

Review

Not peer-reviewed version

Gestures as Scaffolding to Learn Vocabulary in a Foreign Language

Ana Belén García-Gámez and [Pedro Macizo](#) *

Posted Date: 31 October 2023

doi: 10.20944/preprints202310.1975.v1

Keywords: foreign language learning; language learning strategies; iconic gestures



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

Gestures as Scaffolding to Learn Vocabulary in a Foreign Language

Ana Belén García-Gámez ^{1,2} and Pedro Macizo ^{1,2*}

¹ Mind, Brain and Behavior Research Center (CIMCYC), Granada, Spain

² University of Granada; anabgarcia92@gmail.com (A.G.); pmacizo@ugr.es (P.M.)

* Correspondence: pmacizo@ugr.es; Tel.: +34 686103200

Abstract: This review paper investigates the influence of gestures on foreign language (FL) vocabulary learning through a series of experiments conducted in our laboratory. The manipulation of the gesture-word relationship was a consistent factor across the studies. Firstly, we examined the impact of gestures on noun and verb learning. The results revealed that participants exhibited better learning outcomes when FL words were accompanied by congruent gestures compared to a no gesture condition. This suggests that gestures have a positive effect on FL learning when there is a meaningful connection between the words and the accompanying gestures. However, in general, the recall of words in conditions where gestures were incongruent or lacked meaning was lower than in the no gesture condition. This indicates that under certain circumstances, gestures may have a detrimental impact on FL learning. We analyzed these findings in terms of their implications for facilitating or interfering with FL acquisition. Secondly, we addressed the question of whether individuals need to physically perform the gestures themselves to observe the effects of gestures on vocabulary learning. To explore this, participants were divided into two experimental groups. In one group, participants learned the words by actively performing the gestures ("do" learning group), while the other group simply observed the gestures performed by others ("see" learning group). The processing of congruent gestures facilitated the recall of FL words in both the "see" and "do" learning groups. However, the interference effect associated with processing incongruent gestures was more pronounced in the "see" learning group than in the "do" learning group. Thus, the performance of gestures appears to mitigate the negative impact that gestures may have on the acquisition of FL vocabulary. In conclusion, our findings suggest that iconic gestures can serve as an effective tool for learning vocabulary in a FL, particularly when the gestures align with the meaning of the words. Furthermore, the active performance of gestures helps counteract the negative effects associated with inconsistencies between gestures and word meanings. Consequently, if a choice must be made, a FL learning strategy in which learners acquire words while making gestures congruent with their meaning would be highly desirable.

Keywords: foreign language learning; language learning strategies; iconic gestures

1. Introduction

Experimental science has still a lot of questions to solve in the fields of language learning and multilingualism. Due to the global and multicultural ambiance we are involved nowadays, it is mandatory to be able to communicate in different languages. For this reason, there is a research question which has gained more and more importance in the last decades. What is the best way to learn a foreign language (FL)? Different settings including immersion and studying abroad programs emerge as favorable options for language acquisition [1–7], however these alternatives may not consistently be accessible to those seeking to acquire a new language. In addition, the issue of language learning does not only affect the younger generations. Nowadays, the adult population finds itself confronted by the challenge of knowing languages other than their mother tongues. In this context, research becomes essential in order to provide strategies to guide and facilitate learning

in the usual linguistic contexts of the speakers. The emergence of this increasing necessity creates a direct practical application framework and a way in which investigation can directly impact the language learning experience and facilitates the new FL learners experience.

Early techniques for acquiring FL vocabulary employed a first language (L1)-FL words association strategy aimed at establishing connections between newly acquired FL words and their corresponding lexical translations in the native language [8–11]. To illustrate, a native speaker of Spanish would learn that the English translation for “fresa” (L1) is strawberry (FL). Going a step further in the word association strategy, the keyword method [12,13] involves the utilization of a mnemonic method based on selecting a L1 word phonetically resembling a portion of a FL word (the keyword). In this approach, learners initially associate the spoken FL word with the keyword, followed by connecting the keyword to the L1 translation of the target word in the FL. For instance, the Spanish word “cordero” meaning “lamb”, associated with the word “cord” (phonetically related keyword). Previous research has confirmed the effectiveness of these methods during the early phases of FL learning, attributed to the establishment of lexical associations between L1 and FL [14,15]. Nevertheless, when proficient bilingual individuals seek to express themselves in FL, the most optimal processing route is the direct access to FL words from their associated concepts. The reliance on cross-linguistic lexical connections (L1-FL) and the retrieval of L1 words becomes superfluous when bilinguals communicate in FL [16]. In addition, when learning programs based on the reinforcement of semantic connections are compared with lexically-based learning procedures at the earliest stages of FL acquisition, advantages have been found associated with conceptually mediated strategies [8,9,15–19].

In this context, understanding how fluent bilinguals process FL can serve as a foundation for identifying learning methodologies that can replicate this processing pattern, thus constituting effective learning strategies. Research from earlier studies confirmed the notion that building connections between FL words and concepts thrives through training protocols involving semantic processing [9,17,20–24]. These conceptually mediated strategies usually consist of protocols involving multimedia learning. The term “multimedia” encompasses the incorporation of five distinct types of stimuli within the learning protocol, which can be presented in combination or isolation: text, audio, image, animation and captions/subtitles [25–27]. Today, the principles of the Cognitive Load Theory [28] have largely been surpassed. This theory supports that working memory and cognitive channels have finite capacity and can become overloaded when presented with redundant information [28]. Now, we know that when new information is presented in combination with different multimedia modalities learners have the opportunity to form numerous mental representations of the target knowledge, organize and integrate them into their long-term memory [29–31] ultimately improving learning outcomes. Notably, the presentation of FL words alongside images denoting their meanings (picture association method) outperforms the presentation of FL words with translations in the L1 (word association method) when behavioral measures are collected in single-word tasks [8–10,15,18,32], and tasks with words embedded in the context of sentences [10]. Moreover, electrophysiological measures are also sensitive to benefit of the picture association method even after a single and brief learning session [15]. Likewise, the act of envisioning the meanings of words to be learned in a FL enhances the acquisition process [33,34]. More relevant for the purpose of this manuscript, the integration of words and gestures has been proposed as a powerful FL learning tool, as these elements (words and gestures) interact to construct an integrated memory representation of the new words meaning [35–40]. Theoretically supporting these robust findings regarding the effectiveness of semantically related learning strategies, the Dual Coding Theory [41] suggests that forming mental images while learning may contribute to acquiring new words. According to this theory, the integration of verbal information, visual imagery, and movement heightens the likelihood of remembering new words compared to relying solely on verbal glosses.

In this paper, our objective is to conduct a comprehensive review of the topic, beginning with a general overview of the influence of movements on learning. We will then progressively delve deeper into the examination of how gestures specifically impact the acquisition of vocabulary in a FL. Then,

this review will center on detailing the contributions of studies conducted in our lab to understand the role of gestures in FL vocabulary acquisition.

Movements and Language

As humans, we possess the capacity to execute various types of gestures contingent upon the context or situation we encounter. This variability is influenced by the level of consciousness invested in the action, as well as the intended purpose for which the gestures are enacted. For example, there are large differences between the type of gestures we naturally do while teaching in a classroom setting or when trying to explain something to our colleagues relative to the type of gestures that we do when we are playing with a child and we try to represent a lion or a snake. In 1992, McNeill [38] proposed a simple *gesture taxonomy* encompassing many of the possible movements that we can do in a natural communication context. Firstly, representational gestures encompass iconic gestures, employed to visually illustrate spoken content by employing hand movements to refer to tangible entities and/or actions. Additionally, metaphorical gestures fall under this category, conveying abstract concepts by expressing concrete attributes that can be associated with them. This classification introduces two further gesture types: deictic gestures, wherein one or more fingers point towards a reference, and beat gestures, constituted by hand movements that mirror speech prosody and accentuate its emphasis. It might be mentioned that iconic gestures can be distinguished from emblematic gestures, which are culturally specific and involve bodily motions conveying messages akin to words, such as the "good" sign (thumb up, closed fist).

Across all spoken languages, individuals complement their speech with visual-manual communication [42–44]. This form of multimodal interaction, known as co-speech gestures, encompasses spoken language, facial expressions, body movements, and, notably, hand movements. Together, these visual and auditory components constitute an interconnected stream of information that enhances the process of communication [45]. In fact, the significance of movement in language processing has been substantiated by numerous studies [46,47]. As an interesting example, Glenberg and colleagues [48] observed a correlation between movement performance and language comprehension. In their study, participants were tasked with transferring 600 beans (individually) from a larger container to a narrower one, either toward or away from their bodies based on the container's location. Subsequently, sentences describing movements, both meaningful and meaningless, were presented, and participants judged their plausibility. The results revealed that the time taken to evaluate the sentences depended on whether the direction of the bean's movement matched the direction described in the sentence (toward or away from the body). Thus, the execution of actions influenced language comprehension. Notably, Marstaller and Burianová [49] demonstrated that the right auditory cortex and the left posterior superior temporal brain areas appear to be selectively activated for multisensory integration of sounds from spoken language and accompanying gestures.

Numerous *theoretical frameworks* exist to elucidate the connections between gestures and speech. These frameworks primarily delve into the underlying representations involved in gesture processing. These models can be differentiated by considering the interplay of visuospatial and linguistic information. Some perspectives propose that gestures' representations are rooted in visuospatial images (e.g., the sketch model by de Ruiter [50]; the interface model by Kita and Özyürek [51]; the gestures-as-simulated-action (GSA) framework by Hostetter & Alibali [52]). Conversely, other models emphasize the close interrelation between gesture representations and linguistic information (e.g., the interface model by Kita and Özyürek [51]; the growth point theory by McNeill [38,39]). Another distinction among models lies in how gestures and speech are processed, whether as separate entities or as a unified process. Some models propose that gestures and speech are processed independently (e.g., the lexical gesture process model by Krauss and colleagues [53]), interacting when forming communicative intentions (sketch model by de Ruiter [50]), or during the conceptualization phase (interface model by Kita and Özyürek [51]) to facilitate effective communication. On the contrary, other models posit that gestures and speech function collaboratively within a single system (the growth point theory by McNeill [38,39]; the GSA

framework by Hostetter and Alibali [52]). For instance, the gesture-in-learning-and-development model (Goldin-Meadow [54,55]) suggests that children process gestures and speech autonomously, and these elements integrate into a unified system in proficient speakers [56,57]. Beyond the mechanics of gesture and speech processing, a pertinent question arises regarding the role of gestures in communication. Research has established that listeners extract information from gestures [58–62]. This aligns with the fact that gestures frequently emerge during speech planning, leading many models to emphasize the communicative value of gestures (e.g., the sketch model by de Ruiter [50]; the interface model by Kita and Özyürek [51]; the growth point theory by McNeill [38,39]; and the GLD framework by Goldin-Meadow [54,55]).

In conclusion, it is reasonable to infer that the gestures we utilize while attempting to convey a concept, as well as the natural use of deictic or beat gestures within our spontaneous linguistic expressions, play a role in both language production and comprehension, ultimately enhancing the overall communication process.

Gestures as FL learning tool

Several studies have underscored the significance of different types of gestures in FL learning (e.g., [63–70], for reviews). Generally, it is widely accepted that gestures have a positive impact on vocabulary acquisition, advocating for their integration into FL instruction, aligned with a natural language teaching approach (see [71–74], however, see [35,75], for evidence about the limited effect of gestures on learning segmental phonology). More relevant to the purpose of this paper, past studies have illustrated the role of iconic gestures in language comprehension [76] (for a comprehensive review, see [77]), as well as in language production (see [78], for a review on gestures in speech).

In cognitive psychology, three main perspectives offer explanations for the beneficial role of iconic gestures in FL vocabulary learning.

The *self-involvement explanation* posits that gestures promote participant engagement in the learning task, enhancing attention and favoring FL vocabulary acquisition [79]. This increase in attention is primarily attributed to heightened perceptual and attentional processes during the proprioception of movements associated with gestures or the use of objects to perform actions [80]. Here, the motor component itself might not be cause of improvement [81], rather it is the multisensory information conveyed by the gesture that leads to enhanced semantic processing and greater attention [82]. Consequently, according to this perspective, learning new FL words with gestures aids vocabulary acquisition regardless of whether a gesture is commonly produced within a language or signifies the same meaning as the word being learned. Increased attentional processing contributes to word retention [83], highlighting its role in encoding and information retrieval [84]. To illustrate, if the learner needs to acquire the word “teclear” in Spanish whose translation in English is “to type”, the mere fact of performing a movement associated with the new word would facilitate the process independently of any other intrinsic characteristic of the gesture.

The *motor-trace perspective* asserts that the physical element of gestures is encoded in memory, leaving a motor trace that assists in acquiring new FL words [85,86]. In this view, physical enactment is crucial as it allows the formation of a motor trace tied to the word's meaning. Recent neuroscientific studies, employing techniques like repetitive transcranial magnetic stimulation, support the role of the motor cortex in comprehending written words [87]. Additionally, evidence suggests that familiar gestures might engage procedural memory due to their reliance on well-defined motor programs [88]. Consequently, the interplay of procedural and declarative memory could enhance vocabulary learning. Hence, well-practiced familiar gestures would facilitate the FL learning process to a greater extent than unfamiliar gestures (e.g., the gesture of typing on a keyboard vs. touching the right and the left cheeks with the right forefinger sequentially). Thus, the extent of an individual's familiarity with certain gestures would determine the facilitation effect. However, according to this perspective, the impact of gestures operates independently of their meaning and well-practiced gestures might benefit learning regardless of whether they match the new word's meaning.

The *motor-imagery perspective* suggests that gestures are tied to motor images that contribute to a word's meaning [89]. Specifically, executing a gesture while processing a word fosters the formation of a visual image linked to the word's meaning, enhancing its semantic content [66,90]. Neurobiological evidence, including functional connectivity analyses, suggests the involvement of the hippocampal system in binding visual and linguistic representations of words learned with pictures [91]. According to this view, the facilitation effect of gestures is heightened when they align with the meaning of the words being learned, compared to cases where gestures and word meanings don't match. This constitutes the primary point of disagreement with the motor-trace theory. Additionally, this perspective points out that learning words with gestures of incongruent meanings can lead to semantic interference and reduced recall due to the creation of a dual task scenario which ultimately has an adverse impact on the learning process [92,93].

It's important to note that, from our view, these three perspectives are not mutually exclusive, but rather highlight different aspects of gestures' effect on FL learning. A gesture accompanying a word could increase self-involvement (gestures enhancing attention to FL learning), create a motor trace (meaningful movements), and/or evoke a semantic visual image integrated with the word's meaning.

Empirical evidence regarding the role of gestures in FL learning

Empirical evidence pertaining to the role of movements in FL instruction indicates that enhanced vocabulary learning outcomes are observed when participants learn FL words alongside gestures that embody the practical use of objects whose names are learned [94–96]. The significance of movements in FL vocabulary acquisition has been examined in previous learning protocols. For instance, many years ago, Asher [97] was a pioneer in introducing movements in the FL learning process. He presented the Total Physical Response strategy as an effective means of acquiring new vocabulary. This strategy involved a guided approach where students received instructions in the target language (FL). For example, children were taught the Japanese word "tobe" (meaning 'to jump' in English), and each time they heard this word, they physically performed the corresponding gesture (to jump). The author observed an advantageous impact linked with the integration of gestures in FL word instruction, an effect that has been demonstrated across various educational domains, including online courses, language learning, and technology implementation [98,99] (although see [100] for an alternative view).

Then, the first empirical study focused on the impact of iconic gestures on learning a FL was developed by Quinn-Allen [101]. Within this study, English-speaking participants were tasked with acquiring French expressions under two distinct conditions. In the control condition, the participants were presented with French sentences (e.g., *Il est très collet monté?* - He takes himself seriously), which they then had to repeat in French. Conversely, in the experimental condition, learners were provided with the sentences accompanied by a symbolic gesture illustrating the intended meaning (e.g., Head up, one hand in front of neck, the other hand lower as if adjusting a tie), and they were subsequently required to reproduce the gesture. The pattern of outcomes revealed that sentences presented alongside gestures showed enhanced recall compared to the control condition where participants simply had to repeat the sentences. The efficacy of gestures observed in Quinn-Allen's study aligns with the self-involvement theory [79]. As mentioned above, this theory posits that participants become more engaged in the learning process when they receive and produce gestures compared to the control condition. However, the enhanced learning effect could equally be explained by the motor-trace account and the motor-imagery explanation. The gestures employed in this study, like the action of adjusting a tie, are defined as conventional gestures; the type of gestures which are commonly used in interpersonal communication. This aspect might elucidate the learning enhancement under the motor-trace perspective. Furthermore, the gestures are congruent in meaning with the sentence being acquired ("He takes himself seriously"), supporting the explanation proposed by the motor-imagery account. In this line, this work laid the foundation to study the effect of gestures in FL acquisition, however, the causes underlying this improvement associated with the use of gestures during FL learning remain unclear. Further experimental work has resolved this question

by including several gesture conditions and manipulating the correspondence between the gestures and new words meaning [66,90,95,102].

Another benchmark study in the field is the work by Macedonia and collaborators [90]. In this study, a group of German speakers learned new words in Vimmi (an artificial language developed by the authors) that served as FL. The new nouns to be learned were accompanied by either meaningful congruent iconic gestures (e.g., the term "suitcase" paired with a gesture mimicking an actor lifting an imaginary suitcase) or meaningless gestures (e.g., "suitcase" with a gesture involving touching one's own head). The results revealed a superior recall of words associated with iconic gestures when compared to those coupled with meaningless gestures. These findings suggest that gestures introduce something else beyond merely involving the participant in the task. The simple engagement associated with the gesture's performance cannot explain the advantages found in the iconic gestures condition. However, in this case, both the motor-trace and motor-imagery theories could potentially explain the heightened performance observed in the context of iconic gestures relative to meaningless gestures. Iconic gestures might enhance FL learning due to their semantic richness and their higher frequency of use compared to meaningless gestures, resulting in stronger motor activation.

Other researchers have employed additional experimental conditions to distinguish between explanations grounded in the motor component of gestures and those derived from the motor-imagery account. Studies including monolingual speakers have explored congruity effects in communication by deliberately mismatching the semantics of words and the meanings conveyed by gestures [94,103–106] (see [107] for congruity effects in an unfamiliar language context). Kelly and colleagues [65,108] employed an event-related potential study alongside a Stroop-like paradigm. Participants were presented with words (e.g., "cut") and corresponding gestures, which could be either congruent (e.g., a cutting movement) or incongruent (a drinking movement). The study revealed an attenuated N400 response to words paired with congruent gestures compared to incongruent ones, displaying, thus, a semantic integration effect [109]. Additionally, participants were faster responding in the congruent condition compared to the incongruent one. These results suggest that gestures are integrated within new words meaning producing benefits when gestures and words meaning match and an interference when learners perceive a conceptual mismatch. In this case, the gestures used in both conditions, congruent and incongruent were familiar to participants as there was an equal level of engagement. Hence, the results obtained in this study could not be explained by the self-involvement or the motor-trace accounts. The motor-imagery theory would be the unique pointing out differences between the use of congruent and incongruent iconic gestures that rely on the meaning match or mismatch between gestures and words.

Taken together, previous studies have confirmed the positive effect associated with the use of congruent gestures on FL vocabulary learning. However, considering the experimental conditions included in past studies there is a lack of empirical evidence comparing the consequences of several conditions in a single study within a single participants sample. For example, in different studies, the meaningless gesture condition is not included in the experimental design. In this context, it is not possible to determine the degree to which the congruent and incongruent conditions produce a facilitation or an interference effect respectively due to the lack of a baseline condition (see [110] for a review of a comparable experimental paradigm concerning Stroop tasks). We addressed this concern by developing Experiments 1 and 2.

In general, previous studies seem to point out that congruent iconic gestures produce a learning facilitation effect with increased recall when movements accompany new words in the codification process. However, does this effect solely manifest when FL learners perform gestures themselves? Previous research indicates that the mere observation of movements can activate brain areas associated with motor actions [111,112]. Hence, while executing movements appears to enhance learning, it remains unclear whether a learning approach involving self-generated actions would yield an additional benefit beyond the mere observation of movements. This possibility will be explored in the next paragraphs.

Effects of “seeing” and “acting” gestures while learning a FL

In the realm of education, the potential advantages of learning through actions compared to observation-based learning have been under discussion for decades [113]. The perspective of "learning-by-doing" advocates for active individual engagement in the learning process by executing actions while learning is taking place. Learning-by-doing can have a positive influence on the formation of neural networks underlying knowledge acquisition and the performance of cognitive skills [114]. This beneficial effect has been demonstrated across various educational domains such as online courses, language acquisition, learning through play, and new technology utilization [98,99,115]. However, as reported in the next paragraphs, empirical evidence is not fully consistent about the advantages of self-generated movements compared to the mere observation of actions performed by others.

In this context, various studies have examined the differences obtained when participants reproduce the experimental tasks by themselves or when they merely observe the experimenter [116,117]. Previous studies have pointed out that self-generated movements enhance cognitive processing. Out of the linguistic context, Goldin-Meadow and colleagues [114], directly compared the impact of self-generated gestures versus observing another individual producing them when children were required to perform a mental transformation task. In their study, children had to perform a mental rotation task, determining whether two shapes presented in different orientations were the same figure. This task was chosen due to the close connection between mental rotation and motor processing. When mentally rotating a target, premotor areas involved in action planning become active [48,118], and participants naturally and spontaneously used gestures when explaining how they solved the task [119]. Goldin-Meadow et al. [114] demonstrated that children achieved better results when instructed to enact the gesture required to solve the transformation task, rather than simply observing the experimenter's movements.

If we move to the field of language learning, other empirical studies have also highlighted the importance of self-generated movements in acquiring linguistic material. The production of spontaneous gestures seems to be more beneficial than observing non-spontaneous movements [18]. Comparing pictures and gestures, the production of gestures associated to new words is more efficient than the use of picture-word associations. This effect is interpreted as a result of the inclusion of motor traces in the semantic of new words while acquisition is taking place [120]. In 2011, James and Swain [121] planned a study in which children were taught action words associated with tangible toys. Children who manipulated the objects during learning exhibited activation in motor brain areas when hearing the words they had learned. Thus, performing motor actions enhances new word learning, and this benefit is likely attributed to the formation of a motor trace that become activated during subsequent information recall. In the same line and at a higher level of linguistic processing, Engelkamp and colleagues [122] required participants to learn sentences while physically performing the actions described in these sentences, as opposed to a solely listening and memorizing condition. The findings revealed that sentence recall was higher when participants engaged in actions during the learning phase. Again, this was interpreted as the performance of actions promoting the formation of a motor trace that facilitated information retention and recall.

On the other hand, Stefan and colleagues [112] reported motor cortex involvement during the observation of movement (e.g., simple repetitive thumb movements), leading to a specific memory trace in the motor cortex akin to the activation pattern during actual motor actions. If there is an activation overlap between movements observation and performance, is the advantage associated to generating movements that clear? In fact, previous research has found confounding results or similar outcomes when participants engage in producing actions or when they solely observe actions performed by others [111]. At the lowest level of linguistic processing, the production versus the mere observation of hand gestures have limited impact on learning segmental phonology or phonetic distinction in a FL [35,75]. In more advanced linguistic stages, where hand gestures have demonstrated a positive influence on learning, various studies show that self-generated movements and the observation of gestures yield comparable outcomes. For instance, learning of anatomy lectures was similar when the instructor executed movements related to the lecture content compared

to the situation in which the students imitated these motor actions [123]. In the specific context of FL acquisition, Baills and collaborators [124] found similar results during the learning of Chinese tones and words when pitch gestures (metaphoric gestures mimicking speech prosody) were used. Participants exhibited similar performance levels when required to learn through observation (Experiment 1) or production (Experiment 2). In a recent study, undergraduate English speakers were presented with 10 Japanese verbs acoustically, while an instructor accompanied them with iconic gestures. Comparable outcomes were observed when participants learned the words solely by observing the instructor's gestures or by mimicking her movements [125]. In a midway position, Glenberg and colleagues ([46], Experiment 3) investigated the impact of movements on sentence reading comprehension in children and found intermediate outcomes. Children were presented with narratives set in a particular scenario (a farm) involving different elements (e.g., a sheep or a tractor). One group manipulated the objects mentioned in the text, while another group imagined doing so. The manipulation condition yielded better results, while the imagined condition showed modest improvement compared to a read-only condition.

In summary, the self-generation of movements during learning appears to yield positive effects in both non-linguistic tasks [114] and FL instruction [18,120,122]. The creation of a more comprehensive semantic representation in memory, encompassing both verbal and motor information, facilitates greater accessibility to previously acquired knowledge. This underlies the advantageous impact of performing movements while learning is taking place [121,122]. Nevertheless, other studies propose that merely observing gestures is sufficient for learning, regardless of whether participants engage in the gestures themselves [125]. This conflicting pattern of outcomes may stem from methodological differences between studies, such as participant demographics, including children [114] versus undergraduate students [75], and the nature of the learning tasks, ranging from dialogic tasks [18] to segmental phonology [35]. Notably, many studies supporting the positive effects of self-generated gestures involve semantically rich materials like words [120] or sentences [122], while some studies indicating no differences between gesture observation and production focus on non-semantic linguistic levels (e.g., segmental phonology, [35,75]) or the manipulation of the gesture conditions is conducted in different experiments [124]. An additional purpose of our experimental series was to address these aspects comprehensively, investigating the performance versus observation of gestures in different groups while participants learn words accompanied by iconic gestures conveying semantic information.

Gestures in verbs and nouns learning

When considering studies on FL vocabulary learning, there are many theoretical and practical issues that must be attended. In the case of learning and gestures, the close relationship between movements and verbs has an special role regarding codification and recall processes. The GSA framework stablish specific predictions about the role of gestures on learning different types of word. This theory posits that gestures emerge from simulated action and perception, which underlie mental imagery and language production [52]. While viewing the size or shape of an object (nouns) does involve simulated movements, the connection between verbs and actions is more pronounced. Consequently, gestures would exert a more significant influence on verb learning compared to noun learning. In fact, it has been proposed that nouns and verbs lexical acquisition mechanisms might implement a bipolar approach. In this way, the cognitive mechanisms for nouns and verbs acquisition would be different at least for children [126].

In a study conducted by Hadley and colleagues [127] with preschool children, the impact of gestures on teaching different word types was directly examined. Results revealed that while concrete nouns exhibited higher learning rates, employing gestures during verb instruction acted as a scaffold for the accompanying verbal contents. This insight elucidates why the majority of research evaluating gesture effects in FL learning has employed verbs as instructional material [63,95,107]. Many verbs (e.g., those describing actions involving manipulable objects) closely correlate with movements [95]. In fact, prior studies confirm that the semantic representation of verbs inherently includes a gestural or motor component [63,128–130].

However, beyond gestures, it has been found that nouns are easier to learn than verbs. Concrete nouns possess distinct perceptual attributes that facilitate the acquisition of new words, whereas verbs convey dynamic information that facilitates the extraction of the motion meaning [131]. Numerous studies have demonstrated that children acquired English nouns more effortlessly than verbs in a natural context [127,132–134]. However, this phenomenon seems to be culturally specific and although it appears in English speaking cultures, the effect is blurred in oriental cultures (see [134–136] for the absence of this advantage in Korean and Mandarin languages). A possible explanation for these crosslinguistic differences would be the especial emphasis that English speakers place on nouns when they interact with children during the acquisition of L1 words.

To the best of our knowledge, the study by García-Gómez and Macizo in 2019 [137] was the first research that directly compared nouns and verbs in the context of adult individuals learning FL words with accompanying gestures. This study will be reviewed in the current paper.

The current series of studies

In the studies described below, we wanted to shed light on strategies that could help learners on the road to FL vocabulary acquisition. In this way, the current review paper explores the potential role that gestures have on FL vocabulary learning. Taking into account all the discussion previously offered the introduction section, we firstly present the experimental conditions that we have used across studies and then, specific factors and predictions addressed in each experiment. The whole of our research allowed us to provide specific contributions to the field and to formulate further predictions described in the subsequent paragraphs.

Previous research suggests that gestures play a role in facilitating the acquisition of FL vocabulary. The available data indicates that gestures contribute to the formation of a motor trace linked to word meanings, involving procedural memory which, in turn, supports FL acquisition [88]. However, it is important to note that alternative explanations cannot be dismissed, and these accounts may complement each other in elucidating the impact of gestures on FL vocabulary learning. Furthermore, it is plausible that gestures might have both positive and negative effects on the learning process, as individuals potentially face a dual-task situation while learning with gestures.

In our studies, native Spanish speakers (L1) learned foreign words (FL) in an artificial language, Vimmi, developed by Macedonia and collaborators [90], over three consecutive learning sessions. The words were presented in various conditions: alone (no gesture condition), accompanied by meaningless or unfamiliar gestures (meaningless gesture condition), or presented alongside iconic gestures, where the meanings of the gestures were either semantically congruent or incongruent with the meanings of the words (congruent and incongruent conditions, respectively). Finally, our participants were evaluated after each learning session by completing a forward (L1-FL) and backward (FL-L1) translation task.

When considering the theoretical perspectives on the role of gestures in FL vocabulary acquisition (previously described), namely the self-involvement account (Helstrup, 1987), the motor-trace account [85,86], and the motor-imagery account [89], it becomes possible to formulate specific predictions attending to our experimental conditions (see Figure 1 to observe the predictions of each of the theories for the learning conditions included in our studies).

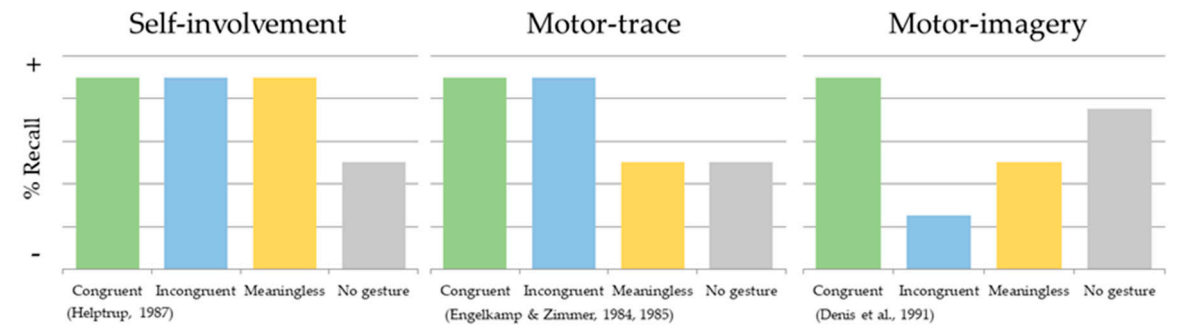


Figure 1. Visual description of expected results (percentage of FL recall after instruction) for each learning condition used in our studies according to the three main theories about the role of gestures on FL vocabulary acquisition.

- If gestures primarily serve to enhance the participant's engagement in the learning tasks, then all conditions involving gestures should lead to improved FL vocabulary acquisition compared to the condition without gestures.
- If the motor trace left by gestures aids participants in learning new words, familiar gestures (in the congruent and incongruent conditions) should be associated with better FL vocabulary learning outcomes than less familiar gestures (meaningless condition) and the condition without gestures.
- Furthermore, the motor-imagery account suggests that learning meaningful gestures could either facilitate or interfere with vocabulary acquisition, depending on the alignment between the gesture's meaning and the meaning of the FL word being learned. Thus, congruent gestures may enhance vocabulary learning, while incongruent gestures might hinder the acquisition of new words. Conversely, meaningless gestures could become distinctive and contribute to the encoding and recall of FL words [138].

As noted before, it is worth noting that these perspectives are not mutually exclusive. For example, the acquisition of FL words paired with congruent gestures might involve a trade-off between the positive impact of strengthening the connection between semantic information and FL words and the negative consequence of engaging participants in a dual-task situation.

In Experiments 1 and 2, we established a direct comparison between nouns and verbs learning. The prevailing notion is that verbs might be more difficult to acquire compared to nouns [129,135,139]. Additionally, it is commonly accepted that action verbs inherently incorporate a gestural or motor component within their mental representation [128,130]. Therefore, disparities between these two words types may hinge on whether they involve overt bodily movements, as is the case with action verbs. For instance, De Grauwe et al. [63] observed that the comprehension of motion verbs (e.g., "to throw") in FL triggered activation in motor and somatosensory brain regions. Consequently, it is plausible to speculate that the impact of gestures in FL vocabulary acquisition might be more pronounced with verbs than with nouns.

In addition, there exists a debate surrounding the impact of self-generated movement compared to the mere observation of gestures on FL vocabulary acquisition. To directly address this issue, two experimental groups were included in Experiment 3. Participants were randomly assigned to the "see" and "do" learning group. In the "see" learning group, learners were required to read aloud Spanish-Vimmi word pairs (L1-FL) while simultaneously observing and mentally envisioning themselves replicating the gestures depicted in a video. On the contrary, in the "do" learning group, learners were instructed to say aloud the word pairs in both Spanish and Vimmi (L1-FL) and physically mimic the gestures as they were presented on the screen. Only action verbs served as learning material in this experiment.

As previously mentioned, certain studies demonstrate enhanced learning outcomes associated with the active generation of gestures [18,120,122]. Conversely, in other research, no discernible difference is observed between the act of viewing gestures and the act of performing gestures by FL learners [35,75,125]. In this work, we controlled for several factors that might explain differences found in previous studies (type of learning material, type of gestures or word-gesture meaning relation). Unlike the vast majority of research on the field, we explore the effect of self-generated movements in adult population [46,114,120,121]. It is crucial to emphasize this methodological distinction because adults generally possess more experience in executing actions compared to children. Furthermore, for adults, the semantic content of gestures and the visual imagery associated with words tend to be rich [140]. Given this consideration, merely observing movements resembling action verbs might suffice for adult individuals to harness the beneficial effects of gestures in FL learning. Consequently, the act of observing gestures could serve to strengthen the connections between FL words and the semantic system in a manner akin to actually performing the gestures

themselves, as proposed by Sweller et al. [125]. In other words, we hypothesized that adult participants may not exhibit any discernible difference between the act of viewing gestures and the act of physically performing those gestures in terms of FL vocabulary acquisition.

Regarding the learning results, our general hypothesis was fitted to the principles of the Revised Hierarchical Model (RHM) [16]. The translation task allowed us to establish conclusions at the level of lexical and semantic access to the new learned words. First, we expected an asymmetrical effect, as seen in prior studies, with more efficient performance in backward translation compared to the forward translation [16]. This prediction arises from the fact that forward translation necessitates more semantic processing than backward translation, making more challenging to translate from L1 to FL than the reverse. In general, it was predicted that in learning conditions where semantic processing is promoted, facilitation would be particularly noticeable in forward translation because it is semantically mediated [16,141,142]. In terms of potential interactions between the translation direction and the gesture conditions (congruent, incongruent, meaningless, and no gestures), we anticipated that forward translation would be more affected by the semantic congruence (congruent and incongruent conditions) compared to the condition with meaningless gestures.

Finally, we anticipated (a) a positive effect (learning facilitation) when employing congruent gestures during FL word acquisition (congruent condition), and (b) a negative effect (learning interference) associated with gestures unrelated to word meanings (incongruent and meaningless conditions). This pattern of results may align with the postulates of the motor-imagery account [138] and with previous studies that have illustrated both the benefits and drawbacks of using gestures in FL vocabulary acquisition [90,94,95]. In addition, regarding the focus of Experiment 3, if the mere act of observing gestures proved to be sufficient to enhance vocabulary acquisition, the outcomes would be consistent across both learning methods. Conversely, if active engagement in gestures during instruction, as seen in the “do” learning group, maximized learning, we would anticipate a higher learning rate in this group compared to the “see” learning group.

2. Materials and Methods

Participants

First of all, regarding the type of population, in all of our studies, we selected Spanish “monolingual” speakers. All the participants were young-adult students from the University of Granada that received course credits as reward for participating. Nowadays it is confusing to classify a Spanish person as monolingual. As is often the case in other countries, the younger generations are exposed to foreign language learning in regular education. In this context, we decided to recruit participants who were as minimally proficient in any FL as possible. To achieve this, we established the following inclusion criteria regarding the participants FL proficiency level. On a daily basis, they needed to confirm that:

1. They had no contact with any language other than Spanish, whether spoken or sign language.
2. Their most recent exposure to a FL had to occur during high school.
3. They had never received any formal instruction in a FL beyond regular education.
4. They had never obtained a certification in any FL.

The language proficiency level of the participants was relevant for the experimental design in these studies. Previous research has shown that there are differences in the way new languages are acquired depending on how proficient speakers are in other languages [143–145]. For instance, learning a third language (L3) is generally assumed to be less demanding or costly than learning a L2 [146–148]. Existing literature suggests that bilingual experience confers individuals with tools that facilitate the L3 learning process (for a comprehensive review, see [149]).

A total of 25 individuals participated in Experiment 1, consisting of 21 women and 4 men. The average age of the participants was 21.72 years, ($SD = 3.17$). In Experiment 2, a total of 32 students were involved (6 men and 26 women). Their mean age was 20.97 years ($SD = 3.21$). Finally, in Experiment 3, 31 participants were recruited. Sixteen of them (15 women and 1 man) were randomly

assigned to the “do” group. Their mean age was 21.12 ($SD = 2.53$). The remaining 15 participants (13 women and 2 men) formed the “see” learning group. Their mean age was 21.13 years ($SD = 2.72$).

All participants provided written informed consent before engaging in the experiment. None of the participants reported history of language disabilities, and they all had either normal visual acuity or corrected-to-normal visual acuity. The data obtained in the studies were treated anonymously by the researchers of this paper.

Materials

Regarding the materials used in the experimental design, the same gestures pool was used across studies. This material was specifically created by the experimenters. In addition, 40 Spanish nouns and 40 Spanish verbs were selected to act as L1 words. Participants had to learn the translation for these words in a FL. As FL, we selected an artificial corpus described below (Vimmi, developed by Macedonia et al. [90]).

The gestures used in the studies included only hand movements. We implemented iconic gestures depicting common actions people normally perform when interacting with objects (e.g., mimicking writing a letter or brushing hair) [38,150]. The videos showing the hand gestures were recorded by the first author of all the studies. These gestures were used in the congruent and incongruent gesture conditions. Furthermore, in the meaningless condition, gestures consisted of small hand movements without iconic or metaphoric connections to the meanings of the accompanying word (e.g., forming a fist with one hand and raising the fingers of the other hand). Care was taken to ensure that these meaningless gestures shared similar characteristics with meaningful gestures, such as hand configuration, the use of a simple movement trajectory, and spatial location. In this condition, 10 different movements were selected, and all participants were exposed to the same set of gestures.

Furthermore, we aimed to ensure that the congruent, incongruent, and meaningless conditions varied in the extent to which the semantics of the word corresponded to the accompanying gesture. To achieve this, a group of 15 Spanish participants who did not participate in any of the experiments took part in a pilot study. They were presented with a video displaying a gesture (without sound) at the top of the screen and a word written in Spanish at the bottom of the screen. They were then instructed to rate the degree of alignment between the meaning of the word and the gesture, using a scale from 1 (indicating a high mismatch) to 9 (indicating a high match).

In Experiment 1, there were differences between nouns included in the congruent condition, meaningless condition, and incongruent condition. The gesture-word pairs were rated higher in the congruent condition compared to the meaningless condition, and the incongruent condition. The incongruent condition and the meaningless condition also differed. 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. In Experiments 2 and 3, significant differences were found between verbs included in the three conditions: congruent condition, meaningless condition, and incongruent condition. Specifically, the gesture-word pairs received significantly higher ratings in the congruent condition compared to both the meaningless condition and the incongruent condition. Furthermore, the incongruent condition and the meaningless condition also differed. Thus, the three conditions involving gestures in the study differed in terms of the association between the word's meaning and the accompanying gesture.

As FL to be learned, we selected an artificial language called Vimmi [66,90]. The corpus of Vimmi words has been designed to eliminate factors that could potentially bias the learning of particular items. This includes avoiding patterns like the co-occurrence of syllables and any resemblance to words from languages such as Spanish, English, and French. Vimmi words were meticulously chosen to be pseudowords in the L1 of the participants (Spanish), thus maintaining proper orthography and phonology in Spanish but without meaning (see [151], for a discussion of these variables in vocabulary learning). As reported by Bartolotti and Marian [152], divergent linguistic structures across languages can hinder new vocabulary acquisition. Vimmi matches the orthotactic probabilities of Latin languages such as Spanish, French or Italian. When the set of Vimmi words was selected in

our studies, different linguistic variables were controlled such as number of graphemes, phonemes and syllables.

Furthermore, Spanish words were selected to act as L1. In Experiment 1, all of them represented concrete nouns denoting objects that could be manipulated with the hands (e.g., spoon, comb, etc.). In Experiments 2 and 3, the same pool of 40 Spanish verbs were used to be coupled with the FL translations. Linguistic variables such as number of graphemes, phonemes and syllables, lexical frequency, familiarity or concreteness were carefully controlled.

Experimental Design

The studies involved four FL vocabulary learning conditions that were manipulated within participants as follows. In Figure 2 we show an example of the experimental design for each learning condition:

- a) *Congruent condition*: L1-FL word pairs were presented with gestures that reflected the common use of objects whose names had to be learned in FL. For example, "teclado (keyboard in Spanish)-saluzafo (Vimmi translation)" was coupled with the gesture of typing with both hands fingers as if we had a keyboard in front of us.
- b) *Incongruent condition*: L1-FL word pairs were linked to gestures associated with the use of an object different from that denoted by the L1 word. For instance, "teclado-saluzafo" was paired with the gesture of striking something with a hammer.
- c) *Meaningless condition*: L1-FL word pairs were paired with unfamiliar gestures. For instance, "teclado-saluzafo" was accompanied by a gesture of touching the forehead and then one ear with the right forefinger.
- d) *No gesture condition*: Participants had to learn Spanish (L1)-Vimmi (FL) word pairs without the use of gestures. For example, they had to associate "teclado" with "saluzafo".



Figure 2. Learning conditions used in the series of studies. Spanish (L1) – Vimmi (FL) words (nouns in this example) are coupled with different gesture conditions. In the example, *teclado* ('keyboard' in Spanish) – *saluzafo* (a Vimmi word) were accompanied by (a) the gesture of moving both hands fingers as if you were typing on a keyboard (congruent condition); (b) the gesture of striking something with a hammer (incongruent condition); (c) the gesture of moving the hand from the forehead to the ear (meaningless condition); (d) no gesture (no gesture condition).

To create the learning material, the 40 Spanish words were randomly paired with the 40 Vimmi words that were previously selected. This resulted in a total of 40-word pairs, with each pair consisting of a L1-Spanish word and a FL-Vimmi word. These 40-word pairs were then randomly divided into four sets, each containing 10-word pairs. Each set of 10 pairs was associated with one of the gesture conditions: congruent, incongruent, meaningless, or no gesture.

To ensure a balanced distribution of gesture conditions across the word sets, four lists of materials were created. Hence, a word pair (e.g., "teclado-saluzafo") was linked with a congruent gesture in list 1, an incongruent gesture in list 2, a meaningless gesture in list 3, and presented without gesture in list 4. Each participant was randomly assigned to one of these four lists. Consequently, across the lists, all 40-word pairs were evenly distributed over the four training conditions, ensuring that across participants, every word pair appeared in all of the training conditions.

In this series of studies, we equated the Spanish nouns and verbs across the four sets of word pairs in lexical variables. There were no significant differences across the nouns sets in terms of: number of graphemes, number of phonemes, number of syllables, lexical frequency, familiarity, or concreteness. Likewise, Vimmi words in the four sets were made equivalent in terms of number of graphemes, number of phonemes, and number of syllables. Furthermore, we ensured that the similarity between the Spanish words and Vimmi words was consistent across sets of word pairs. The number of shared phonemes between the Spanish and Vimmi words remained the same across the four sets, both when phoneme position was considered and when it was not.

Procedure

The learning phase involved three training sessions conducted on three consecutive days. In each session, participants first completed the FL training and then the evaluation of new words learning. The two phases were separated by a 15-minute break. E-prime experimental software was used for stimulus presentation and data acquisition [153]. Participants were informed that the training sessions would be recorded on video to ensure they followed the instructions provided by the experimenter. The procedure performed in this study was approved by the Ethical Committee on Human Research at the University of Granada (Spain) associated with the research project (Grant PSI2016-75250-P; number issued by the Ethical Committee: 86/CEIH/2015) awarded to Pedro Macizo. It was conducted in accordance with the 1964 Helsinki Declaration and its subsequent amendments.

The design of the learning and evaluation phases remained consistent across experimental studies. However, we introduced certain variations that are specified in the next sections.

Vocabulary Learning Phase

The learning phase lasted approximately 1 hour per session. We employed a stimulus presentation procedure organized by experimental conditions, following a similar approach to that used in other studies with various gesture conditions [90]. Participants were presented with a block of 40 Spanish–Vimmi word pairs, with each block containing 10 word pairs for each of the four learning conditions (congruent gestures, incongruent and meaningless gestures, and no gestures). This block was repeated 12 times, resulting in each participant receiving 480 trials, where the 40 word pairs were presented 12 times. Short breaks were introduced between learning blocks. Word pairs were randomly presented within each condition, and the order of learning conditions within a block was counterbalanced. This blocked design was adopted to minimize the cognitive effort associated with constantly switching between conditions where participants had to perform gestures and those without gestures.

The gestures were recorded on video by the experimenter and varied in nature, being congruent, incongruent, or meaningless depending on the accompanying word and thus, the learning condition. Each recorded gesture had a duration of 5 seconds and the gesture was repeated twice. The video appeared in the middle top part of the screen. In addition, in all experimental conditions, participants were presented with a Spanish–Vimmi (L1-FL) word pair visually displayed at the bottom of the screen. These word pairs were presented with the experimenter in a static position (no gesture condition) or accompanied by a gesture (remaining conditions) (see Figure 2).

Participants were instructed to read aloud each L1-FL word pair twice. In the gesture conditions, participants were also required to produce the gesture that accompanied the word pair each time they vocalized it. Participants initiated the gesture production simultaneously with the vocalization of the L1-FL word pair, ensuring that each gesture and word pair were produced twice. For example, when participants received the word pair “teclado-saluzafó” along with the congruent gesture (Figure 2a), they were instructed to vocalize the word pair while simultaneously performing the corresponding gesture of moving fingers as if typing on a keyboard. After producing the word pair twice while mimicking the experimenter movements also twice, participants pressed the space bar to proceed to the next trial.

As previously mentioned, we implemented some variations to adapt the procedure to the goal of each study. Experiment 1 included nouns as learning material while in Experiments 2 and 3, verbs

served as words to be learned. In addition, the learning procedure previously described was the same for all experiments but in the case of the “see” learning group (Experiment 3), the participants were instructed to verbally say aloud each L1–FL pair twice while mentally simulating the gesture presented with the word pair, but without performing it overtly. For instance, when participants encountered the word pair “teclado-saluzafó” paired with the congruent gesture, they were expected to say aloud this word pair while concurrently generating a mental image of themselves moving fingers as if typing on a keyboard. This was maintained in the 3 gesture conditions (congruent, incongruent and meaningless conditions) but the participants did not perform any movement in the condition without gestures as occurred in the rest of the experiments and in the “do” learning group. The reason for instructing participants in the “see” learning group to mentally recreate the gesture was to ensure their engagement with the gesture and prevent them from solely focusing on learning the words in Vimmi. While it was challenging to confirm a priori whether participants were indeed mentally visualizing the required gesture, post-experiment observations, indicate that participants did, in fact, follow the experimenter's instructions and mentally executed the prescribed gestures.

Evaluation Phase

To evaluate the acquisition of Vimmi words two tests were employed in all the studies: Translation from Spanish into Vimmi (forward translation from L1 to FL) and translation from Vimmi into Spanish (backward translation from FL to L1). These tasks have been used previously in FL learning studies [19,154].

To prevent any potential order/practice effects, the order of presenting the translation tests was randomized across the three training sessions and among participants. In each translation task, participants received 40 Spanish words for forward translation and 40 Vimmi words for backward translation. On each trial, a word appeared in the center of the screen until the participants' response. Oral translations were recorded for subsequent accuracy analysis, and response times (RTs) were recorded from word presentation to the start of oral translation. The learning assessment took approximately 10 minutes, depending on individual performance.

3. Results

In this section, the common issues regarding the analysis protocol are described first. Then, the results of Experiments 1 and 2 are presented followed by the outcomes of the two learning groups evaluated in Experiment 3.

While recall percentages (Recall %) serve as the primary measure of vocabulary acquisition, we also analyzed RTs in all the studies. RTs associated with correct translations were subjected to trimming following the procedure outlined by Tabachnick and Fidell [155] to remove univariate outliers. Raw scores were converted to standard scores (z-scores), and data points that, after standardization, deviated by 3 standard deviations from the normal distribution were identified as outliers. Outliers were removed until no observations exceeded the 3 *SD* threshold. In all analyses, we maintained a significance level of $\alpha = .05$. Only correct responses were included in the RT analyses. Data points were excluded from RT analyses under the following situations: (1) participants produced nonverbal sounds that triggered the voice key, (2) participants stuttered or hesitated in producing the word, (3) participants produced something other than the required word. Some minor errors were allowed and considered correct responses depending on the length of the correct word to be produced:

- For monosyllabic words, the replacement of a vowel.
- For disyllabic words, the replacement of a vowel or a consonant, but not both.
- For words with three or more syllables, the inversion of a vowel and a consonant or the replacement of a vowel or a consonant.

Low-level (sublexical) errors, such as the replacement of vowels that did not introduce new semantic content, were considered correct. Since the FL words were in an artificial language (Vimmi),

when participants replaced a vowel in Vimmi (e.g., 'rel') to create a legal word (e.g., 'rol'), this type of response was regarded as an error (comprising less than 5% of total errors).

Results of Experiment 1

Several factors treated as within-participant variables were entered in Analyses of variance (ANOVA). These factors were the translation test (comprising forward translation and backward translation), the training session (comprising the first session, second session, and third session), and the learning condition (which included congruent, incongruent and meaningless gestures, and the no gestures condition). Initially, the order in which participants received the translation tasks, either in a forward-backward translation order or a backward-forward translation order, was included as a between-subject factor in the analysis. However, the translation order was not significant nor did it interact with other factors, so this factor was not considered any further.

Response time analysis revealed that participants learned the new words properly across the three acquisition sessions. They responded faster on the last learning session compared to the first day of training. The recall analysis showed a wider array of main effects and interactions. In line with the response time results, participants remembered more words at the end of the training compared to the first evaluation phase. Figure 3 illustrate the interaction between translation direction and learning condition in each learning session. In the forward translation direction (L1-FL), performance improved across sessions. The main effect of learning condition was modulated by the translation test (see Figure 4). The comparison between learning conditions showed better recall in the congruent gestures condition compared to the meaningless gestures condition. The meaningless, incongruent and no gesture conditions showed similar outcomes. Compared to the no gesture condition, the recall was lower in the meaningless condition. The backward translation direction revealed a similar session effect showing that the more participants trained, the more they learned. Comparing learning conditions, congruent gestures improved performance compared to the meaningless condition. No differences were observed between the meaningless and incongruent conditions. Compared to the no gesture condition, the recall was lower in the incongruent and meaningless conditions.

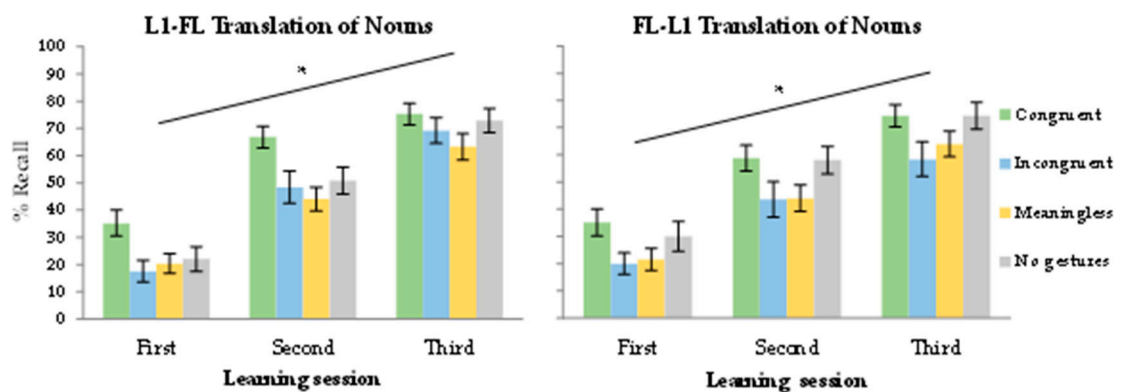


Figure 3. Recall percentages (% Recall) obtained in Experiment 1 during the translation of nouns from L1 to FL and from FL to L1 across training sessions (first, second, and third) and learning conditions (congruent, incongruent, meaningless, and no gestures). Vertical lines represent standard errors. * $p < .05$.

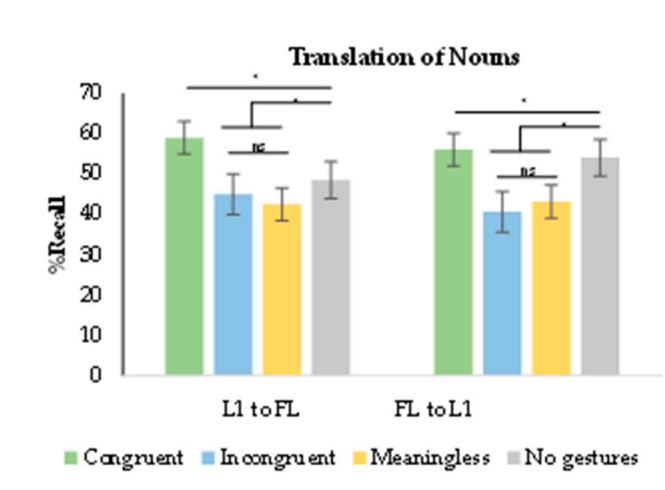


Figure 4. Recall percentage (% Recall) of nouns as a function of translation direction (L1 to FL, FL to L1) and gesture conditions (congruent, incongruent, meaningless, and no gestures). Standard errors are plotted in vertical lines. * $p < .05$, $^{ns}p > .05$.

In Experiment 1, when investigating the learning of FL nouns across various training conditions, two primary effects emerged. When participants learned FL words in the congruent condition, there was a notably higher recall percentage compared to the meaningless condition. However, the recall of FL words was lower in both the incongruent and meaningless conditions relative to the no-gesture condition, as shown in the backward translation task. This pattern of results suggests that these conditions had a detrimental impact on the learning process. Altogether, these findings reveal two contrasting effects when comparing different methods of FL learning: facilitation and interference. Figure 4 shows the % Recall of nouns as a function of translation direction and gesture conditions (learning session collapsed).

The facilitation effect observed with congruent gestures appears to lend support to the motor-imagery account of gestures' role in FL vocabulary learning (as proposed by [89]). The shared semantic meaning between gestures and L1 words seemed to enhance the acquisition of FL words. Another perspective, the motor-trace theory [85,86], could also explain the facilitation effect, as congruent gestures were familiar and could activate motor traces associated with the words. However, this perspective would not account for the interference effect found with incongruent gestures relative to the no gesture condition, given that incongruent gestures were also common and familiar to the participants. Furthermore, the self-involvement explanation could not accommodate the observed pattern of results, as there were clear differences between conditions involving gestures (as indicated by Helstrup [79]).

Notably, the magnitude of the interference effect was similar in both the incongruent and meaningless conditions when compared to the learning of nouns in the no-gesture condition. This observation suggests that the negative impact of gestures in these two conditions may be attributed to the fact that participants were engaged in a dual-task setting, which heightened the difficulty of the learning process (unless it provided consistent information as in the congruent condition).

Results of Experiment 2

The same analysis protocol and set of factors used in Experiment 1 were employed in Experiment 2. The sole distinction between the two studies was the category of words participants were tasked with learning. In Experiment 2, students acquired verbs instead of nouns. When comparing the acquisition of nouns and verbs, previous research has consistently shown that verbs tend to be more challenging to learn [129,135,139]. To address this inherent difficulty associated with verb acquisition, one potential approach is to incorporate gestures during the learning process. It has been posited that the semantic representation of verbs intrinsically involves a motor component

[128,130]. Consequently, in Experiment 2, we evaluated whether the use of gestures could alleviate the difficulty of learning verbs vs. nouns.

When it comes to RTs, learners responded faster in the last training session compared to the first learning day. Regarding learning conditions, congruent gestures exhibited faster responses than the incongruent, meaningless and no gesture conditions. No differences were found between any other learning condition. Recall was better in the last learning day compared to the beginning of the training (see Figure 5 to observe the % Recall across sessions, translation directions and learning conditions). There was also a main effect of translation direction that interacted with the learning condition factor. This effect is shown in Figure 6 (learning session collapsed). In the forward translation, the session effect previously described persisted. Better recall was observed in the congruent condition relative to the meaningless and no gesture condition. No differences were found between the incongruent, meaningless and no gesture conditions. The backward translation direction showed again the increasement in percentages of recall across learning sessions. Better recall was observed in the congruent condition relative to the no gesture and meaningless conditions. No differences were observed between meaningless and incongruent gestures. However, compared to the no gestures condition, the recall was lower in the incongruent condition and the meaningless condition.

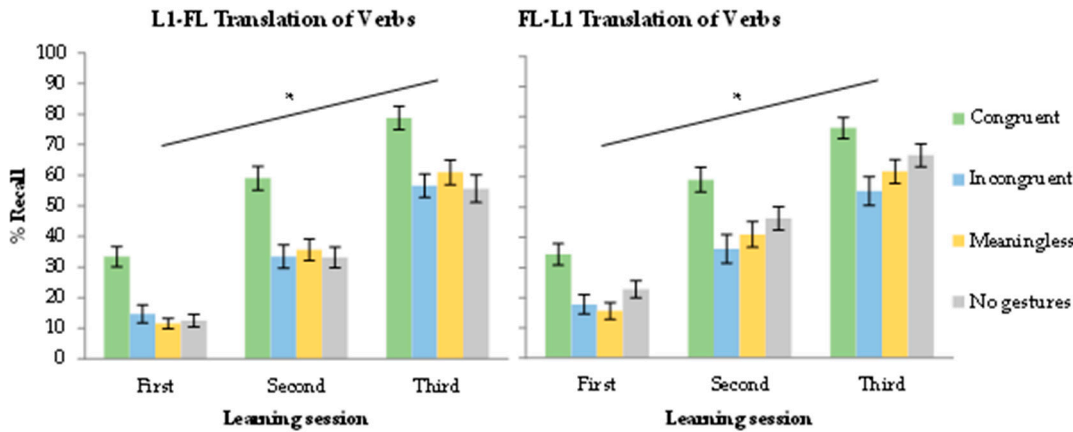


Figure 5. Recall percentages (% Recall) obtained in Experiment 2 during the translation of verbs from L1 to FL and from FL to L1 across training sessions (first, second, and third) and learning conditions (congruent, incongruent, meaningless, and no gestures). Vertical lines represent standard errors. * $p < .05$.

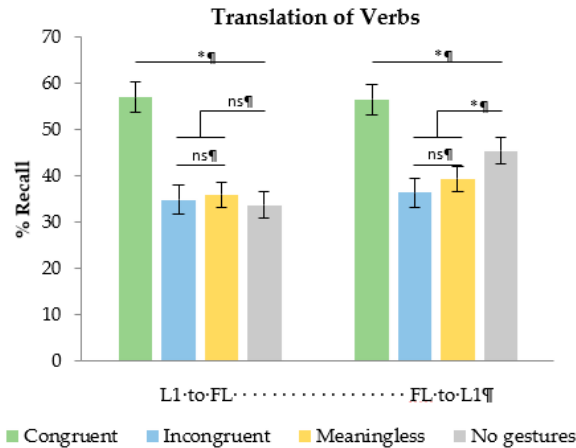


Figure 6. Recall percentage (% Recall) of verbs as a function of translation direction (L1 to FL, FL to L1) and gesture conditions (congruent, incongruent, meaningless, and no gestures). Standard errors are plotted in vertical lines. * $p < .05$, $^{ns}p > .05$.

In Experiment 2, we focused on investigating the role of gestures in the acquisition of FL verbs. The observed pattern of results closely mirrored what we found in Experiment 1, which pertained to the learning of FL nouns.

Once again, we identified a facilitation effect stemming from the integration of gestures into the learning process. Specifically, congruent gestures proved to enhance the acquisition of FL verbs when compared to learning without gestures or with meaningless gestures. However, as in Experiment 1, we also encountered an interference effect when participants were exposed to incongruent and meaningless gestures, as opposed to the no-gesture condition, particularly evident in the FL to L1 translation task. The discussion section will provide a more in-depth explanation of these two effects, shedding further light on their underlying mechanisms.

Between Experiments Comparison

To establish comparisons between word types, which constituted one of the primary objectives of this experimental series, we introduced the word type factor (nouns, verbs) in a new analysis along with the factors previously mentioned. In the forward translation direction, participants recalled more nouns than verbs. Also, the learning condition interacted with the word type factor. In the no gesture and incongruent conditions, better recall was noted for nouns than verbs. However, no differences between nouns and verbs were obtained for congruent and meaningless gestures. Considering the backward translation, no main effects or interactions were significant. Figure 7 shows the results found for nouns and verbs recall in each of the learning conditions.

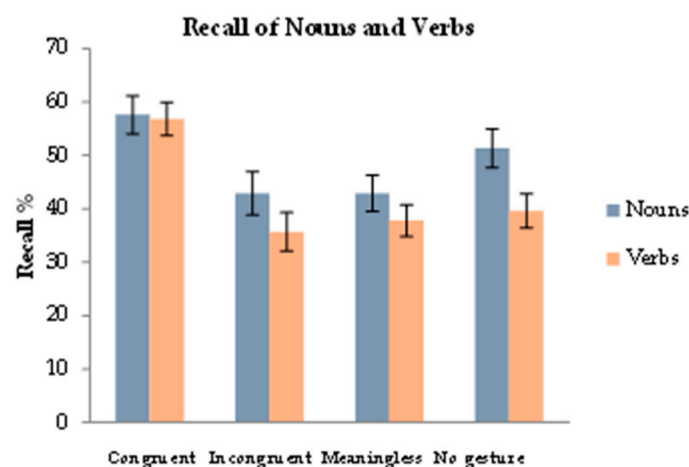


Figure 7. Comparison of recall percentages (% Recall) of nouns (Experiment 1) and verbs (Experiment 2) across gesture conditions (congruent, incongruent, meaningless, and no gestures). Standard errors are plotted in vertical lines.

Results of Experiment 3

We conducted ANOVAs including translation direction (forward translation, backward translation), training session (first session, second session, third session), and learning condition (congruent, incongruent, meaningless, no gestures) that were treated as within-participants factors, while the learning group ("see", "do") was considered a between-groups variable in Experiment 3. Consistently with Experiments 1 and 2, the more participants trained the faster and more accurate they were.

Regarding RTs, the main effect of translation direction was significant so participants were faster in the backward than in the forward translation. The learning condition was also significant and

learners were faster in the congruent condition compared to the meaningless condition and the no gesture condition. No differences were found between incongruent and meaningless gestures. Finally, marginal differences were obtained between meaningless gestures and the no gesture condition with participants responding subtly slower to the meaningless condition. It might be highlighted that participants in the “do” learning group were faster than the ones who learned under the “see” instruction across learning sessions.

Recall results showed that participants were more accurate in the congruent and no gesture conditions compared to the meaningless condition. No differences appeared between the incongruent and meaningless conditions. Finally, the congruent condition exhibited advantages compared to the no gesture condition. In this case, the meaningless condition seemed to be the most detrimental to the participants performance while the congruent gestures facilitated the learning process compared to the no gesture condition. Differences emerged between learning groups with participants included in the “do” group being more accurate than participants in the “see” group. The learning group interacted with learning condition and translation direction and hence, the analyses for each learning group were conducted separately. In the “see” group congruent gestures produced a facilitation effect compared to meaningless gestures. Also, participants remembered more words in the no gesture condition compared to the meaningless condition. Marginal differences were obtained when the incongruent and meaningless conditions were compared being the meaningless gestures the most detrimental learning situation. The final comparison between congruent and no gesture conditions showed a marginal effect. The “do” learning group revealed an interaction between learning condition and translation direction. In this way, the learning condition effect was explored for each translation direction separately. In the forward translation, there was a facilitation effect associated with the use of congruent gestures. No differences were obtained between incongruent, meaningless and no gesture conditions. The learning interference associated with the use of meaningless gestures that appeared in the “see” group, was not present in the case of the “do” group. Advantages were associated with congruent gestures compared to the no gesture condition. Finally, in the backward translation direction, congruent gestures showed better recall patterns compared to meaningless gestures. Differences appeared between incongruent and meaningless gestures but similar results were obtained when the congruent and no gesture conditions were compared (see Figure 8).

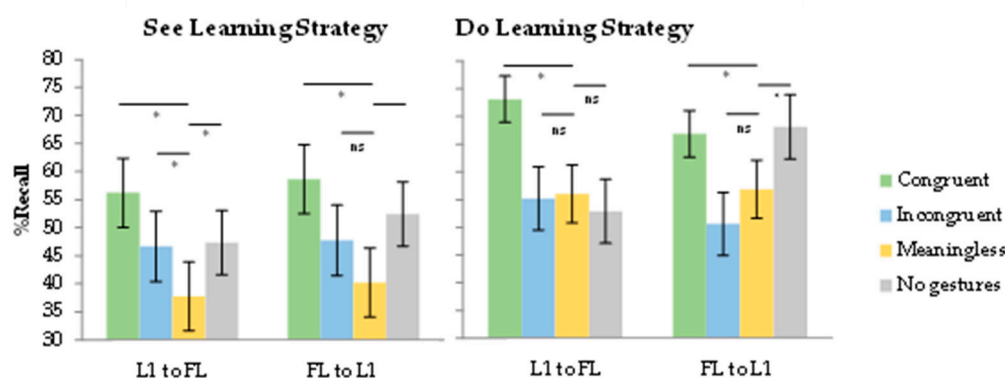


Figure 8. Recall percentage (% Recall) of the ‘see’ (left graph) and ‘do’ (right graph) learning groups as a function of translation direction (L1 to FL, FL to L1) and the gesture conditions (congruent, incongruent, meaningless, and no gestures). Standard error is plotted in vertical lines. * $p < .05$, ns $p > .05$.

4. Discussion

It is commonly accepted that movements appear to play a significant role in a variety of cognitive processes. A facilitative effect has been observed for different types of movements, not only in educational settings but also in clinical contexts such as developmental disorders and aphasia

treatments [156–158]. For instance, pointing movements, defined as deictic gestures, and beat gestures, which reflect prosody and emphasize speech, have demonstrated positive effects on language learning and development [68,69,159]. Particularly noteworthy are iconic gestures, referring to concrete entities or actions. These gestures have been used in many studies to investigate how they enhance memory consolidation in language production and comprehension [76,78] and in FL vocabulary acquisition [69,120]. The findings from the present series of studies indicate that iconic gestures play a crucial role in enhancing semantic processing and linking the semantic system with the lexicon in FL learning. Several pieces of evidence support this conclusion.

Previous studies explored the comparison between congruent and incongruent conditions, employing either speech or gestures, particularly in the context of verb learning [95] and noun learning [90]. Additionally, some works have compared congruent gestures to meaningless gestures when presented alongside written words [102] and sentences [76]. Previous research has also examined conditions with gestures versus conditions without gestures [66]. Moreover, some studies have considered all three conditions —congruent, incongruent, and no gestures— [94]. In addition, the comparison between gesture imitation and picture observation [120] and the contrast between the observation and production of non-iconic gestures, such as pitch gestures and beat gestures [68,124] have been explored. Sweller and colleagues [125] directly assessed the impact of self-performed gestures and gesture observation when using iconic gestures as learning materials along with acoustically presented FL words. As far as we know, our series of studies addressed for the first time the role of four gesture condition (congruent, incongruent, meaningless and no gesture) on FL vocabulary learning comparing nouns and verbs acquisition. In addition, the effects of observing versus producing gestures taking into account the different semantic relationships that can be established between gestures and words (congruent, incongruent, meaningless) had not been evaluated until the publication of our work on this issue. The current experimental design and manipulations holds significant potential for advancing our understanding of the role of gestures in FL vocabulary learning.

Coming back to the theories explaining the potential role that gestures have on FL vocabulary acquisition, the self-involvement explanation failed to accommodate the pattern of results. In line with this perspective, gestures can be seen as a means to enhance the participant's engagement in the learning task. Consequently, whenever gestures are integrated into the learning process, one would expect to observe an improvement in word acquisition. However, in these experiments clear differences were observed among gesture conditions [79].

Firstly, and highly important, regarding the impact of different gesture conditions in our study, in the congruent gesture condition, where both the gesture and the word shared a common meaning, there was a notable increase in word learning [65,94,96]. In general, the congruent gesture condition exhibited a higher recall percentage and faster response times compared to the remaining gesture conditions. This suggests that congruent gestures promote semantic processing, ultimately benefiting the learning outcomes. Additionally, processing congruent gestures alongside word learning mitigated the negative effects of dual-task performance (processing both gestures and words) [160,161]. This effect might be supporting the motor-imagery theory concerning the role of gestures in FL vocabulary acquisition [89]. The shared semantic meaning between gestures and L1 words appeared to enhance the acquisition of FL words. Consistent with previous findings, the FL learning facilitation effect resulting from the processing of congruent gestures remained consistent regardless of the type of training (“see” learning vs. “do” learning). This suggests that mere exposure to gestures is sufficient to observe the beneficial effects of gestures on vocabulary acquisition in a FL [125]. While our behavioral studies did not provide evidence of brain activity, our results aligns with outcomes from various reports, which demonstrate that the mere observation of actions triggers a pattern of brain activation akin to that observed during the performance of motor actions in the motor cortex [112]. Consequently, processing gestures, whether through observation or performance, may enrich the encoding of words to be learned by incorporating sensorimotor networks and procedural memory into the semantic/declarative memory linked to the word meanings [88]. Hence, congruent

gestures can enhance the semantic processing of words and this pattern of findings align with the motor-trace [85,86] and motor-imagery account [89].

Considering the incongruent and meaningless conditions, both the motor-trace [85,86] and motor imagery accounts [89] establish specific predictions about the learning results for each gesture condition. The motor-trace theory supports that only when learners perform familiar gestures, a facilitation effect would be found. Hence, this theory would suggest a reduced performance in the meaningless and no gesture conditions compared to the congruent and incongruent conditions. On the other hand, the motor-imagery account might expect the higher learning interference associated with incongruent gestures followed also by reduced performance in the meaningless conditions compared to the no gesture condition. This difference would arise because, in the incongruent condition, there is higher discrepancy between the meaning conveyed by the gestures and the meaning of the FL words being learned.

Results revealed that the incongruent and meaningless gesture conditions worsened performance compared to the no-gesture condition, revealing a learning interference effect (although no differences were found in the L1-FL translation of verbs in Experiment 2 and in the “do” learning group) [65,90,93,94,105]. The primary distinction between the incongruent and meaningless conditions when compared to the no-gesture condition was the level of engagement or involvement of the participants in a dual-task versus a single-task learning context, respectively. This interference would be indicative of the difficulty associated to the integration of the meaning of words and gestures in working memory [105]. Therefore, the absence of alignment between the information activated by the gesture and the word would promote a conflict situation in working memory, impeding the learning and subsequent recall of FL words. These outcomes would be in accordance with the hypothesis of the motor-imagery account. However, the “see” learning group exhibited an additional interference effect (only in the L1-FL translation direction), with lower recall of words in the meaningless gesture condition compared to the no-gesture condition. As meaningless gestures lacked semantic content, the interference observed in the meaningless gesture condition could be attributed to the conflict between motor traces activated by observing meaningless movements and the processing of action verbs. Therefore, the use of gestures may have a negative impact during learning, particularly when instructors perform meaningless gestures. In addition, results revealed no differences between these two conditions (incongruent and meaningless conditions) (except a small effect found in the “see” learning group included in Experiment 3). Specifically, participants found themselves in a dual-task situation where they had to process both the gesture and the corresponding L1 word simultaneously as they learned the FL word. This dual coding requirement incurred a cost, which is consistent with findings from other studies that have highlighted the difficulty of encoding the meaning of a FL message when learners are simultaneously engaged in a concurrent task [160–162]. These results might not totally align with any of the theories predictions considering incongruent and meaningless gestures although the motor-imagery theory would explain in part the pattern of results.

Furthermore, our studies revealed differences between the learning of nouns and verbs. Typically, nouns exhibited a higher learning rate than verbs, primarily due to the greater semantic content of nouns [135,139,163,164]. We confirmed this pattern of results in the current study. Specifically, in the forward translation task, we observed that the recall of verbs was lower compared to the recall of nouns in the absence of gestures (no gesture condition). Furthermore, when we examined the impact of gestures, we found similar facilitation and interference effects in both the learning of FL nouns and verbs. However, upon closer examination, an interesting observation emerged: the inherent difficulty associated with the learning of verbs seemed to diminish when congruent gestures were incorporated into the training. In essence, the use of gestures in the acquisition of FL vocabulary appeared to alleviate the intrinsic challenges typically associated with learning verbs. This finding aligns again with the motor-imagery account, as the distinctions between nouns and verbs disappeared when the gestures conveyed the same meaning as the FL words to be learned. Interestingly, there is another hypothesis regarding the role of gestures on FL learning that can explain the positive effect associated with learning verbs accompanied by gestures. In fact, the

gestures for conceptualization hypothesis [165] points out that gestures have their origins in practical actions that encompass bodily movements and motor-related content. Importantly, the meaning of verbs inherently involves motoric information. Therefore, gestures may directly engage in the simulation of the meaning of verbs, which, in turn, would facilitate the learning of this specific category of words. This hypothesis aligns well with the effects observed in our study, highlighting the unique role of gestures in the acquisition of verbs.

The translation task participants performed at the end of each learning session gave us more information about the development of lexical and semantic links through the acquisition process. To better understand these effects, our results can be accommodated within the RHM hypothesis [16]. In short, this model proposes the existence of a network of interconnected elements, including L1 words, FL words, and a shared semantic system. However, the strength of these connections varies depending on the stage of FL vocabulary acquisition. In the early stages, the links between the semantic system and FL words are relatively weak, leading FL learners to predominantly rely on a lexical route for processing from FL to L1. As proficiency in the FL language increases, the connections between FL words and the semantic system become more robust, while the weight of the lexical route diminishes. This model has garnered substantial support from previous research. For instance, the model finds support in studies involving unbalanced bilinguals, where an asymmetry is observed in translation tasks. These individuals tend to exhibit faster performance in backward translation (which involves the lexical route from FL to L1) compared to forward translation (which relies on the semantic route from L1 to FL). This phenomenon has been documented in research by Kroll and Stewart [16] and contributes to our understanding of how bilinguals process and translate words between two languages depending on their FL proficiency levels. In our studies, the learning procedure including congruent gestures proved to foster the semantic route of processing. Previous research has demonstrated that gestures enhance the encoding of words to be learned by incorporating sensorimotor networks and procedural memory into the semantic/declarative memory associated with the word meanings [88]. Consequently, gestures play a role in enriching the semantic processing of FL words. The findings from our current studies provide evidence suggesting that the use of gestures is associated with semantic connections in FL learning. As previously mentioned, gestures effectively relieve the difficulty typically linked to the learning of verbs. This effect was particularly noticeable in forward translation, a task that relies more heavily on semantic mediation compared to backward translation (as discussed in [166]). Furthermore, our exploration of the characteristics of nouns and verbs in our study material revealed a notable difference in concreteness. Nouns were found to have a higher concreteness value compared to verbs. Concreteness is a variable known to influence FL vocabulary learning, with concrete words activating the semantic system more strongly than abstract words. This difference in activation makes concrete words more readily acquired by FL learners [167].

At first glance, it may appear that we reach a suitable learning strategy, but as we delve deeper, new questions come to light. What can happen if we remove the physical motor component of the equation? In our third experiment, our primary focus was to directly compare the effects of self-performed iconic gestures with the mere observation of gestures when acquiring vocabulary in a FL. Participants were randomly included in two groups. The “do” learning group was instructed to execute the gestures associated with the FL words (as in Experiments 1 and 2), while the “see” learning group solely observed the gestures without performing any physical movement. As experimenters, we took the risk of giving participants in the “see” learning group the instruction to imagine themselves performing the gestures. We did it to ensure that participants in the “see” condition did not exclusively focus on learning FL words while neglecting the processing of gestures. The study's results unveiled a higher recall of FL words in the “do” learning group compared to the “see” learning group (12% advantage). However, the difference between training types was significant in the analyses conducted at the item level but not in the participant-level analyses. In any case, participants in the “do” learning group retrieved FL words significantly faster than those in the “see” learning group. Consequently, the training approach centered on self-generated gestures somewhat facilitated the retrieval of new vocabulary in a FL. It is plausible that requesting

participants in the “see” group to imagine the gestures might have reduced the potential for a more consistent learning group effect. Nevertheless, previous studies comparing direct manipulation of objects with an imagined manipulation condition have shown superior outcomes when participants physically engage in the motor activity [46].

Regarding verb processing, Hauk and colleagues [130] found that passive reading of action verbs (e.g., to lick, pick, or kick) differentially activated brain areas along the associated motor strip, overlapping with areas activated during actual tongue, finger, and foot movements, respectively. Concerning the observation of movements, Buccino and colleagues [168] reported that observing both object and non-object-related actions led to somatotopically organized activation of the premotor cortex, similar to the classical motor cortex homunculus. Thus, the results from the “see” learning group suggest that processing 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 a FL [169]. However, the interference effects observed in the “see” learning group with meaningless gestures compared to the no-gesture condition partly diminished when participants performed the gestures while learning, as in the “do” learning group. The reduction of this interference effect in the “do” learning group was evident in the forward translation task but not in the backward translation task. We may wonder about the cognitive mechanism responsible for minimizing the negative impact of gestures that do not align in meaning with the words, especially in the “do” learning group. The advantages associated with gesture production during learning have been documented in previous studies. For example, Cook and Goldin-Meadow [170] found that children showed benefits in problem-solving when self-performing gestures. They argued that the facilitative role attributed to acting while learning is due to the reduction of working memory load. In the same line, Experiment 3 may show that the performance of gestures while learning could reduce cognitive effort in working memory, thereby attenuating the conflict arising from 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 widely established that the capacity to resolve conflict situations strongly depends on the availability of resources in working memory, a phenomenon observed in bilingual populations as well [171].

As previously mentioned, in our studies, the recall of FL words was lower, and response latencies were higher for L1–FL translation than for FL–L1 translation in the no-gesture condition. This outcome underscores the greater difficulty associated with forward translation compared to backward translation, even when gestures were not used during the learning of FL words. Furthermore, within the “do” learning group, the reduction of interference effects was more noticeable in the forward translation task (L1–FL translation) compared to the backward translation task (FL–L1) (Figure 8). Specifically, interference effects were observed in the meaningless gesture condition when participants performed the backward translation task. However, in the forward translation task, the recall percentages were similar across the incongruent gesture condition, the meaningless gesture condition, and the no-gesture condition. Again, the key distinction between the forward and backward translation tasks lies in their difficulty levels. As previously noted, the forward translation task, compared to the backward translation task, requires more extensive semantic processing, increasing the cognitive load, especially in the early stages of foreign language learning [16]. Consequently, the impact of gestures on learning and information retrieval appears to be contingent on task difficulty. For instance, Marstaller and Burianová [49] observed that a letter memorization task was more challenging and led to lower recall in participants with low working memory capacity. Importantly, the use of gestures facilitated letter recall compared to a condition without gestures, particularly for participants facing a more challenging task (those with low working memory capacity). In our study, performing gestures during learning mitigated interference associated with processing incongruent and meaningless gestures to a greater extent when task demands were high (forward translation task). Therefore, the facilitative effect of gesture performance during learning appears to be more prominent when the retrieval task demands greater cognitive effort.

At this juncture, one might wonder why the interaction between translation direction and learning condition observed in the “do” learning group (i.e., the attenuation of interference during recall of FL words in the forward translation but not in the backward translation) did not manifest in the “see” learning group. In the “see” group of participants, the recall of FL words was not influenced by translation direction, nor did this variable interact with the learning condition. In fact, it could be argued that the “see” learning group would exhibit an advantage in the translation task because they did not exhibit the typical difficulty observed in forward vs. backward translation [16]. We acknowledge that we do not possess a definitive explanation for this phenomenon. However, it is plausible that the way we presented the verbal material in our study (i.e., the Spanish word and its Vimmi translation, L1–FL word pairs) may have led participants to primarily employ a lexical coding strategy, regardless of the translation direction. In contrast, the performance of gestures in the “do” learning group would have favored semantic processing of the material [89,102]. This enhanced semantic processing in the “do” group may have been more pronounced in the forward vs. backward translation task, as the L1–FL translation requires a greater degree of conceptual information retrieval than FL–L1 translation [16,142]. Nevertheless, we acknowledge the need for caution in interpreting this explanation.

On another note, the pattern of results found in Experiment 2 can be compared with the learning outcomes of the “do” learning group in Experiment 3, as both experiments employed the same design. In general, both studies revealed similar main effects and interactions between variables. However, some differences were noted in the final outcomes. Specifically, in Experiment 2, the comparison between the congruent and no-gesture conditions in the backward translation direction revealed an advantage associated with performing congruent gestures, evident in both response times and accuracy. In contrast, in the “do” learning group, no differences were observed between the congruent and no-gesture conditions in the backward translation direction. While tentative, one possible explanation for these differences may be variations in the overall rate of FL learning across the studies. Participants in Experiment 2 exhibited poorer performance across the three learning sessions (i.e., slower response times and lower accuracy) compared to participants in Experiment 3. Therefore, participants in Experiment 2 may have had more room to benefit from performing congruent gestures during FL learning.

Future research endeavors could delve deeper into the relationship between incongruent and meaningless gestures, as both conditions seem to exhibit similar pattern of processing when associated with new words to be learned. Additionally, it would be intriguing to design a long-term assessment program to determine how the recall of new words evolve over time when learners acquire new FL words across conditions. Moreover, the effect of gesture imitation and gesture observation strategies on the learning of nouns, which are typically the first words children learn in language development, could be evaluated. Finally, approaches rooted in conceptual material processing, such as using images alongside words, as seen in Tellier [120], have also shown promise in supporting FL learning. A future direction for exploration could involve assessing the potential cumulative benefits of combining these two instructional aids (gestures and images) on FL vocabulary acquisition. Finally, an additional crucial factor to consider in future studies is the presence of individual differences among learners, which can influence the effectiveness of specific strategies for FL acquisition [172–174].

5. Conclusions

As mentioned in the introduction section, the theories explaining the potential role gestures have on FL vocabulary learning are not mutually exclusive. In fact, these accounts highlight different aspects of gestures' effect on FL learning. Accompanying a word with a gesture may enhance self-involvement (thereby increasing attention to FL learning), establish a meaningful motor trace, and/or evoke a semantic visual image that is integrated with the word's meaning. In our studies, we have found that the use of gestures congruent in meaning with the target words greatly facilitates vocabulary acquisition in a FL. This advantage extends to the realm of verb learning, where congruent gestures alleviate the difficulties encountered compared to noun learning. In addition, this

effect is more pronounced when learners actively engage in gesture production. Moreover, gesture production seems to counteract potential adverse effects linked to unrelated gestures. Hence, when deciding between various FL learning methodologies incorporating gestures, our recommendation is to implement a training protocol that encourages participants to generate gestures aligned with the words they are learning. Finally, we would like to point out that learning a new language is not as easy as rubbing the magic lamp but it can be easier with the appropriate learning methods.

Funding: This research was funded by Spanish Ministry of Economy and Competitiveness Grant PSI2016-75250-P to P.M. Preparation of this paper was supported by a grant awarded to P.M. by the Spanish Ministry of Science and Innovation (PID2019-111359GB-I00/AEI/10.13039/501100011033). All procedures performed in this study involving human participants were in accordance with the ethical standards of the research ethical committee at the University of Granada (number issued by the ethical committee: 86/CEIH/2015) and with the 1964 Helsinki declaration and its later amendments.

Data Availability Statement: Data will be accessible along the revision process.

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

References

1. Breiner-sanders, K.; Richter, J.; Chi, T.R. Total language immersion programs: Outcomes and assessments—the Middlebury experience. In *Annual Meeting of the American Council on the Teaching of Foreign Languages (ACTFL)*, Dallas, TX. **1999**
2. Coleman, J.A. The current state of knowledge concerning student residence abroad. *The year abroad: Preparation, monitoring, evaluation*, **1995**, pp. 17-42.
3. Coleman, J.A. Language learning and study abroad: The European perspective. *Frontiers: The interdisciplinary journal of study abroad*, **1998**, pp. 167-203. <https://doi.org/10.36366/frontiers.v4i1.67>
4. Dewey, D. The effects of study context and environment on the acquisition of reading by students of Japanese as a second language during study-abroad and intensive domestic immersion. *Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA*, **2002**.
5. Freed, B.F.; Segalowitz, N.; Dewey, D.P. Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in second language acquisition*, **2004**, 26, pp. 275-301. <https://doi.org/10.1017/S0272263104262064>
6. Genesee, F. Second language learning in school settings: Lessons from immersion. In *Bilingualism, multiculturalism, and second language learning*. Psychology Press, **2014**, pp. 183-201.
7. Marian, V.; Shook, A.; Schroeder, S.R. Bilingual two-way immersion programs benefit academic achievement. *Bilingual research journal*, **2013**, 36, pp. 167-186. <https://doi.org/10.1080/15235882.2013.818075>
8. Comesaña, M.; Perea, M.; Piñero, A.; Fraga, I. Vocabulary teaching strategies and conceptual representations of words in L2 in children: Evidence with novice learners. *Journal of Experimental Child Psychology*, **2009**, 104, pp. 22-33. <https://doi.org/10.1016/j.jecp.2008.10.004>
9. Comesaña, M.; Soares, A.P.; Sánchez-Casas, R.; Lima, C. 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*, **2012**, 103, pp. 378-392. <https://doi.org/10.1111/j.2044-8295.2011.02080.x>
10. García-Gámez, A.B.; Macizo, P. The way in which foreign words are learned determines their use in isolation and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **2020**, 46, pp. 364–379. <https://doi.org/10.1037/xlm0000721>
11. García-Gámez, A.B.; Macizo, P. Seeing or acting? The effect of performing gestures on foreign language vocabulary learning. *Language Teaching Research*, **2021**. <https://doi.org/10.1177/13621688211102>
12. Atkinson, R.C.; Raugh, M.R. An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory*, **1975**, 1, pp.126–133. <https://doi.org/10.1037/0278-7393.1.2.126>
13. Raugh, M.R.; Atkinson, R.C. A mnemonic method for learning a second-language vocabulary. *Journal of Educational Psychology*, **1975**, 67, pp. 1-16. <https://doi.org/10.1037/h0078665>
14. Atkinson, R. C. Mnemotechnics in second-language learning. *American Psychologist*, **1975**, 30, pp.821–828. <https://doi.org/10.1037/h0077029>

15. García-Gámez, A.B.; Macizo, P. Lexical and semantic training to acquire words in a foreign language: An electrophysiological study. *Bilingualism: Language and Cognition*, **2022**, *25*, pp. 768-785. <https://doi.org/10.1017/S1366728921000456>
16. Kroll, J.F.; Stewart, E. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of memory and language*, **1994**, *33*, pp. 149-174. <https://doi.org/10.1006/jmla.1994.1008>
17. Altarriba, J.; Mathis, K.M. Conceptual and lexical development in second language acquisition. *Journal of memory and language*, **1997**, *36*, pp. 550-568. <https://doi.org/10.1006/jmla.1997.2493>
18. Morett, L.M. In hand and in mind: Effects of gesture production and viewing on second language word learning. *Applied Psycholinguistics*, **2018**, *39*, pp. 355–381. <https://doi.org/10.1017/S0142716417000388>
19. Poarch, G.J.; Van Hell, J.G.; Kroll, J.F. Accessing word meaning in beginning second language learners: Lexical or conceptual mediation? *Bilingualism: Language and Cognition*, **2015**, *18*, pp. 357-371. <https://doi.org/10.1017/S1366728914000558>
20. Barcroft, J. Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, **2002**, *5*, 2pp. 323-363. <https://doi.org/10.1111/0023-8333.00186>
21. de Groot, A.M.B; Poot, R. Word translation at three levels of proficiency in a second language: The ubiquitous involvement of conceptual memory. *Language learning*, **1997**, *47*, pp. 215-264. <https://doi.org/10.1111/0023-8333.71997007>
22. Finkbeiner, M.; Nicol, J. Semantic category effects in second language word learning. *Applied psycholinguistics*, **2003**, *24*, pp. 369-383. <https://doi.org/10.1017/S0142716403000195>
23. Kroll, J.F.; Michael, E.; Sankaranarayanan, A. A model of bilingual representation and its implications for second language acquisition. In *Foreign Language Learning*, Healy, A. F., Ed; Erlbaum: Mahwah NJ, 2013. pp. 365-395.
24. Wimer, C.C.; Lambert, W.E. The differential effects of word and object stimuli on the learning of paired associates. *Journal of Experimental Psychology*, **1959**, *57*, pp. 31-36. <https://doi.org/10.1037/h0046822>
25. Khezrlou, S.; Ellis, R.; Sadeghi, K. Effects of computer-assisted glosses on EFL learners' vocabulary acquisition and reading comprehension in three learning conditions. *System*, **2017**, *65*, pp. 104-116. <https://doi.org/10.1016/j.system.2017.01.009>
26. Liu, Y.; Jang, B.G.; Roy-Campbell, Z. Optimum input mode in the modality and redundancy principles for university ESL students' multimedia learning. *Computers & Education*, **2018**, *127*, pp. 190-200. <https://doi.org/10.1016/j.compedu.2018.08.025>
27. Montero Pérez, M.; Peters, E.; Clarebout, G.; Desmet, P. Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology*, **2014**, *18*, pp. 118-141.
28. Sweller, J. Cognitive load during problem solving: Effects on learning. *Cognitive science*, **1988**, vol. 12, pp. 257-285. [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7)
29. Aldera, A.S.; Mohsen, M.A. Annotations in captioned animation: Effects on vocabulary learning and listening skills. *Computers & Education*, **2013**, *68*, pp. 60-75. <https://doi.org/10.1016/j.compedu.2013.04.018>
30. Peters, E. The effect of imagery and on-screen text on foreign language vocabulary learning from audiovisual input. *Tesol Quarterly*, **2019**, *53*, pp. 1008-1032. <https://doi.org/10.1002/tesq.531>
31. Teng, F. Incidental vocabulary learning for primary school students: the effects of L2 caption type and word exposure frequency. *The Australian Educational Researcher*, **2019**, *46*, pp. 113-136. <https://doi.org/10.1007/s13384-018-0279-6>
32. Lotto, L.; de Groot, A.M.B. Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language learning*, **1998**, *48*, pp. 31-69. <https://doi.org/10.1111/1467-9922.00032>
33. Ellis, N.C.; Beaton, A. Psycholinguistic determinants of foreign language vocabulary learning. *Language learning*, **1993**, *43*, pp. 559-617. <https://doi.org/10.1111/j.1467-1770.1993.tb00627.x>
34. Wang, A.Y.; Thomas, M.H. Effects of keyword on long-term retention: Help or hindrance? *Journal of Educational Psychology*, **1995**, *87*, pp. 468-475. <https://doi.org/10.1037/0022-0663.87.3.468>
35. Hirata, Y.; Kelly, S.D.; Huang, J.; Manansala, M. Effects of hand gestures on auditory learning of second-language vowel length contrasts. *Journal of Speech, Language, and Hearing Research*, **2014**, *57*, pp. 2090-2101. https://doi.org/10.1044/2014_JSLHR-S-14-0049
36. Kelly, S.D.; Özyürek, A.; Maris, E. Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological science*, **2010**, *21*, pp. 260-267. <https://doi.org/10.1177/0956797609357327>

37. Macedonia, M. Bringing back the body into the mind: gestures enhance word learning in foreign language. *Frontiers in psychology*, **2014**, 5, pp. 1467. <https://doi.org/10.3389/fpsyg.2014.01467>
38. McNeill, D. *Hand and mind: What gestures reveal about thought*. University of Chicago press: Chicago, 1992.
39. McNeill, D. *Gesture and thought*. Chicago: University of Chicago Press: Chicago, 2005
40. McNeill, D., Levy, E.; Duncan, S. Gesture in discourse. In Tannen, D. Hamilton, H.; Schiffrin D. Eds.; *Handbook of discourse analysis*, Wiley: London, 2015. pp. 262–319.
41. Paivio, A. Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, **1991**, 45, pp. 255-287. <https://doi.org/10.1037/h0084295>
42. Macuch Silva, V.; Holler, J.; Ozyurek, A.; Roberts, S.G. Multimodality and the origin of a novel communication system in face-to-face interaction. *Royal Society open science*, **2020**, 7. <https://doi.org/10.1098/rsos.182056>
43. Ortega, G.; Özyürek, A. Systematic mappings between semantic categories and types of iconic representations in the manual modality: A normed database of silent gesture. *Behavior Research Methods*, **2020**, 52, pp. 51-67. <https://doi.org/10.3758/s13428-019-01204-6>
44. Özyürek, A. Hearing and seeing meaning in speech and gesture: Insights from brain and behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **2014**, 369. <https://doi.org/10.1098/rstb.2013.0296>
45. Holle, H.; Gunter, T.C. The role of iconic gestures in speech disambiguation: ERP evidence. *Journal of cognitive neuroscience*, **2007**, 19, pp. 1175-1192. <https://doi.org/10.1162/jocn.2007.19.7.1175>
46. Glenberg, A.M.; Gutierrez, T.; Levin, J.R.; Japuntich, S.; Kaschak, M.P. Activity and imagined activity can enhance young children's reading comprehension. *Journal of educational psychology*, **2004**, pp. 424-436. <https://doi.org/10.1037/0022-0663.96.3.424>
47. Koriat, A.; Pearlman-Avnion, S. Memory organization of action events and its relationship to memory performance. *Journal of Experimental Psychology: General*, **2003**, 132, pp. 435-454. <https://doi.org/10.1037/0096-3445.132.3.435>
48. Glenberg, A.M.; Sato, M.; Cattaneo, L. Use-induced motor plasticity affects the processing of abstract and concrete language. *Current Biology*, **2008**, 18. <https://doi.org/10.1016/j.cub.2008.02.036>
49. Marstaller, L.; Burianová, H. The multisensory perception of co-speech gestures—A review and meta-analysis of neuroimaging studies. *Journal of Neurolinguistics*, **2014**, 30, pp. 69-77. <https://doi.org/10.1016/j.jneuroling.2014.04.003>
50. De Ruiter, J. The production of gesture and speech. In D. McNeill Ed.; *Language and gesture*. Cambridge University Press: Cambridge, 2000. pp. 284–311.
51. Kita, S.; Özyürek, A. 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*, **2003**, 48, pp. 16-32. [https://doi.org/10.1016/S0749-596X\(02\)00505-3](https://doi.org/10.1016/S0749-596X(02)00505-3)
52. Hostetter, A.B.; Alibali, M.W. Visible embodiment: Gestures as simulated action. *Psychonomic bulletin & review*, **2008**, 15, pp. 495-514. <https://doi.org/10.3758/PBR.15.3.495>
53. Krauss, R.M.; Chen, Y.; Gottesman, R.F.; McNeill, D. *Language and gesture*, 2000.
54. Goldin-Meadow, S. Beyond words: The importance of gesture to researchers and learners. *Child development*, **2000**, 71, pp. 231-239. <https://doi.org/10.1111/1467-8624.00138>
55. Goldin-Meadow, S. *Hearing gesture: How our hands help us think*. Harvard University Press: Cambridge, MA. 2003
56. Butcher, C.; Goldin-Meadow, S. Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill Ed.; *Language and gesture*, Cambridge University Press: Cambridge, 2000. pp. 235-258
57. Özçalışkan, Ş.; Gentner, D.; Goldin-Meadow, S. Do iconic gestures pave the way for children's early verbs? *Applied psycholinguistics*, **2014**, 35, pp. 1143-1162. <https://doi.org/10.1017/S0142716412000720>
58. Alibali, M.W.; Flevares, L.M.; Goldin-Meadow, S. Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of educational psychology*, 1997, 89, pp. 183-193. <https://doi.org/10.1037/0022-0663.89.1.183>
59. Cassell, J.; McNeill, D.; McCullough, K.E. Speech-gesture mismatches: Evidence for one underlying representation of linguistic and nonlinguistic information. *Pragmatics & cognition*, **1999**, 7, pp. 1-34. <https://doi.org/10.1075/pc.7.1.03cas>

60. Goldin-Meadow, S.; Wein, D.; Chang, C. Assessing knowledge through gesture: Using children's hands to read their minds. *Cognition and Instruction*, **1992**, *9*, pp. 201-219. https://doi.org/10.1207/s1532690xci0903_2
61. Holler, J.; Shovelton, H.; Beattie, G. Do iconic hand gestures really contribute to the communication of semantic information in a face-to-face context? *Journal of Nonverbal Behavior*, **2009**, *33*, pp. 73-88. <https://doi.org/10.1007/s10919-008-0063-9>
62. Singer, M.A.; Goldin-Meadow, S. Children learn when their teacher's gestures and speech differ. *Psychological science*, **2005**, *16*, pp. 85-89. <https://doi.org/10.1111/j.0956-7976.2005.00786.x>
63. De Grauwe, S.; Willems, R.M.; Rueschemeyer, S.A.; Lemhöfer, K.; Schriefers, H. Embodied language in first-and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia*, **2014**, *56*, pp. 334-349. <https://doi.org/10.1016/j.neuropsychologia.2014.02.003>
64. Gullberg, M. Gestures and second language acquisition. In Müller, C., Cienki, A., E. Fricke., et al. Eds.; *Body language communication: An international handbook on multimodality in human interaction*, De Gruyter Mouton, 2014. pp. 1868-1875
65. Kelly, S.D.; Creigh, P.; Bartolotti, J. Integrating speech and iconic gestures in a Stroop-like task: evidence for automatic processing. *Journal of cognitive neuroscience*, **2010**, *22*, pp. 683-694. <https://doi.org/10.1162/jocn.2009.21254>
66. Macedonia, M.; Knösche, T.R. Body in mind: How gestures empower foreign language learning. *Mind, Brain, and Education*, **2011**, *5*, pp. 196-211. <https://doi.org/10.1111/j.1751-228X.2011.01129.x>
67. Macedonia, M.; Kriegstein, K. Gestures enhance foreign language learning. *Biolinguistic*, **2012**, *6*, pp. 393-416. <https://hdl.handle.net/20.500.12528/1489>
68. Morett, L.M. 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*, **2014**, *98*, pp. 834-853. <https://doi.org/10.1111/modl.12125>
69. So, W.C.; Sim Chen-Hui, C.; Low Wei-Shan, J. Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall? *Language and Cognitive Processes*, **2012**, *27*, pp. 665-681. <https://doi.org/10.1080/01690965.2011.573220>
70. McCafferty, S.; Stam, G. Gesture, second language acquisition and classroom research. Taylor & Francis, 2008.
71. Asher, J.J.; Price, B.S. The learning strategy of the total physical response: Some age differences. *Child development*, **1967**, 1219-1227. <https://doi.org/10.2307/1127119>
72. Carels, P.E. Pantomime in the foreign language classroom. *Foreign Language Annals*, **1981**, *14*, pp. 407-411. <https://doi.org/10.1111/j.1944-9720.1981.tb01660.x>
73. Krashen, S.D.; Terrell, T.D. The natural approach: Language acquisition in the classroom. Alemany Press: San Francisco, 1983.
74. Macedonia, M.; Bergmann, K.; Roithmayr, F. Imitation of a pedagogical agent's gestures enhances memory for words in second language. *Science Journal of Education*, **2014**, *2*, pp. 162-169. <https://hdl.handle.net/11858/00-001M-0000-002B-563A-8>
75. Kelly, S.D.; Hirata, Y.; Manansala, M.; Huang, J. Exploring the role of hand gestures in learning novel phoneme contrasts and vocabulary in a second language. *Frontiers in Psychology*, **2014**, *5*, pp. 673. <https://doi.org/10.3389/fpsyg.2014.00673>
76. Straube, B.; Green, A.; Weis, S.; Kircher, T. A supramodal neural network for speech and gesture semantics: an fMRI study. *PloS one*, **2012**, *7*. <https://doi.org/10.1371/journal.pone.0051207>
77. Yang, J.; Andric, M.; Mathew, M.M. The neural basis of hand gesture comprehension: a meta-analysis of functional magnetic resonance imaging studies. *Neuroscience & Biobehavioral Reviews*, **2015**, *57*, pp. 88-104. <https://doi.org/10.1016/j.neubiorev.2015.08.006>
78. Goldin-Meadow, S.; Alibali, M.W. Gesture's role in speaking, learning, and creating language. *Annual review of psychology*, **2013**, *64*, pp. 257-283. <https://doi.org/10.1146/annurev-psych-113011-143802>
79. Helstrup, T. One, two, or three memories? A problem-solving approach to memory for performed acts. *Acta Psychologica*, **1987**, *66*, pp. 37-68. [https://doi.org/10.1016/0001-6918\(87\)90017-5](https://doi.org/10.1016/0001-6918(87)90017-5)
80. Bäckman, L.; Nilsson, L.G.; Chalom. New evidence on the nature of the encoding of action events. *Memory & Cognition*, **1986**, *14*, pp. 339-346. <https://doi.org/10.3758/BF03202512>
81. Kormi-Nouri, R.; Nilsson, L. G. The motor component is not crucial! In H. D. Zimmer, R.L. Cohen, M. J. Gynn, J. Engelkamp, R. Kormi-Nouri, & M. A. Foley Eds.; *Memory for action: A distinct form of episodic memory?* Oxford University Press: New York, 2001. pp. 97-111.

82. Knopf, M. Gedächtnis für Handlungen: Funktionsweise und Entwicklung (Unpublished doctoral dissertation, Ruprecht-Karls-Universität, Heidelberg) **1992**.
83. Craik, F.I.M.; Tulving, E. Depth of processing and the retention of words in episodic memory. *Journal of experimental Psychology: general*, **1975**, *104*, pp. 268-294. <https://doi.org/10.1037/0096-3445.104.3.268>
84. Muzzio, I.A.; Levita, L.; Kulkarni, J.; Monaco, J.; Kentros, C.; Stead, M.; ... Kandel, E.R. Attention enhances the retrieval and stability of visuospatial and olfactory representations in the dorsal hippocampus. *PLoS biology*, **2009**, *7*. <https://doi.org/10.1371/journal.pbio.1000140>
85. Engelkamp, J.; Zimmer, H.D. Motor programme information as a separable memory unit. *Psychological Research*, **1984**, *46*, pp. 283-299. <https://doi.org/10.1007/BF00308889>
86. Engelkamp, J.; Zimmer, H.D. Motor programs and their relation to semantic memory. *German Journal of psychology*, **1985**, *9*, pp. 239-254.
87. Vukovic, N.; Feurra, M.; Shpektor, A.; Myachykov, A.; Shtyrov, Y. Primary motor cortex functionally contributes to language comprehension: An online rTMS study. *Neuropsychologia*, **2017**, *96*, pp. 222-229. <https://doi.org/10.1016/j.neuropsychologia.2017.01.025>
88. Macedonia, M.; Mueller, K. Exploring the neural representation of novel words learned through enactment in a word recognition task. *Frontiers in psychology*, **2016**, *7*, pp. 953. <https://doi.org/10.3389/fpsyg.2016.00953>
89. Denis, M.; Engelkamp, J.; Mohr, G. Memory of imagined actions: Imagining oneself or another person. *Psychological Research*, **1991**, *53*, pp. 246-250. <https://doi.org/10.1007/BF00941394>
90. Macedonia, M.; Müller, K.; Friederici, A.D. The impact of iconic gestures on foreign language word learning and its neural substrate. *Human brain mapping*, **2011**, *32*, pp. 982-998. <https://doi.org/10.1002/hbm.21084>
91. Takashima, A.; Bakker, I.; Van Hell, J.G.; Janzen, G.; McQueen, J.M. Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. *NeuroImage*, **2014**, *84*, pp. 265-278. <https://doi.org/10.1016/j.neuroimage.2013.08.023>
92. Cook, S.W.; Yip, T.K.; Goldin-Meadow, S. Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and cognitive processes*, **2012**, *27*, pp. 594-610. <https://doi.org/10.1080/01690965.2011.567074>
93. Yap, D.F.; So, W.C.; Melvin Yap, J.M.; Tan, Y.Q.; Teoh, R.L.S. Iconic gestures prime words. *Cognitive science*, **2011**, *35*, pp. 171-183. <https://doi.org/10.1111/j.1551-6709.2010.01141.x>
94. Feyereisen, P. Further investigation on the mnemonic effect of gestures: Their meaning matters. *European Journal of Cognitive Psychology*, **2006**, *18*, pp. 185-205. <https://doi.org/10.1080/09541440540000158>
95. Kelly, S.D.; McDevitt, T.; Esch, M. Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and cognitive processes*, **2009**, *24*, pp. 313-334. <https://doi.org/10.1080/01690960802365567>
96. Macedonia, M.; Klimesch, W. Long-term effects of gestures on memory for foreign language words trained in the classroom. *Mind, Brain, and Education*, **2014**, *8*, pp. 74-88. <https://doi.org/10.1111/mbe.12047>
97. Asher, J.J. The Learning Strategy of the Total Physical Response: A Review. *The modern language journal*, **1966**, *50*, pp. 79-84.
98. Aleven, V.; Koedinger, K.R. An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive science*, **2002**, *26*, pp. 147-179. https://doi.org/10.1207/s15516709cog2602_1
99. Bessen, James. Learning by doing: the real connection between innovation, wages, and wealth. Yale University Press: Yale, 2015.
100. Nakatsukasa, K. Gesture-enhanced recasts have limited effects: A case of the regular past tense. *Language Teaching Research*, **2021**, *25*, pp. 587-612. <https://doi.org/10.1177/1362168819870283>
101. Allen, L.Q. The effects of emblematic gestures on the development and access of mental representations of French expressions. *The Modern Language Journal*, **1995**, *79*, pp. 521-529. <https://doi.org/10.1111/j.1540-4781.1995.tb05454.x>
102. Krönke, K.M.; Mueller, K.; Friederici, A.D.; Obrig, H. Learning by doing? The effect of gestures on implicit retrieval of newly acquired words. *Cortex*, **2013**, *49*, pp. 2553-2568. <https://doi.org/10.1016/j.cortex.2012.11.016>
103. Barbieri, F.; Buonocore, A.; Dalla Volta, R.; Gentilucci, M. How symbolic gestures and words interact with each other. *Brain and language*, **2009**, *110*, pp. 1-11. <https://doi.org/10.1016/j.bandl.2009.01.002>
104. Bernardis, P.; Gentilucci, M. Speech and gesture share the same communication system. *Neuropsychologia*, **2006**, *44*, pp. 178-190. <https://doi.org/10.1016/j.neuropsychologia.2005.05.007>

105. Bernardis, P.; Salillas, E.; Caramelli, N. Behavioural and neurophysiological evidence of semantic interaction between iconic gestures and words. *Cognitive neuropsychology*, **2008**, *25*, pp. 1114-1128. <https://doi.org/10.1080/02643290801921707>
106. Chieffi, S.; Secchi, C.; Gentilucci, M. Deictic word and gesture production: their interaction. *Behavioural brain research*, **2009**, *203*, pp. 200-206. <https://doi.org/10.1016/j.bbr.2009.05.003>
107. Kircher, T.; Straube, B.; Leube, D.; Weis, S.; Sachs, O.; Willmes, K.; ... Green, A. Neural interaction of speech and gesture: differential activations of metaphoric co-verbal gestures. *Neuropsychologia*, **2009**, *47*, pp. 169-179. <https://doi.org/10.1016/j.neuropsychologia.2008.08.009>
108. Kelly, S.; Healey, M.; Özyürek, A.; Holler, J. The processing of speech, gesture, and action during language comprehension. *Psychonomic bulletin & review*, **2015**, *22*, pp. 517-523. <https://doi.org/10.3758/s13423-014-0681-7>
109. Kutas, M.; Hillyard, S.A. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, **1980**, *207*, pp. 203-205. <https://doi.org/10.1126/science.7350657>
110. MacLeod, C.M. Half a century of research on the Stroop effect: an integrative review. *Psychological bulletin*, **1991**, *109*, pp. 163. <https://doi.org/10.1037/0033-2909.109.2.163>
111. Rizzolatti, G.; Craighero, L. The mirror-neuron system. *Annu. Rev. Neurosci.*, **2004**, *27*, pp. 169-192. <https://doi.org/10.1146/annurev.neuro.27.070203.144230>
112. Stefan, K.; Cohen, L.G.; Duque, J.; Mazzocchio, R.; Celnik, P.; Sawaki, L.; ... Classen, J. Formation of a motor memory by action observation. *Journal of Neuroscience*, **2005**, *25*, pp. 9339-9346. <https://doi.org/10.1523/JNEUROSCI.2282-05.2005>
113. Goldin-Meadow, S. The role of gesture in communication and thinking. *Trends in cognitive sciences*, **1999**, *3*, pp. 419-429. [https://doi.org/10.1016/S1364-6613\(99\)01397-2](https://doi.org/10.1016/S1364-6613(99)01397-2)
114. Goldin-Meadow, S.; Levine, S.C.; Zinchenko, E.; Yip, T.K.; Hemani, N.; Factor, L. Doing gesture promotes learning a mental transformation task better than seeing gesture. *Developmental science*, **2012**, *15*, pp. 876-884. <https://doi.org/10.1111/j.1467-7687.2012.01185.x>
115. Cutting, J.; Iacovides, I. Learning by doing: Intrinsic Integration directs attention to increase learning in games. *Proceedings of the ACM on Human-Computer Interaction*, **2022**, *6*, pp. 1-18. <https://doi.org/10.1145/3549503>
116. Cohen, R.L. On the generality of some memory laws. *Scandinavian Journal of Psychology*, **1981**, *22*, pp. 267-281. <https://doi.org/10.1111/j.1467-9450.1981.tb00402.x>
117. Engelkamp, J.; Zimmer, H.D. Zum Einfluß von Wahrnehmen und Tun auf das Behalten von Verb-Objekt-Phrasen [The influence of perception and performance on the recall of verb-object phrases]. *Sprache & Kognition*, **1983**, *2*, pp.117-127.
118. Ganis, G.; Keenan, J.P.; Kosslyn, S.M.; Pascual-Leone, A. Transcranial magnetic stimulation of primary motor cortex affects mental rotation. *Cerebral Cortex*, **2000**, *10*, pp. 175-180. <https://doi.org/10.1093/cercor/10.2.175>
119. Chu, M.; Kita, S. Spontaneous gestures during mental rotation tasks: insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General*, **2008**, *137*, pp. 706-723. <https://doi.org/10.1037/a0013157>
120. Tellier, M. The effect of gestures on second language memorisation by young children. *Gesture*, **2008**, *8*, pp. 219-235. <https://doi.org/10.1075/gest.8.2.06tel>
121. James, K.H.; Swain, S.N. Only self-generated actions create sensori-motor systems in the developing brain. *Developmental science*, **2011**, *14*, pp. 673-678. <https://doi.org/10.1111/j.1467-7687.2010.01011.x>
122. Engelkamp, J.; Zimmer, H.D.; Mohr, G.; Sellen, O. Memory of self-performed tasks: Self-performing during recognition. *Memory & Cognition*, **1994**, *22*, pp. 34-39. <https://doi.org/10.3758/BF03202759>
123. Cherdieu, M.; Palombi, O.; Gerber, S.; Troccaz, J.; Rochet-Capellan, A. Make gestures to learn: Reproducing gestures improves the learning of anatomical knowledge more than just seeing gestures. *Frontiers in psychology*, **2017**, *8*, pp. 1689. <https://doi.org/10.3389/fpsyg.2017.01689>
124. Baills, F.; Suarez-Gonzalez, N.; Gonzalez-Fuente, S.; Prieto, P. Observing and producing pitch gestures facilitates the learning of Mandarin Chinese tones and words. *Studies in Second Language Acquisition*, **2019**, *41*, pp. 33-58. <https://doi.org/10.1017/S0272263118000074>
125. Sweller, N.; Shinooka-Phelan, A.; Austin, E. The effects of observing and producing gestures on Japanese word learning. *Acta psychologica*, **2020**, *207*. <https://doi.org/10.1016/j.actpsy.2020.103079>

126. Hastings, P.M.; Lytinen, S.L. Objects, actions, nouns, and verbs. In Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society. Routledge, 2019. pp. 397-402.
127. Hadley, E.B.; Dickinson, D.K.; Hirsh-Pasek, K.; Golinkoff, R.M.; Nesbitt, K.T. Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly*, **2016**, *51*, pp. 181-198. <https://doi.org/10.1002/rrq.130>
128. Boulenger, V.; Hauk, O.; Pulvermüller, F. Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. *Cerebral cortex*, 2009, *19*, pp. 1905-1914. <https://doi.org/10.1093/cercor/bhn217>
129. Childers, J.B.; Tomasello, M. Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental psychology*, **2002**, *38*, pp. 967-978. <https://doi.org/10.1037/0012-1649.38.6.967>
130. Hauk, O.; Johnsrude, I.; Pulvermüller, F. Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, **2004**, *41*, pp. 301-307.
131. Golinkoff, R.M.; Chung, H.L.; Hirsh-Pasek, K.; Liu, J.; Bertenthal, B.I.; Brand, R.; ... Hennon, E. Young children can extend motion verbs to point-light displays. *Developmental psychology*, **2002**, *38*, pp. 604-614. <https://doi.org/10.1037/0012-1649.38.4.604>
132. Fernald, A.; Morikawa, H. Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child development*, **1993**, *64*, pp. 637-656. <https://doi.org/10.1111/j.1467-8624.1993.tb02933.x>
133. Goldfield, B.A. Noun bias in maternal speech to one-year-olds. *Journal of child language*, **1993**, *20*, pp. 85-99. <https://doi.org/10.1017/S0305000900009132>
134. Tardif, T.; Shatz, M.; Naigles, L. Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language*, **1997**, *24*, pp. 535-565. <https://doi.org/10.1017/S030500099700319X>
135. Gentner, D.; Boroditsky, L. Individuation, relativity, and early word learning. In M. Bowerman & S. C. Levinson Eds.; *Language acquisition and conceptual development*, Cambridge University Press: Cambridge, 2001. pp. 215-256.
136. Choi, S.; Gopnik, A. Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of child language*, **1995**, *22*, pp. 497-529. <https://doi.org/10.1017/S0305000900009934>
137. García-Gámez, A.B.; Macizo, P. Learning nouns and verbs in a foreign language: The role of gestures. *Applied Psycholinguistics*, **2019**, *40*, pp. 473-507. <https://doi.org/10.1017/S0142716418000656>
138. Worthen, J. B. 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*, Oxford University Press: New York, 2006. pp. 133-156.
139. Gentner, D. Some interesting differences between verbs and nouns. *Cognition and brain theory*, **1981**, *4*, pp. 161-178.
140. Bauer, P.J.; Pathman, T.; Inman, C.; Campanella, C.; Hamann, S. Neural correlates of autobiographical memory retrieval in children and adults. *Memory*, **2017**, *25*, pp. 450-466. <https://doi.org/10.1080/09658211.2016.1186699>
141. Kroll, J.F.; Michael, E.; Tokowicz, N.; Dufour, R. The development of lexical fluency in a second language. *Second language research*, **2002**, *18*, pp. 137-171. <https://doi.org/10.1191/0267658302sr2010a>
142. Kroll, J.F.; Tokowicz, N. The development of conceptual representation for words in a second language. In J. Nicol Ed.; *One mind, two languages: Bilingual language processing*, Blackwell Publishers: Malden, 2001. pp. 49-71.
143. Hammarberg, B. Roles of L1 and L2 in L3 production and acquisition. In C. Baker & N.H. Hornberger Eds.; *Cross-linguistic influence in third language acquisition*, Multilingual Matters LTD: Sydney, 2001. pp. 21-41
144. Serfaty, J.; Serrano, R. Lag effects in grammar learning: A desirable difficulties perspective. *Applied Psycholinguistics*, **2022**, *43*, pp. 513-550. <https://doi.org/10.1017/S0142716421000631>
145. Tremblay, Annie. On the second language acquisition of Spanish reflexive passives and reflexive impersonals by French-and English-speaking adults. *Second Language Research*, **2006**, *22*, pp. 30-63. <https://doi.org/10.1191/0267658306sr2600a>
146. Bartolotti, J.; Marian, V. Language learning and control in monolinguals and bilinguals. *Cognitive science*, **2012**, *36*, pp. 1129-1147. <https://doi.org/10.1111/j.1551-6709.2012.01243.x>
147. Bartolotti, J.; Marian, V. Learning and processing of orthography-to-phonology mappings in a third language. *International journal of multilingualism*, **2019**, *16*, pp. 377-397. <https://doi.org/10.1080/14790718.2017.1423073>

148. Grey, S. What can artificial languages reveal about morphosyntactic processing in bilinguals? *Bilingualism: language and cognition*, **2020**, 23, pp. 81-86. <https://doi.org/10.1017/S1366728919000567>
149. Hirosh, Z.; Degani, T. Novel word learning among bilinguals can be better through the (dominant) first language than through the second language. *Language Learning*, **2021**, 71, pp. 1044-1084. <https://doi.org/10.1111/lang.12457>
150. Kendon, A. The study of gesture: Some remarks on its history. In J. N. Deely Ed.; *Semiotics*, Plenum Press: New York, 1981. pp. 153–166.
151. Jalbert, A.; Neath, I.; Bireta, T.J.; Surprenant, A.M. When does length cause the word length effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **2011**, 37, pp. 338. <https://doi.org/10.1037/a0021804>
152. Bartolotti, J.; Marian, V. Wordlikeness and novel word learning. In *Proceedings of the Annual Meeting of the Cognitive Science Society*. 2014
153. Schneider, W.; Eschman, A.; Zuccolotto, A. E-prime (version 2.0). Computer software and manual]. Pittsburgh, PA: Psychology Software Tools Inc, 2002.
154. Kroll, J.F.; De Groot, A.M. *Handbook of bilingualism: Psycholinguistic approaches*. Oxford: Oxford University Press, 2005.
155. Fidell, L.S.; Tabachnick, B.G. Preparatory data analysis. In Schinka, J.A. & Velicer, W.F. Eds.; *Handbook of psychology: Research methods in psychology*, John Wiley & Sons, Inc.: Canada, 2003. pp. 115-141.
156. Botting, N.; Riches, N.; Gaynor, M.; Morgan, G. Gesture production and comprehension in children with specific language impairment. *British Journal of Developmental Psychology*, **2010**, 28, pp. 51-69. <https://doi.org/10.1348/026151009X482642>
157. Hogrefe, K.; Ziegler, W.; Wiesmayer, S.; Weidinger, N.; Goldenberg, G. The actual and potential use of gestures for communication in aphasia. *Aphasiology*, **2013**, 27, pp. 1070-1089. <https://doi.org/10.1080/02687038.2013.803515>
158. Kelly, S.D.; Manning, S.M.; Rodak, S. Gesture gives a hand to language and learning: Perspectives from cognitive neuroscience, developmental psychology and education. *Language and Linguistics Compass*, **2008**, 2, pp. 569-588. <https://doi.org/10.1111/j.1749-818X.2008.00067.x>
159. Kushch, O.; Igualada, A.; Prieto, P. Prominence in speech and gesture favour second language novel word learning. *Language, Cognition and Neuroscience*, **2018**, 33, pp. 992-1004. <https://doi.org/10.1080/23273798.2018.1435894>
160. Van Patten, B. Attending to form and content in the input: An experiment in consciousness. *Studies in second language acquisition*, **1990**, 12, pp. 287-301. <https://doi.org/10.1017/S0272263100009177>
161. Wong, W. Modality and attention to meaning and form in the input. *Studies in second language acquisition*, **2001**, 23, pp. 345-368. <https://doi.org/10.1017/S0272263101003023>
162. Bransdorfer, R.L. *Communicative value and linguistic knowledge in second language oral input processing*. (Unpublished doctoral dissertation) University of Illinois at Urbana–Champaign. 1991.
163. Gentner, D. Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj Ed.; *Language development*, Erlbaum: New York, 1982. pp. 301–334
164. Lennon, P. Getting “easy” verbs wrong at the advanced level. *International Review of Applied Linguistics in Language Teaching*, **1996**, 34, pp. 23–36. <https://doi.org/10.1515/iral.1996.34.1.23>
165. Kita, S.; Alibali, M.W.; Chu, M. How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological review*, **2017**, 124, pp. 245-266. <https://doi.org/10.1037/rev0000059>
166. Kroll, J.F.; Van Hell, J.G.; Tokowicz, N.; Green, D.W. The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition*, **2010**, 13, pp. 373-381. <https://doi.org/10.1017/S136672891000009X>
167. Kaushanskaya, M.; Rehtzigel, K. Concreteness effects in bilingual and monolingual word learning. *Psychonomic bulletin & review*, **2012**, 19, pp. 935-941. <https://doi.org/10.3758/s13423-012-0271-5>
168. Buccino, G.; Binkofski, F.; Fink, G.R.; Fadiga, L.; Fogassi, L.; Gallese, V.; ... Freund, H.J. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European journal of neuroscience*, **2001**, 13, pp. 400-404. <https://doi.org/10.1111/j.1460-9568.2001.01385.x>
169. Huang, X.; Kim, N.; Christianson, K. Gesture and vocabulary learning in a second language. *Language Learning*, **2019**, 69, pp. 177-197. <https://doi.org/10.1111/lang.12326>

170. Cook, S.W.; Goldin-Meadow, S. The role of gesture in learning: Do children use their hands to change their minds? *Journal of cognition and development*, **2006**, *7*, pp. 211-232. https://doi.org/10.1207/s15327647jcd0702_4
171. Morales, J.; Calvo, A.; Bialystok, E. Working memory development in monolingual and bilingual children. *Journal of experimental child psychology*, **2013**, *114*, pp. 187-202. <https://doi.org/10.1016/j.jecp.2012.09.002>
172. Schmidt, R. Attention, awareness, and individual differences in language learning. In Chan W.M.; Chin, K.N.; Bhatt, S. & Walker, I. Eds.; *Perspectives on individual characteristics and foreign language education*, De Gruyter Mouton: Germany, 2012.
173. Xu, J. Predicting homework time management at the secondary school level: A multilevel analysis. *Learning and individual differences*, **2010**, *20*, pp. 34-39. <https://doi.org/10.1016/j.lindif.2009.11.001>
174. Rivera, M.; Paolieri, D.; Iniesta, A.; Pérez, A. I.; Bajo, T. Second language acquisition of grammatical rules: The effects of learning condition, rule difficulty, and executive function. *Bilingualism: Language and Cognition*, **2023**, *26*, pp. 639-652. <https://doi.org/10.1017/S1366728922000815>

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.