Neurolinguistics*, *22*(3), 281-300. doi: 10.1016/j.jneuroling.2008.08.001

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, *118*(1), 43-71. doi: 10.1037//0096-3445.118.1.43

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology, 114*(2), 187-202. doi: 10.1016/j.jecp.2012.09.002

Moreno, E. M., Federmeier, K. D., & Kutas, M. (2002). Switching languages, switching palabras (words): An electrophysiological study of code switching. *Brain and Language*, *80*(2), 188-207. doi: 10.1006/brln.2001.2588

Moreno-Martínez, F. J., Montoro, P. R., & Rodríguez-Rojo, I. C. (2014). Spanish norms for age of acquisition, concept familiarity, lexical frequency, manipulability, typicality, and other variables for 820 words from 14 living/nonliving concepts. *Annals of Telecommunications*, *46*(4), 1088-1097. doi: 10.3758/s13428-013-0435-x

Morett, L. M. (2014). When hands speak louder than words: The role of gesture in the communication, encoding, and recall of words in a novel second language. *The Modern Language Journal*, *98*(3), 834-853. doi: 10.15845/bells.v5i0.688

Muzzio, I. A., Levita, L., Kulkarni, J., Monaco, J., Kentros, C., Stead, M., ... & Kandel, E. R. (2009). Attention enhances the retrieval and stability of visuospatial and olfactory representations in the dorsal hippocampus. *PLoS Biology, 7*(6), e1000140. doi: 10.1371/journal.pbio.1000140

Myles, F., Hooper, J., & Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning*, *48*(3), 323-364. doi: 10.1111/0023-8333.00045

Nation, P., & Waring, R. (1997).Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge: Cambridge University Press.

Neill, W. T. (1997). Episodic retrieval in negative priming and repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(6), 1291-1305. doi: 10.1037/0278-7393.23.6.1291

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: Effects of proficiency. *Journal of Cognitive Neuroscience*, *17*(8), 1212-1228. doi: 10.1162/0898929055002436

Özçaliskan, S., Gentner, D., & Goldin-Meadow, S. (2014). Do iconic gestures pave the way for children's early verbs? *Applied Psycholinguistic, 35*(6), 1143-1162. doi:10.1017/S0142716412000720

Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: Insights from brain and behaviour. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 369*(1651), 20130296. doi:10.1098/rstb.2013.0296

Padilla, F., Bajo, M. T., & Macizo, P. (2005). Articulatory suppression in language interpretation: Working memory capacity, dual tasking and word knowledge. *Bilingualism: Language and Cognition*, *8*(3), 207-219. doi: 10.1017/S1366728905002269

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.

Palmer, S. D., van Hooff, J. C., & Havelka, J. (2010). Language representation and processing in fluent bilinguals: Electrophysiological evidence for asymmetric mapping in bilingual memory. *Neuropsychologia*, *48*(5), 1426-1437. doi: 10.1016/j.neuropsychologia.2010.01.010

Pérez, M.-Á., & Navalón, C. (2003). Normas españolas de 290 nuevos dibujos: Acuerdo en la denominación, concordancia de la imagen, familiaridad, complejidad visual y variabilidad de la imagen. *Psicológica*, *24*, 215-241.

Phillips, N. A., Klein, D., Mercier, J., & de Boysson, C. (2006). ERP measures of auditory word repetition and translation priming in bilinguals. *Brain Research*, *1125*(1), 116-131. doi: 10.1016/j.brainres.2006.10.002

Pivneva, I., Mercier, J., & Titone, D. (2014). Executive control modulates cross-language lexical activation during L2 reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning Memory and Cognition*, *40*(3), 78-796. doi: 10.1037/a0035583

Poarch, G., Van Hell, J., & Kroll, J. (2014). Accessing word meaning in beginning second language learners: Lexical or conceptual mediation? *Bilingualism: Language and Cogntion, 18*(3), 357-371. doi: 10.1017/S1366728914000558

Potter, M. C., & Faulconer, B. A. (1975). Time to understand pictures and words. *Nature, 253*, 437-438. doi: 10.1038/253437a0

Potter, M. C., Nieuwenstein, M., & Strohminger, N. (2008). Whole report versus partial report in RSVP sentences. *Journal of Memory and Language*, *58*(4), 907-915. doi:10.1016/j.jml.2007.12.002

Potter, M., So, K.F., Von Eckardt, B., & Feldman, B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23*(1), 23-38. doi: 10.1016/S0022-5371(84)90489-4

Prior, A., MacWhinney, B., & Kroll, J. F. (2007). Translation norms for English and Spanish: The role of lexical variables, word class, and L2 proficiency in negotiating translation ambiguity. *Behavior Research Methods*, *39*(4), 1029-1038. doi: 10.3758/bf03193001

Proverbio, A. M., Leoni, G., & Zani, A. (2004). Language switching mechanisms in simultaneous interpreters: An ERP study. *Neuropsychologia*, *42*(12), 1636-1656. doi: 10.1016/j.neuropsychologia.2004.04.013

Quinn-Allen, L. (1995). The effect of emblematic gestures on the development and access of mental representations of french expressions. *The Modern Language Journal, 7*(4), 521-529. doi: 10.1111/j.1540-4781.1995.tb05454.x

Raugh, M. R., & Atkinson, R. C. (1975). A mnemonic method for learning a second-language vocabulary. *Journal of Educational Psychology*, *67*(1), 1-16.doi: 10.1037/h0078665

Rickard, T. C. (2007). Forgetting and learning potentiation: Dual consequences of between-session delays in cognitive skill learning. *Journal of Experimental Psychology: Learning Memory and Cognition*, *33*(2), 297-304. doi: 10.1037/0278-7393.33.2.297

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Revision of Neuroscience*, *27*, 169-192. doi: 10.1146/annurev.neuro.27.070203.144230

Rodríguez-Fornells, A., Cunillera, T., Mestres-Missé, A., & de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3711-3735. doi: 10.1098/rstb.2009.0130

Rugg, M. D. (1990). Event-related brain potentials dissociate repetition effects of high-and low-frequency words. *Memory & Cognition*, *18*(4), 367-379. doi: 10.3758/BF03197126

Schneider, W. (1995). *Micro Experimental Laboratory.* Pittsburgh: Psychology Software Tools Inc.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user’s guide* (Version 1.1). Pittsburg: Psychology Software Tools.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, *55*(2), 197-212. doi: 10.1016/j.jml.2006.03.004

Schwieter, J., & Sunderman, G. (2009). Concept selection and developmental effects in bilingual speech production. *Language Learning, 59*(4), 897-927. doi: 10.1111/j.1467-9922.2009.00529.x

Schwid, S. R., Tyler, C. M., Scheid, E. A., Weinstein, A., Goodman, A. D., & McDermott, M. P. (2003). Cognitive fatigue during a test requiring sustained attention: a pilot study. *Multiple Sclerosis Journal, 9*(5), 503-508. doi: 10.1191/1352458503ms946oa

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371-388). New York: Oxford University Press.

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer between picture naming and translation: A test of asymmetries in bilingual memory. *Psychological Science, 6*(1), 45-49. doi: 10.1111/j.1467-9280.1995.tb00303.x

Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language, 51*(3), 381-398. doi: 10.1016/j.jml.2004.05.003

Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science, 16*(2), 85-89. doi:10.3758/s13423-016-1145-z

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174-215. doi: 10.1037/0278-7393.6.2.174

So, W. C., Chen-Hui, C., & Low Wei-Shan, J. (2012). Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall? *Language and Cognitive Processes, 27*(5), 665-681. doi: 10.1080/01690965.2011.573220

Statista Portal (2017). Retrieved from https://es.statista.com/buscar/?q=language+ learning+europe

Staub, A., Grant, M., Astheimer, L., & Cohen, A. (2015). The influence of cloze probability and item constraint on cloze task response time. *Journal of Memory and Language*, *82*, 1-17. doi: 10.1016/j.jml.2015.02.004

Stefan, K., Cohen, L. G., Duque, J., Mazzocchio, R., Celnik, P., Sawaki, L., ... & Classen, J. (2005). Formation of a motor memory by action observation. *Journal of Neuroscience*, *25*(41), 9339-9346. doi: 10.1523/JNEUROSCI.2282-05.2005

Straube, B., Green, A., Weis, S., & Kircher, T. (2012). A supramodal neural network for speech and gesture semantics: An FMRI study. *PLoS ONE, 7*(11), e51207. doi:10.1371/journal.pone.0051207

Sunderman, G., & Kroll, J. (2006). First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition, 28*(3), 387-422. doi: 10.10170/S0272263106060177

Swain, M., & Lapkin, S. (1982). E*valuating Bilingual Education: A Canadian Case Study. Multilingual Matters 2.* Clevedon: Multilingual Matters.

Tabachnick, B., & Fidell, L. (2001). *Using multivariate statistics* (4th edition). Boston: Allyn & Bacon.

Takashima, A., Bakker, I., Van Hell, J. G., Janzen, G., & McQueen, J. M. (2014). Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. *NeuroImage, 84*, 265-278. doi: 10.1016/j.neuroimage.2013.08.023

Talamas, A., Kroll, J. F., & Dufour R. (1999). Form related errors in second language learning: A preliminary stage in the acquisition of L2 vocabulary. *Bilingualism: Language and Cognition, 2*(1), 45-58. doi: 10.1017/S1366728999000140

Tamminen, J., & Gaskell, M. G. (2013). Novel word integration in the mental lexicon: Evidence from unmasked and masked semantic priming. *Quarterly Journal of Experimental Psychology*, *66*(5), 1001-1025. doi: 10.1080/17470218.2012.724694

Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language, 24*(3), 535-565. doi: 10.1017/S030500099700319X

Taylor, W. L. (1953). "Cloze Procedure": A New Tool for Measuring Readability. *Journalism Bulletin*, *30*(4), 415-433. doi: 10.1177/107769905303000401

Tellier, M. (2008). The effect of gestures on second language memorization by young children. *Gesture, 8*(2), 219-235. doi: 10.1075/gest.8.2.06tel

Titone, R. (1972). *Early bilingualism.* Bruxelles: Charles Dessart.

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2011). Bilingual lexical access during l1 sentence reading: the effects of L2 knowledge, semantic constraint, and L1-L2 intermixing. *Journal of Experimental Psychology: Learning Memory and Cognition*, *37*(6), 1412-1431. doi: 10.1037/a0024492

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in second language acquisition*, *27*(2), 173-204. doi: 10.1017/S0272263105050102

Tonzar, C., Lotto, L., & Job, R. (2009). L2 vocabulary acquisition in children: Effects of learning method and cognate status. *Language Learning, 59*(3), 623-646. doi: 10.1111/j.1467-9922.2009.00519.x

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition*, *18*(1), 3-24. doi: 10.1017/S1366728914000522

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language*, *64*(1), 88-107. doi: 10.1016/j.jml.2010.08.006

Van Assche, E., Duyck, W., & Hartsuiker, R. J. (2016). Context effects in bilingual sentence processing: task specificity. In *Methods in Bilingual Reading Comprehension Research* (pp. 11-31). New York: Springer.

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science*, *20*(8), 923-927. doi: 10.1111/j.1467-9280.2009.02389.x

Van Hell, J., & Candia-Mahn, A. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning, 47*(3), 507-546. doi: 10.1111/0023-8333.00018

Van Hell, J. G., & De Groot, A. M. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1*(3), 193-211.doi: 10.1017/S1366728998000352

Van Hell, J. G., & De Groot, A. M. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica*, *128*(3), 431-451. doi: 10.1016/j.actpsy.2008.03.010

Van Patten, B. (1990). Attending to form and content in the input. *Studies in Second Language Acquisition, 12*(3), 287-301. doi: 10.1017/S0272263100009177

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*(2), 394-417. doi: 10.1037/02787393.25.2.394

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, *18*(4), 380-393. doi: 10.3758/BF03197127

Verhoef, K., Roelofs, A., & Chwilla, D. J. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition*, *110*(1), 84-99. doi: 10.1016/j.cognition.2008.10.013

Vukovic, N., Feurra, M., Shpektor, A., Myachykov, A., & Shtyrov, Y. (2017). Primary motor cortex functionally contributes to language comprehension: An online rTMS study. *Neuropsychologia, 96*, 222-229. doi: 10.1016/j.neuropsychologia.2017.01.025

Wagner, S. M., Nusbaum, H., & Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is hand waving spatial? *Journal of Memory & Language, 50*(4), 395-407. doi: 10.1016/j.jml.2004.01.002

Wang, A., & Thomas, M. (1995). Effects of keyword on long-term retention: Help or hindrance? *Journal of Educational Psychology, 87*(3), 468-475. doi:10.1037/0022-0663.87.3.468

Webster, N. (1961). *Webster's elementary dictionary*. Springfield: G. & C. Merriam Co.

Wesely, P. M. (2009). The language learning motivation of early adolescent french and spanish elementary immersion program graduates. *Foreign Language Annals*, *42*(2), 270-286. doi: 10.1111/j.1944-9720.2009.01021.x

Willems, R. M., Özyürek, A., & Hagoort, P. (2008). Seeing and hearing meaning: ERP and fMRI evidence of word versus picture integration into a sentence context. *Journal of Cognitive Neuroscience*, *20*(7), 1235-1249. doi: 10.1162/jocn.2008.20085

Wimer, C. C., & Lambert, W. E. (1959). The differential effect of word and object stimuli on the learning of paired associates. *Journal of Experimental Psychology*, *57*(1), 31-36. doi: 10.1037/h0046822

Wong, W. (2001). Modality and attention to meaning and form in the input. *Studies in Second Language Acquisition, 23*(3), 345-368. doi: 10.1017/S0272263101003023

Worthen, J. B. (2006). Resolution of discrepant memory strengths: An explanation of the effects of bizarreness on memory. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 133-156). New York: Oxford University Press.

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

Yang, J., Andric, M., & Mathew, M. M. (2015). The neural basis of hand gesture comprehension: A meta-analysis of functional magnetic resonance imaging studies. *Neuroscience and Behavioral Reviews, 57,* 88-104. doi: 10.1016/j.neubiorev.2015.08.006

Yap, D., So, W., Yap, J., Tan, Y., & Teoh, R. (2010). Iconic gestures prime words. *Cognitive Science, 35*(1), 171-183. doi: 10.1111/j.1551-6709.2010.01141.x

Yum, Y. N., Midgley, K. J., Holcomb, P. J., & Grainger, J. (2014). An ERP study on initial second language vocabulary learning. *Psychophysiology*, *51*(4), 364-373. doi: 10.1111/psyp.12183

APPENDIX

Appendix 1. *Words (Nouns and Verbs) and Gestures Used in the Study*

Spanish Nouns	English Translation	Congruent Gestures
Abanico	Fan	Put the right hand as if holding a fan and turn the wrist repeatedly
Aguja	Needle	Hold an imaginary piece of fabric with the left hand and perform the sewing movement with an imaginary needle in the right hand between the thumb and the forefinger
Anillo	Ring	With the left hand, collocate the thumb and the forefinger as if holding a ring and introduce it in the right ring finger
Armónica	Harmonica	Use both hands to hold an imaginary harmonica and move it in front of the mouth while expelling air
Auriculares	Earphones	Take imaginary earphones with both hands and introduce them into the ears
Bayeta	Wiper	Stretch both arms, close the fists and turn the right hand forward and the left one backward as if draining a wiper
Cerillas	Match	Fist the left hand, put together the thumb and forefinger finger tips and pass them by the left thumb
Champú	Shampoo	Separate the fingers of both hands, put them in your head and perform a little circular movement

Cigarrillo	Cigarette	Draw a "V" with the right and middle forefingers, put them next to the mouth, breathe, separate the fingers from the mouth and breathe out
Clínex	Paper tissue	With the nose between the fingers of both hands perform a slight downward movement of the head.
Coche	Car	Put the hands as if holding a steering wheel and turn it right and left
Coleta	Ponytail	Grab the hair with the right hand next to the head and the left one on the hair tips
Cremallera	Zip	Take an imaginary zip and perform the movement of raising the zip on your chest
Cuchara	Spoon	Make the imaginary movement of taking a spoon and approach it to the mouth
Dardo	Dart	With the right hand take an imaginary dart and throw it in front of your face
Flauta	Flute	Hold an imaginary flute with both hands and move your fingers as if you were playing the flute
Fregona	Mop	Make the movement of taking an imaginary broom with both hands and passing it from front to back
Foto	Photo	Take an imaginary camera in front of your face and press an imaginary button with the right forefinger
Lápiz	Pencil	Take an imaginary pen with the right hand between the thumb and the forefinger and perform left-right writing movements
Lentillas	Contact lens	Open your left eye with the left hand and approximate the right forefinger to the left eye

Libro	Book	Stretch the left hand as if holding an imaginary opened book and perform the movement of turning pages on the left hand
Martillo	Hammer	Take the handle of an imaginary hammer and hit the air in front of your face
Mechero	Lighter	Close the right hand and pass the thumb by the forefinger
Micrófono	Microphone	Take an imaginary microphone with the right hand and move your mouth as if you were speaking
Pesa	Weight	Stretch the right arm, close the hand upwards and move the elbow up and down
Peine	Comb	With an imaginary comb in the right hand, comb the hair from the front to the back
Perfume	Perfume	Take an imaginary perfume bottle and press the dispenser in front of your neck
Pistola	Gun	Stretch the thumb and the forefinger and perform a light down movement in an horizontal plane respect to the floor
Platillos	Cymbals	Put the thumb and the forefinger of both hands together to draw a circle, after that, bring them together
Pomo	Knob	Take with the right hand a spherical knob and turn the hand right
Secador	Hairdryer	Take the handle of an imaginary hairdryer and move it around the head
Sello	Stamp	Stretch the left hand as if it was a sheet, close the right hand and hit it on the left hand
Silbato	Whistle	Take an imaginary whistle between the thumb and the forefinger, approximate them to the mouth and blow

Sombrero	Hat	Take the hat visor between the thumb and the forefinger and move them to the head as if you were putting on the hat
Tambor	Drum	Take drumsticks with both hands and hit an imaginary drum in front of you
Tapón	Cap	Put the thumb and the forefinger of both hands together to draw a circle, and move the right hand to the left as if you were covering the left circle
Teclado	Keyboard	Move the fingers of both hands as if you were typing on a keyboard
Tijeras	Scissors	Draw a "V" with the right forefinger and middle finger, put them in a horizontal plane, approach and separate them
Vaso	Glass	Take an imaginary glass with the right hand and put it to the mouth as if you are going to drink
Violín	Violin	Collocate the head as if you were holding an imaginary violin between it and the left shoulder, put the left hand as if you were touching the neck, take an imaginary bow with the right thumb and forefinger and move it by the violin
Sin significado1	Meaningless 1	Close the right fist and move the arm from the left to the right
Sin significado 2	Meaningless 2	Cross the arms in front of the chest
Sin significado 3	Meaningless 3	Stretch the hand's palms, put the right one in front of the right eye and the left one covering the mouth
Sin significado 4	Meaningless 4	Draw an imaginary triangle with the right forefinger in front of the face

Sin significado 5	Meaningless 5	Touch the chin and after that the forehead with the right forefinger
Sin significado 6	Meaningless 6	Stretch the right arm and hand with the palm of the hand facing outwards and move it from left to right
Sin significado 7	Meaningless 7	With the right forefinger, touch the left cheek and after that, the right one
Sin significado 8	Meaningless 8	Touch the forehead and the right ear with the right forefinger
Sin significado 9	Meaningless 9	Stretch the left hand, clap the hands twice, first with right fist closed and after that, with the right hand stretched
Sin significado 10	Meaningless 10	Put the palms of the hands together and rotate them from one side and to another

Note. Left Column: Spanish nouns used in Experiment 1. Middle column: English translations. Right column: approximate description of congruent gestures.

Spanish Verbs	**English translation**	**Gestures**
Afeitar	To shave	Make the movement of grabbing an imaginary razor and passing it by the right cheek
Apretar	To squeeze	Stretch the right hand with the palm up and close the fist strongly
Barrer	To sweep	Make the movement of taking an imaginary broom with both hands and passing it from front to back
Batir	To Beat	Put the left hand as if was an imaginary bowl and with the right one perform the movement of a whisk over the bowl.

Beber	To drink	Stretch the thumb and the little finger, close the rest of the fingers of the right hand and move it towards the mouth
Besar	To kiss	Put the right hand fingers on the lips and whilst separating them, give a kiss
Botar	To bounce	Stretch the right hand in horizontal plane with respect to the floor with the palm of the hand facing the floor and move the hand up and down
Callar	To silence	Put the right forefinger in the lips in a vertical plane
Comer	To eat	Make the imaginary movement of taking a spoon and putting it to the mouth
Coser	To cough	Hold an imaginary piece of fabric with the left hand and perform the cough movement with an imaginary needle in the right hand holding it between the thumb and the forefinger
Disparar	To shoot	Stretch the thumb and the forefinger and perform a slight downward movement in an horizontal plane with respect to the floor
Dormir	To sleep	Put together the palms of the hands, put them to the right and nod the head to the right a little and close the eyes
Escribir	To write	Take an imaginary pen with the right hand between the thumb and the forefinger and perform left-right writing movements in front of your face
Escuchar	To listen	Stretch the right forefinger and touch the right ear

Escurrir	To drain	Fist both hands in front of your chest with 15 centimeters distance between them and spin them in opposite directions
Estirar	To stretch	Put in contact the thumb and the forefinger of both hands, put the hands together and separate the hands
Estornudar	To sneeze	With the nose between the fingers of both hands perform a slight downward movement of the head.
Estrujar	To wring out	Put the palms of the hands one in front of the other separated by approximately 15 centimeters with separation between the fingers, approximate the hands at the same time that you fold the fingers and rotate the hands in opposite directions
Fotografiar	To photograph	Take an imaginary camera in front of your face and press an imaginary button with the right forefinger
Fumar	To smoke	Draw a “V” with the right and middle forefingers and put them next to the mouth, breathe, separate the fingers from the mouth and breathe out
Guiñar	To wink	Wink with the right eye
Llorar	To cry	Close both fists, put the thumb and the forefinger in contact with the cheeks and perform a rotating movement

Maquillar	To apply makeup	Hold an imaginary eye liner with the right hand between the thumb and the forefinger and draw a line following the eye line next to the eye
Mirar	To look	Stretch the right forefinger and put it under the right eye
Morder	To bite	Show the teeth and open and close the jaw
Negar	To deny	Perform a lateral movement with the head
Peinar	To brush	With an imaginary comb in the right hand, comb the hair from the front to the back
Pintar	To draw	Stretch the left hand, hold a brush with the right hand between the thumb and the forefinger and draw with slow movements on the left hand
Rechinar	To grind	Show the teeth and move the lower jaw from right to left
Recortar	To trim	Draw a “V” with the right forefinger and middle finger, put them in an horizontal plane, approach and separate them
Relamer	To lick your lips	Stick out the tongue and pass it by the top lip
Rezar	To pray	Put the palms of the hands together in a vertical plane in front of your chest
Romper	To break	With the palms of the hands facing the floor, close the fists and put them together sideways. Separate the fists sideways

Sacudir	To shake	Close the hands with the palms one in front of the other, and move them towards and outwards from the body
Silbar	To wistle	Position the lips as if kissing and expel air
Sonreir	To smile	Show your teeth in a big smile
Soplar	To blow	Expel air through the mouth
Teclear	To type	Move both hands fingers as if you were typing on a keyboard
Toser	To cough	Put the right fist with the thumb next to the mouth and perform a slight downward movement of the head
Untar	To spread	Stretch both hands, and with the right hand perform the movement of a knife spreading something on the left hand
Sin significado1	Meaningless 1	Close the right fist and move the arm from the left to the right
Sin significado 2	Meaningless 2	Cross the arms in front of the chest
Sin significado 3	Meaningless 3	Stretch the palms of the hands, and put the right one in front of the right eye and the left one covering the mouth.
Sin significado 4	Meaningless 4	Draw an imaginary triangle with the right forefinger in front of the face
Sin significado 5	Meaningless 5	Touch the chin and then the forehead with the right forefinger
Sin significado 6	Meaningless 6	Stretch the right arm and hand with the palm of the hand facing the floor and move it from left to right
Sin significado 7	Meaningless 7	With the right forefinger, touch the left cheek and then the right one
Sin significado 8	Meaningless 8	Touch the forehead and the right ear with the right forefinger

Sin significado 9	Meaningless 9	Stretch the left hand, clap the hands twice, first with right fist closed and then with the right hand stretched
Sin significado 10	Meaningless 10	Put the palms of the hands together and rotate them from one side and to another

Note. Left Column: Spanish verbs used in Experiments 2 and 3. Middle column: English translations. Right column: approximate description of congruent gestures.

Appendix 2. *Stimulus Material Used in the Training Tasks*

Spanish words	Vimmi words	Semantic "yes"	Semantic "no"	Lexical "yes"	Lexical "no"
oso (bear)	miresado	animals	body	s	m
perro (dog)	fukepa	animals	body	p	f
caballo (horse)	bati	animals	fruits	b	t
cabra (goat)	doba	animals	fruits	b	d
burro (donkey)	boreda	animals	kitchen	o	a
ratón (mouse)	fesuti	animals	kitchen	t	f
cebra (zebra)	dirube	animals	music	b	d
cerdo (pork/pig)	wepuda	animals	music	d	p
conejo (rabbit)	ean	animals	vehicles	e	a
gato (cat)	geloro	animals	vehicles	o	e
cabeza (head)	urabe	body	animals	b	r
dedo (finger)	detu	body	animals	d	t
labio (lip)	bikute	body	fruits	b	t
pierna (leg)	ganuma	body	fruits	n	g
nariz (nose)	seza	body	kitchen	a	e
ojo (eye)	kadonega	body	kitchen	o	e
brazo (arm)	zuowe	body	music	o	u
mano (hand)	wari	body	music	a	i
oreja (ear)	iol	body	vehicles	o	i
pie (foot)	furome	body	vehicles	e	o
manzana (apple)	nobani	fruits	animals	n	b
naranja (orange)	wubonige	fruits	animals	n	b
cereza (cherry)	loeke	fruits	body	e	o
fresa (strawberry)	toari	fruits	body	a	o
limón (lemon)	tofita	fruits	kitchen	i	a
piña (pineapple)	deschoga	fruits	kitchen	a	o
pera (pear)	pigemola	fruits	music	p	l
plátano (banana)	raone	fruits	music	n	r
sandía (watermelon)	peabe	fruits	vehicles	a	e
uva (grape)	lodefawi	fruits	vehicles	a	e
salero (salt shaker)	kiale	kitchen	animals	l	k

sartén (frying pan)	mulogite	kitchen	animals	t	m
cuchara (spoon)	gitu	kitchen	body	u	i
cuchillo (knife)	ruzanego	kitchen	body	u	e
taza (cup)	tedo	kitchen	fruits	t	d
vaso (glass)	esepo	kitchen	fruits	s	p
cazo (dipper)	zagido	kitchen	music	a	i
copa (wineglass)	lofuse	kitchen	music	o	e
frigorífico (fridge)	mogra	kitchen	vehicles	o	a
tenedor (fork)	pukoni	kitchen	vehicles	o	i
arpa (harp)	fapro	music	animals	p	f
campana (bell)	lamube	music	animals	m	l
trompeta (trumpet)	mapusebo	music	body	m	s
violín (violin)	koludi	music	body	l	d
flauta (flute)	lefu	music	fruits	f	l
trompa (horn)	pabezi	music	fruits	p	b
acordeón (accordion)	dikemori	music	kitchen	e	i
guitarra (guitar)	foine	music	kitchen	i	e
piano (piano)	ratube	music	vehicles	a	e
tambor (drum)	lasi	music	vehicles	a	i
avión (plane)	lenope	vehicles	animals	n	p
barco (boat)	beropuga	vehicles	animals	b	p
camión (lorry)	pewo	vehicles	body	o	e
carrito (shopping cart)	tanedila	vehicles	body	i	e
bicicleta (bicycle)	paltra	vehicles	fruits	t	p
moto (motorbike)	mofire	vehicles	fruits	m	f
coche (car)	tizo	vehicles	kitchen	o	i
helicóptero (helicopter)	uteli	vehicles	kitchen	i	u
autobús (bus)	brido	vehicles	music	o	i
tren (train)	dra	vehicles	music	r	d

Note. Spanish words (English translation in brackets) and Vimmi words used in the study. Semantic "yes": the Vimmi word denotes an exemplar of the semantic category. Semantic "no": the Vimmi word refers to an exemplar form another semantic category. Lexical "yes": the grapheme is part of the Spanish translation of the Vimmi word. Lexical "no": the grapheme is not part of the Spanish translation of the Vimmi word. Music: Musical instruments, Body: Body parts, Kitchen: Kitchen utensils.

Appendix 3. *Material used in the Sentence Processing Tasks*

	Plausible	Implausible
Highly predictable Sentences		
Mi madre nunca bebe el café en taza, ella prefiere tomárselo en	esepo (vaso)	zuowe (brazo)
(My mother never drinks coffee in a mug, she prefers to have it in)	glass	arm
Como soy poco besucón, prefiero saludar a la gente dándole la	wari (mano)	pigemola (pera)
(As I don't like kisses, I prefer to greet people by shaking)	hand	pear
La parte del cuerpo encargada del sentido del olfato es la	seza (nariz)	doba (cabra)
(The part of the body in charge of the sense of smell is)	nose	goat
Al pasar por Gran Vía vi a un hombre mayor que pedía dinero tocando el	dikemori (acordeón)	tofita (limón)
(Passingby Gran Vía, I saw an old man who was asking for money playing the)	accordion	lemon

Para conservar durante más tiempo la fruta, la guardo en la parte de abajo del	mogra (frigorífico)	miresado (oso)
(To preserve the fruit longer, I store the fruit in the lower part of the)	fridge	bear
Andrés es militar, es el encargado de despertar a todos tocando la	mapusebo (trompeta)	toari (fresa)
(Andrés is military, he is in charge of waking everyoneup in the morning by playing the)	trumpet	strawberry
Al volcar, se escaparon todos los pollos que transportaba el	pewo (camión)	fapro (arpa)
(All the chickens ran away when it overturned the)	lorry	harp
Al darle un beso le dejó la marca de los	bikute (labios)	raone (plátano)
(When she gave him a kiss, she left the mark of her)	lips	banana
Cuando se porta bien, su madre le deja tomar una copa de nata con	toari (fresa)	mapusebo (trompeta)
(When he is a good boy, his mother leaves him has a cup of cream with)	strawberry	trumpet
La fruta típica de Canarias es el	raone (plátano)	bikute (labios)
(The typical fruit in the Canary Island is the)	banana	lips

La famosa fruta que es el símbolo de la marca Apple es la	nobani (manzana)	ganuma (pierna)
(The famous fruit which is the Apple brand symbol is the)	apple	leg
En las torres de las iglesias suele haber	lamube (campanas)	lodefawi (uvas)
(Usually, in the church towers there are)	bell	grapes
La parte del cuerpo que va desde la muñeca hasta el hombro es el	zuowe (brazo)	esepo (vaso)
(The part of the body from the wrist to the shoulder is the)	arm	glass
El chico le dijo al director de la banda que el bombo pesada demasiado para él y le pidió cambiarlo por un	lasi (tambor)	wepuda (cerdo)
(The boy told the director that the bass drum weighed too much for him and he asked to change it for a)	drum	pig
No suelo apurar el corazón de la manzana ni de la	pigemola (pera)	wari (mano)
(I don’t usually rush the heart neither the apple nor the)	pear	hand

No me gusta mucho la Fanta de naranja, prefiero la de	tofita (limón)	dikemori (acordeón)
(I don’t like the orange Fanta too much, I prefer the Fanta of)	lemon	accordion
En la película del oeste, el jinete iba cabalgando en un bonito	bati (caballo)	mofire (moto)
(In the western, the rider was riding a beautiful)	horse	motorbike
Si quieres que se disuelva bien el azúcar en el café, tienes que removerla bien con la	gitu (cuchara)	iol (oreja)
(If you want to dissolve the sugar in the coffee, you have to stir with the)	spoon	ear
Cuando iba por el campo encontré la madriguera de un	ean (conejo)	ratube (piano)
(When I was walking by the countryside, I found the hole of a)	rabbit	piano
José estaba muy nervioso porque poco antes del concierto se rompió una tecla del	ratube (piano)	ean (conejo)
(José was very nervous because just before the concert it was broken a key of the)	piano	rabbit
Ana se rompió la rodilla y tuvieron que escayolarle toda la	ganuma (pierna)	nobani (manzana)
(Ana broke her knee and they had in a cast her entire)	leg	apple

Los pandas y los pardos son diferentes tipos de	miresado (oso)	mogra (frigorífico)
(Panda and Brown are different types of)	bear	fridge
Elena no pudo ponerse sus pendientes nuevos porque tenía una herida en la	iol (oreja)	gitu (cuchara)
(Elena couldn't use her new earrings because she had an injury in her)	ear	spoon
Mis padres no comen melón pero en verano siempre toman de postre un trozo de	peabe (sandía)	dirube (cebra)
(My parents never eat melon, but in summer they always eat as dessert a piece of)	watermelon	zebra
El mejor amigo del hombre suele decirse que es el	fukepa (perro)	beropuga (barco)
(It's usually said that the man best friend is the)	dog	boat
Cuando fui al puerto, vi en el agua un gran	beropuga (barco)	fukepa (perro)
(When I went to the port, I saw in the water a big)	boat	dog
A María no le gusta la leche de vaca ni la de oveja, prefiere la de	doba (cabra)	seza (nariz)
(María doesn't like the cow's or the sheep's milk, she prefers the milk of the)	goat	nose

Para tener un buen jamón ibérico, hay que darle de comer muchas bellotas al	wepuda (cerdo)	lasi (tambor)
(To have a good Iberian ham, we have to give as food a lot of acorns to the)	pig	drum
Eduardo tuvo que coger el coche porque no le funcionaba el tubo de escape de la	mofire (moto)	bati (caballo)
(Eduardo had to take the car because it didn’t work, the exhaust pipe of the)	motorbike	horse
No me agrada que tengas en casa como mascota un	fesuti (ratón)	pukoni (tenedor)
(I don’t like you have at home as a pet a)	mouse	fork
Me regalaron la figurita de un ángel tocando el	fapro (arpa)	pewo (camion)
(They gives me as a gift, the figurine of an angel playing the)	harp	lorry
El día de noche vieja mi prima casi se atraganta comiéndose las	lodefawi (uvas)	lamube (campanas)
(On New Year's Eve, my cousin almost choked on her)	grapes	bell
Los ingleses suelen beberse el té en una	tedo (taza)	loeke (cerezas)
(British people usually drink tea in a)	cup	cherries

De tanto andar, a Julia le ha salido un juanete en el	furome (pie)	geloro(gato)
(I've been walking so long and for that, I've got a bunion on my)	foot	cat
Todas las noches escucho maullar al	geoloro (gato)	furome (pie)
(Every night I hear the meow of a)	cat	foot
En el concierto de flamenco, me gustó mucho el hombre que tocaba la	foine (guitarra)	wubonige (naranja)
(In the flamenco concert, I really liked the man who was playing the)	guitar	orange
Para ir de un sitio a otro, los padres llevan a sus bebés en un	tanedila (carrito)	brido (autobús)
(To go from one place to another, parents carry their babies in a)	stroller	bus
Juan tuvo conjuntivitis porque le entró mucho polvo en el	kadonega (ojo)	kiale (salero)
(Juan had conjunctivitis because dust got in his)	eye	salt shaker
Para pinchar el tomate de la ensalada suelo usar el	pukoni (tenedor)	fesuti (ratón)
(To prick the tomato in the salad I usually use the)	fork	mouse

Mi asiento estaba en el primer vagón del (My seat was in the first wagon of the)	dra (tren) train	koludi (violín) violin
La DGT controla las carreteras desde el aire con un (The DGT controls the roads from the air with a)	uteli (helicóptero) helicopter	zagido (cazo) saucepan
Cuando fui al aeropuerto de Madrid, vi despegar por primera vez un (When I went to the Madrid airport, I saw for the first time how it takes of a)	lenope (avión) plane	detu (dedo) finger
Para aliñar la ensalada, el camarero me trajo una aceitera, la vinagrera y el (To season the salad, the waiter brought me the oilcan, the cruet and the)	kiale (salero) saltshaker	kadonega (ojo) eye
María tuvo que pararse en medio de la autovía porque se le pinchó una rueda al (María had to stop in the middle of the highway because it had punctured the wheel of her)	tizo (coche) car	lofuse (copa) cup
De tanto freír huevos se ha roto el mango de la (From all the egg frying, it has broken the handle of the)	mulogite (sartén) pan	urabe (cabeza) head

Se lo apoyó entre el hombro y la barbilla y con el arco, comenzó a tocar el	koludi (violín)	dra (tren)
(He put it between his shoulder and chin and with the bow, he began to play the)	violin	train
Como no quedaba ninguna, tuve que comprar una lata de rodajas de	deschoga (piña)	lefu (flauta)
(Since there were none left, I had to buy a can of slices of)	pineapple	flute
El animal parecido al caballo con la piel a rayas blancas y negras es la	dirube (cebra)	peabe (sandía)
(The animal with black and white striped skin similar to the horse is the)	zebra	watermelon
Cuando no había microondas, mi abuela siempre calentaba la leche en un	zagido (cazo)	uteli (helicóptero)
(When there weren't microwaves, my grandmother always heated the milk in a)	saucepan	helicopter
En la mano tenemos cinco	detu (dedos)	lenope (avión)
(In the hand, we have five)	fingers	plane
A Pedro le costó mucho cortar el filete porque no estaba bien afilado el	ruzanego (cuchillo)	boreda (burro)
(Pedro had a hard time cutting the steak because it was not sharp enough the)	knife	donkey

El flautista de Hamelín dirigía a los ratones tocando su	lefu (flauta)	deschoga (piña)
(The Pied Piper guided mousses by playing his)	flute	pineapple
Es muy típica la imagen de un mexicano montado en un	boreda (burro)	ruzanego (cuchillo)
(It is typical the image of a Mexican riding a)	donkey	knife
Aunque a Jose le gustan las guindas rojas, no soporta las	loeke (cerezas)	tedo (taza)
(Although Jose likes the red sour cherries, he cannot stand the)	cherries	cup
Cuando el ciclista iba por la carretera se rompió el pedal al caerse de la	paltra (bicicleta)	pabezi (trompa)
(When the cyclist was on the road, it broke the pedal of the)	bicycle	horn
El camarero me sirvió el vino en una	lofuse (copa)	tizo (coche)
(The waiter poured the wine into a)	cup	car
Llevo diez minutos esperando en la parada del	brido (autobús)	tanedila (carrito)
(I have been waiting for ten minutes in the stop of the)	bus	stroller

El médico me dijo que necesitaba vitamina C, que tomase mucho zumo de	wubonige (naranja)	foine (guitarra)
(The doctor told me that I needed vitamin C, I have to drink a lot of juice of)	orange	guitar
El instrumento que tiene el mismo nombre que la nariz de los elefantes es la	pabezi (trompa)	paltra (bicicleta)
(The instrument which has the same name than the elephant nose is the)	horn	bicycle
Dicen que el animal al que más le gusta el queso es el	fesuti (ratón)	pukoni (tenedor)
(It's said that the animal that likes cheese the most is the)	mouse	fork
De tanto pensar sobre el tema, a Juan le estuvo doliendo toda la tarde la	urabe (cabeza)	mulogite (sartén)
(From so much thinking about it, Juan had the whole afternoon a)	head	pan
Sentences of low predictability		
Isabel sale de trabajar a las 5, no puede llegar a tiempo para coger el	lenope (avión)	detu (dedos)
Isabel finish working at 5, she can't get on time to take the	plane	fingers
La niña empezó a llorar porque a su muñeca se le había perdido un	kadonega (ojo)	kiale (salero)
(The little girl started to cry because she lost her doll's)	eye	salt shaker

Cuando iba corriendo por la calle porque llegaba tarde a clase, se la cayó el	koludi (violín)	dra (tren)
(When he was running down the street because he was late for class, he dropped his)	violin	train
Aunque fui corriendo, cuando llegué ya había cerrado el	brido (autobús)	tanedilla (carrito)
(Although I ran, when I got, it was already closed the)	bus	(stroller)
Como había estudiado, el niño suspendió su examen de	lefu (flauta)	deschoga (piña)
(As the boy didn't study, he failed his exam)	flute	pineapple
A esa chica le hicieron el mismo día que a mí un tatuaje en el	furome (pie)	geloro (gato)
(That girl had her tattoo done the same day as me in the)	foot	cat
Mi amiga tiene en casa un	geloro (gato)	furome (pie)
(My friend has at home a)	cat	foot
Paula tiene 5 años y sus padres han pensado regalarle para su cumpleaños una	paltra (bicicleta)	pabezi (trompa)
(Paula is 5 years old and her parents have thought about giving her a)	bicycle	horn

Después de tanto esperar para hacer la foto, se cruzó un turista y solo salió la (After be waiting for a long time to take the picture , a tourist stood in the middle and I only caught the)	urabe (cabeza) head	mulogite (sartén) pan
Aunque no la conocía de nada, mientras comía le pedí a la chica de al lado el (Although I didn’t know her, while I was eating, I asked her for the)	kiale (salero) salt shaker	kadonega (ojo) eye
La abuela tenía varias, me pero no encontraba ninguna (The grandmother has a lot, but she cannot find any)	mulogite (sartén) pan	urabe (cabeza) head
Los animales tienen algunos (Animals have some)	detu (dedo) finger	lenope (avión) plane
Cuando nació su hijo, a mi amiga le regalaron de todo menos un (When her baby was born, my friend had a lot of gifts but not a)	tanedilla (carrito) stroller	brido (autobús) bus
Cuando vi cómo estaba de usado, el pedí otro (When I saw how used it was, I asked for another)	pukoni (tenedor) fork	fesuti (ratón) mouse
Alicia tuvo la oportunidad de comprobar cómo era el trabajo en la recolección de (Alicia had the opportunity to check how the work was in the harvest of)	wubonige (naranja) orange	foine (guitarra) guitar

En la granja podemos encontrar diferentes animales como un	boreda (burro)	ruzanego (cuchillo)
(In farms, we can find different animals as a)	conkey	knife
Parece que la chica que estaba en la banda no tocaba demasiado la	foine (guitarra)	wubonige (naranja)
(It seems that the girl who was in the band didn't play too much the)	guitar	orange
De los muchos animales que había en el zoo, al niño le gustó más la	dirube (cebra)	peabe (sandía)
(There were a lot of animals in the zoo, but the children liked better the)	zebra	watermelon
Paco intenta no comerlas porque cree que es alérgico a las	lodefawi (uvas)	lamube (campana)
(Paco try not toeat them because he thinks he is allergic to)	grapes	bell
Como iba corriendo se le cayó una	lofuse (copa)	tizo (coche)
(He was running and it dropped a)	cup	car
Como no sabía dónde lo podía hacer, al final puse la mezcla para la tarta en un	zagido (cazo)	uteli (helicóptero)
(As I didn't know where to do it, finally I put the cake mix in a)	saucepan	helicopter
Es increíble que tenga toda la pared de su dormitorio llena de posters de un	tizo (coche)	lofuse (copa)
(It's incredible that he has the wall of his bedroom fullof posters of a)	car	cup

Dijeron en las noticias que había habido un accidente de	dra (tren)	koludi (violín)
(They said in the news that there were an accident of a)	train	violin
No me gustan nada los bombones que tienen dentro	loeke (cerezas)	tedo (taza)
(I didn’t like the chocolates filled with)	cherries	mug
Desde la ventana de casa, vimos pasar muy de cerca un	uteli (helicóptero)	zagido (cazo)
(Since our house window, we saw passing by very close a)	helicopter	saucepan
Cuando llegaron al río, Pedro miró en su mochila y no encontró el	ruzanego (cuchillo)	boreda (burro)
(When they arrived at the river, Pedro lookedin his backpack and didn’t find the)	knife	donkey
A Pilar le gusta coleccionar	tedo (taza)	loeke (cerezas)
(Pilar likes to collect)	mug	cherries
Antes no tenían, pero la banda ha contratado a dos personas para que toquen la	pabezi (trompa)	paltra (bicicleta)
(Until now it doesn’t have one, but the band has hired two persons to play the)	horn	bicycle
Mi madre me mandó al supermercado a comprar café, leche, mermelada y	deschoga (piña)	lefu (flauta)
(My mother makes me go to the supermarket to buy coffee, milk, jam and)	pineapple	flute

Mi amigo Juan dijo que nunca se había montado en un	pewo (camión)	fapro (arpa)
(My friend Juan said that he has never been mounted in a)	lorry	harp
Cuando Jose llegó de su viaje a Japón me pidió la	wari (mano)	pigemola (pera)
(When Jose arrived from his trip, he ask for my)	hand	pear
Me parece muy desagradable el sonido de las	lamube (campanas)	lodefawi (uvas)
(I find very unpleasant the sound of the)	bells	grapes
Mi medio de transporte favorito es el	beropuga (braco)	fukepa (perro)
(My favorite transport is the)	boat	dog
Cuando Javier llegó a la casa vio a una chica que estaba alimentando a su	bati (caballo)	mofire (moto)
(When Javier got home, he saw a girl feeding his)	horse	motorbike
En la excursión al campo, los niños vieron a un	ean (conejo)	ratube (piano)
(On the field trip, children saw a)	rabbit	piano
Como no sabían que regalarle a su sobrina pensaron que podría gustarle un	dikemori (acordeón)	tofita (limón)
(They didn't know what to give her nephew as a present and they thought that she might like a)	accordion	lemon

Para mejorar el sonido de la banda, el director decidió incorporar una	mapusebo (trompeta)	toari (fresa)
(To improve the band sound, the director decided to incorporate a)	trumpet	strawberry
Cuando voy a un italiano siempre me pido raviolis y	pigemola (pera)	wari (mano)
(When I go to a Italian restaurant I always order raviolis and)	pear	hand
Dicen que bueno criar a un	fukepa (perro)	beropuga (barco)
(It’s good to raise a)	dog	boat
Cuando Fernando iba con sus amigos por la sierra se encontraron una	doba (cabra)	seza (zariz)
(When Fernando was walking through the mountains with his friends, they found a)	goat	nose
Aunque parezca raro, me encanta comer la tortilla de patatas con	tofita (limón)	dikemori (acordeón)
(Although it can sound estrange, I really like eating the Spanish omelet with)	lemon	accordion
Ese anciano de la sala de espera no para de decir que le duele el	zuowe (brazo)	esepo (vaso)
(The old man in the waiting room does not stop saying that it hurts his)	arm	glass
Menos mal que no me hizo daño, solo me dio un golpe en la	iol (oreja)	gitu (cuchara)
(It didn’t hurt me. He just slapped me in the)	ear	spoon

En la frutería compré unos (In the greengrocer I bought some)	raone (plátano) banana	bikute (labios) lips
Esta mañana me levanté muy enfadada, no entendía por qué había a las cinco de la madrugada alguien tocando un (I was really angry this morning, I didn't understand why there were a person at five in the morning playing a)	lasi (tambor) drum	wepuda (cerdo) pig
El profesor le dijo que era mejor que se apuntase el año siguiente a clases de (The teacher told him that he'd better go to classes of)	ratube (piano) piano	ean (conejo) rabbit
Según la historia, durante la guerra, a este personaje le cortaron la (According to the history, this character lost his)	seza (nariz) nose	doba (cabra) goat
Aparte de la bebida y la carne hemos pensado comprar para la barbacoa una (Apart from drinks and meat we though to buy for the barbecue a)	peabe (sandía) watermelon	dirube (cebra) zebra
Cuando iba andando por la calle la niña se encontró una (When the little girl was walking on the street, she found a)	gitu (cuchara) spoon	iol (oreja) ear

Esa chica tiene unos graciosos	bikute (labios)	raone (plátano)
(That girl has funny)	lips	banana
Mi padre me mandí a comprar para el cumpleaños un	esepo (vaso)	zuowe (brazo)
(For the birthday my father asked me to buy a)	glass	arm
Como iba muy rápido, se cayó y se hizo daño en la	ganuma (pierna)	nobani (manzana)
(As he was running too fast, he fell and hurt his)	leg	apple
Una de las carnes que se usa habitualmente para el consume humano es la de	wepuda (cerdo)	lasi (tambor)
(One of the meats that people normally consume is that of the)	pig	drum
A Jesús le dijeron que intentase comer siempre como merienda una	nobani (mazana)	ganuma (pierna)
(They told Jesus that he should try to eat always as a snack a)	apple	leg
Los vecinos dijeron que la noche anterior habían visto al ladrón con las llaves de su	mofire (moto)	bati (caballo)
(Neighbors said that the previous night they saw a thief with the keys of her)	motorbike	horse
Asistieron en Madrid a un concierto de	fapro (arpa)	pewo (camión)
(They assisted in Madrid to a concert of)	harp	lorry

Al final, mi vecino tuvo que llamar al técnico porque él era incapaz de arreglar su	mogra (frigorífico)	miresado (oso)
(Finally, my neighbor had to call the technician because he was unable of repairing his)	fridge	bear
En algunos bosques podemos encontrar a un	miresado (oso)	mogra (frigorífico)
(In some forest we can find a)	bear	fridge
Ayer María fue a la mejor pastelería de la ciudad y compró galletas y una tarta de	toari (fresas)	mapusebo (trompeta)
(Yesterday María went to the best bakery of the city and bought cookies and a pie of)	strawberries	trumpet

Note. Sentences were presented in Spanish and ended with a Vimmi word. The approximate English translation reported in brackets is adapted here to end the sentence with the critical word. The syntactic structure of all sentences was correct in Spanish.