Title
The Role of Embodied Manual Action in Second Language Word Learning

Author
Morett, Laura

Publication Date
2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

THE ROLE OF EMBODIED MANUAL ACTION IN SECOND LANGUAGE WORD LEARNING

A dissertation submitted in partial satisfaction
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PSYCHOLOGY

by

Laura Morett

December 2012

The dissertation of Laura Morett is approved:

____
Professor Raymond W. Gibbs, Chair

____
Professor Alan H. Kawamoto

____
Professor Brian J. MacWhinney

____
Tyrus Miller
Vice Provost and Dean of Graduate Studies

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

THE ROLE OF EMBODIED MANUAL ACTION IN SECOND LANGUAGE WORD LEARNING

by

LAURA MORETT

Previous research has provided evidence that mental imagery and embodied action can facilitate lexical learning in a novel language. However, it is unclear *how* these factors interact—as well as *why* they play a role—in lexical learning. Through a set of four experiments, this research demonstrated that neither mental imagery nor embodied action directly promotes the acquisition of second language (L2) words. The results of a fifth experiment provide evidence that both mental imagery and embodied action enhance ASL sign acquisition through enactment. Taken together, these results indicate that the body plays a more pivotal role in the early stages of sign acquisition than it does in the early stages of spoken L2 word learning. Overall, the results provide some support for embodied theories of language acquisition and processing and indicate that the saliency of perceptual experiences underlying lexical items and their referents depends on modality of communication.

The Role of Embodied Manual Action in Second Language Word Learning

Regardless of its form, human communication is profoundly multimodal. As such, not only do people use vocalization to communicate, but also movements of the hands, head, eyes, and other body parts. For successful communication to occur, interlocutors must take all of these signals, which are collectively known as known as paralinguistic information, into account, drawing meaning from them holistically. From a very young age, children learn to use paralinguistic information to interpret utterances, as well as to express themselves nonverbally when they cannot speak. Likewise, there is evidence that adults use paralinguistic information to facilitate their comprehension and production of second languages (L2s). In some populations that cannot easily communicate verbally, such as the Deaf community, these embodied signals develop into an autonomous language (sign language), functioning as an alternative mode of communication. All of these lines of evidence indicate that the body plays an important role in language acquisition and processing.

Despite the intuitive evidence of the body's contribution to human communication, traditional linguistic theories (Chomsky, 1965; Pinker, 1984) have de-emphasized it in favor of focusing on the representation of language within the mind. Implicitly or explicitly, these theories are philosophically based on two broader assumptions about human cognition proposed by Fodor (1983): computational theory and modularity of mind. Computational theory of mind holds that the mind is an information processing system and that cognitive processes are symbol-based computations. Modularity of mind maintains that the mind is composed of innately specified modules that are specialized to process only a specific type of information. According to traditional linguistic theories, language is produced and comprehended by applying a finite set of rules to a set of amodal, symbolic linguistic units (phonemes, words, phrases, etc.) without engaging the perceptual systems of the body. Thus, upon hearing a novel sentence, such as "Colorless green ideas sleep furiously," listeners apply a set of algorithms to determine the grammatical classes of the words, access the meanings of the words from their mental lexicons, and combine these two sources of information to infer the meaning of the sentence. If the assumptions of the computational nature and modularity of mind underlying these traditional linguistic theories are accepted, it follows that the body cannot contribute significantly to language processing. Thus, these theories consider paralinguistic information peripheral to language processing, maintaining that core meaning and structure can only be conveyed via language itself.

On the other hand, many modern linguistic theories are compatible with—and even acknowledge—the idea that the body plays a central role in language processing (Barsalou, 1999; Gibbs, 2006). Unlike the traditional theories discussed above, these theories do not assume computational nature or modularity of mind; instead, they view language processing as a data-driven process that is tightly integrated with other cognitive processes (Spivey, 2007). Moreover, rather than viewing language as a static amodal symbolic system, many contemporary theories view language as a dynamic, multimodal system grounded in perceptual experience. Specifically, they posit that, when comprehending language, listeners draw upon their real and imagined perceptual experiences with referents to mentally simulate meaning. For example, upon hearing the sentence "The foolish blue owl offended the shy, well-dressed pussycat," listeners recall their past experiences with owls, pussycats, and offense from real life to imagine the event conveyed by the sentence, even though they have not directly experienced the event described before[1]. Due to their emphasis on the role of the body in language processing, these theories are collectively known as embodied theories of language.

The Role of Embodied Action in Language Processing

Empirical work supporting embodied theories of language has demonstrated that sensorimotor simulations are activated during language processing. In one set of experiments, participants were slower to verify object properties when they were perceived via different modalities than when they were perceived via the same modality (Pecher, Zeelenberg, & Barsalou, 2003, 2004). For example, when presented with the word "apple," participants verified the property words "shiny" and "green," both of which are perceived via sight, quicker than they verified the property words "shiny" and "tart," the latter of which is perceived via taste. These results are similar to those observed in perceptual verification experiments, in which participants are slower to verify a signal when it is presented in a different modality than the preceding signal than when two signals are presented in the same modality (Spence,

[1] That is, except for in the Neighborhood of Make-Believe, where the event described was commonly experienced by viewers.

Nicholls, & Driver, 2001). The similarity of these two sets of results provides evidence that sensorimotor simulations of referent properties are activated when comprehending language, given that language comprehension reflects the same constraints as physical perception of the referents. They support embodied theories of cognition by indicating that language is processed through direct access of perceptual states.

Research from the realm of cognitive neuroscience has provided evidence demonstrating the instantiation of sensorimotor simulations within the brain, as well as evidence indicating that these simulations underlie language production, as well as language comprehension. Upon presentation of both line drawings and names of objects, the production of color words selectively activated an area of the ventral temporal lobe used to process color information. Similarly, the production of action words activated a region of the middle temporal gyrus anterior to the area used to perceive motion (Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995). A related study showed that naming tools also activated this area as well as the left premotor area (Martin, Wiggs, Ungerleider, & Haxby, 1996). Together, these findings indicate that the parts of the brain active in processing the physical characteristics of concepts also subserve the production of speech referring to these concepts. Moreover, they provide evidence that speaking evokes robust, embodied representations of referents rather than amodal symbols, supporting embodied views of language.

Aside from indicating that language processing involves the activation of sensorimotor simulations, empirical work has also demonstrated that meaningful engagement of the motor system during language processing enhances memory encoding and retrieval. For example, spoken instructions are recalled more accurately when they are acted out than when they are repeated aloud (Svensson & Nilsson, 1989). Moreover, spoken instructions for tasks that have been enacted for a longer time period (30 s.) are more likely to be recalled than those that have been enacted for a brief time period (5 s.) (Cohen & Bryant, 1991), indicating that greater engagement of the motor system produces richer, more robust memories. A separate line of research has provided evidence that adults are more likely to produce sought-after words during speech dysfluencies when they gesture than when they keep their hands still (Frick-Horbury, 2002; Frick-Horbury & Guttentag, 1998), indicating that motor engagement through gesture facilitates lexical access. Taken together, the results of all of this work suggests that the enactment of meaningful motions with the body during language processing allows speakers to tap into their semantic representations more effectively, thereby promoting language processing.

One major way in which the body is involved in language processing is through the means of hand motions, including gestures and signs. These hand motions vary on the basis of several qualities, including conventionalization, semiosis, and relationship to speech. To show how different types of hand movements relate to one another on the basis of these characteristics, Adam Kendon and David McNeill (1992) developed a continuum, which is illustrated below in Figure 1. At one extreme of the continuum lies sign language, which is highly conventionalized, segmented and analytic, and occurs in lieu of speech. At the opposite extreme lies gesticulation, which is unconventionalized, global and synthetic, and occurs concurrently with speech. Although it is not plotted on this continuum, McNeill (2005) points out that speech lies beyond sign language in terms of these characteristics, distinguishing itself as the most conventionalized and analytic of the communicative signals. Although iconicity is not plotted on this continuum, it can be inferred that, due to its global and synthetic (i.e., holistic) nature, gesticulation is highly iconic, whereas sign (and, by extension, speech) is the most arbitrary of the signals. The continuum developed by Kendon and McNeill provides a useful means for conceptualizing how the involvement of the body affects the form of human communicative signals, which in turn provides insight into how cognition is used to interpret them in real time conversation.

The iconicity inherent in many signs provides evidence that the perceptual and motor systems are deeply involved in sign language processing, in addition to spoken language processing. According to embodied theories of language, iconicity is a crucial starting point for language because it illustrates the isomorphism between lexical items and their referents, providing listeners with a tangible link between language form and meaning that they can access through their perception and action of referents in the real world. As languages evolve and become more conventionalized, these form-meaning correspondences become less important within communication, and lexical units become increasingly arbitrary (Corballis, 2003; Steklis & Harnad, 1976). Within sign languages, there is evidence of this pattern within the home sign of individual deaf children, as well as pidgin sign languages created by communities of deaf children, both of which are generally more iconic than conventionalized sign languages (Kendon, 1980; Senghas, Kita, & Özyürek, 2004). Even within highly-conventionalized sign languages, recently-coined signs are more iconic than signs that have been in the language longer (Frishberg, 1975). This pattern demonstrates how iconicity contributes to the phylogenetic development of sign languages, illustrating the role that the body plays in sign language evolution.

Ontogenetically, the effect of iconicity on sign language processing is less clear. On one hand, one study (Orlansky & Bonvillian, 1984) showed that only 30% of these children's first 10 signs are iconic, and that this number increases to only 34% at 18 mos., suggesting that iconic signs are not learned prior to arbitrary signs. Another study (Miller, 1987) showed that 3-year-old hearing children unfamiliar with ASL were unable to reliably select the correct referent of iconic signs on the Peabody Picture Vocabulary Test, a standardized, forced-choice measure of vocabulary development (Dunn & Dunn, 1997). On the other hand, two studies have provided evidence that iconicity affects hearing L2 learners of sign languages' sign processing (Campbell, Martin, & White, 1992; Thompson, Vinson, & Vigliocco, 2009), and a survey study of adult Deaf signers suggests that iconic signs are learned at younger ages than non-iconic signs, which are more distributed (Vinson, Cormier, Denmark, Schembri, & Vigliocco, 2008). Although the ontogenetic evidence of iconicity's role in sign language is mixed, most research seems to suggest that iconic signs are more readily associated with their referents than non-iconic signs by both Deaf children and L2 sign language learners, indicating that the embodied experience underlying iconicity improves sign learnability.

Although the form of most speech is less iconic than sign, speech is frequently accompanied by gesture, which exhibits iconicity through its shape and motion. Through its iconicity of form and the information that it conveys, co-speech gesture provides a paralinguistic signal that is closely related yet complementary to speech. On the basis of form and information, McNeill (1992) classified co-speech gestures into four distinct categories: iconic, metaphoric, deictic, and beat. Iconic gestures represent the physical form and affordances of concrete actions and objects referred to via speech through the means of their hand shape and motion (e.g., raising a cupped hand to the mouth to represent drinking). Metaphoric gestures represent the physical form and affordances of the source domains of conceptual metaphors underlying abstract concepts referred to via speech (e.g., moving outspread hands up and down like a scale to represent justice. Deictic gestures (i.e., pointing with the index finger or another part of the body) indicate a concrete, present (or absent) entity to an interlocutor by directing the listener's attention to it. Beat gestures are simple, rhythmic emphasizing gestures that occur synchronously with speech prosody, calling the listener's attention to select concepts expressed via speech (e.g., the slight movement of one hand in time to the words, "One, two, three").

An important question that is often raised is whether gesture facilitates communication for the speaker or the listener. Enhancement of communication for the speaker suggests that gesture provides a means for the speaker to offload the cognitive processes necessary for speaking, facilitating both thought and speech

(Goldin-Meadow, 2003). Enhancement of communication for the listener suggests that gesture's iconicity conveys more information than speech alone, providing an iconic communicative signal in conjunction with speech that creates rich, multimodal representations that are more likely to be recalled (Cohen & Otterbein, 1992). These possibilities are not mutually exclusive; hypothetically, gesture can serve as both a cognitive aid and an effective communication device. Empirical work has demonstrated that speakers are more likely to recall words in a tip-of-the-tongue state when they gesture (Frick-Horbury & Guttentag, 1998, 1998), and that they perform better on a secondary task when they gesture while talking through a cognitively-demanding primary task (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001), providing evidence that gesture lightens speakers' cognitive load. There is also evidence that speakers produce gestures that are more iconic, elaborate, and larger when they are speaking to an interlocutor who is present and visible, as opposed to an interlocutor who is present but occluded or an interlocutor who is on the phone (Alibali, Heath, & Myers, 2001; Bavelas, Gerwing, Sutton, & Prevost, 2008). Finally, speakers are more likely to accurately recall the meanings of speech accompanied by iconic gestures, particularly if it does not form a cohesive narrative unit (Cohen & Otterbein, 1992). Taken together, all of these findings suggest that gesture facilitates both speaking and listening, enriching communication and representation through its embodied and iconic nature.

The Role of Mental Imagery in Language Processing

One important process involved in language processing that is often overlooked by traditional linguistic theories is mental imagery. Mental imagery plays a key role in embodied theories of language processing because it is incorporated into sensorimotor simulations, which are necessary to interpret language. According to embodied theories of language processing, sensorimotor simulations are "reenactment[s] of perceptual, motor, and introspective states acquired during experience with the world, body, and mind" (Barsalou, 2008) (p. 618). Thus, the body plays an integral role in generating sensorimotor simulations, and as such, is inextricably involved in mental imagery formation. When sighted people interpret language, they reenact their visual experiences with referents perceived via their eyes through the means of mental imagery. For sighted people, mental imagery is an integral component of sensorimotor simulations, given the salience of objects' visual affordances. The contribution of mental imagery to language processing is also compatible with models of working memory such as dual coding theory (Paivio, 1990) and the multicomponent model (Baddeley, 1992), both of which posit that, after passing through working memory separately, visual and verbal information combine to form rich, multimodal representations in long term memory. Theoretically and practically, there can be little doubt that mental imagery—and, by extension, embodied action—play an important role in real-time language processing.

Indeed, there are several compelling lines of research demonstrating that listeners engage in mental imagery generation during real-time language processing. One such line of research focuses on interpretation of language conveying motion through verbs. For example, listeners are faster to respond to sentences conveying motion events (e.g., *The car veered to the right*) when they are presented with visual displays depicting motion in the direction opposite that of the sentence than when the displays depict motion in the same direction as the sentence, because in the latter case, the visual imagery interferes with the mental imagery used to interpret the sentence (Kaschak et al., 2005). Moreover, listeners are faster to respond to sentences conveying fluid fictive motion (e.g., *The road runs through the forest*) than they are to respond to sentences conveying restricted fictive motion (e.g., *The road meanders through the forest*), but no similar pattern of results was observed for sentences with the same meaning that did not convey fictive motion (e.g., *The road is in the forest*) (Matlock, 2004). Corroborating these findings, neural activation has been observed in the visual cortices—as well as in the classical language areas—during the interpretation of sentences conveying both literal and fictive motion (Saygin, McCullough, Alac, & Emmorey, 2009). This finding highlights the roles of perception and action in mental imagery formation, demonstrating the close link between mental imagery and embodied action. Taken together, all of these findings provide evidence that listeners generate mental imagery during language interpretation, facilitating their language processing through their perceptual experience with motion.

Another line of research demonstrating the involvement of mental imagery in language processing focuses on language conveying entities; that is, on nouns. Within this line of research, it has been shown that larger, more perceptually salient properties of entities are processed more quickly through language than smaller, less salient properties (Solomon & Barsalou, 2004). For example, participants are quicker to associate *mane* with *lion* than *tail*. In a similar vein, memory for narratives improves when objects are associated with the main actor. For example, participants are quicker to verify an unrelated event mentioned in a narrative when the main actor puts on his shirt than when he takes it off (Glenberg, Meyer, & Lindem, 1987). Similar to the work on motion representation in language, these behavioral findings indicating that object representations are dependent upon mental imagery has been verified by research on neural activation. Neuroimaging research has shown that the visual cortices are active during the processing of animal and color names, whereas the motor cortex is active during the processing of tool names (Martin et al., 1995, 1996). This activation demonstrates that listeners' engagement of mental and motor imagery is dependent upon their perception of salient object properties, again highlighting the perception-action link and the role of the body in mental imagery generation. Considered as a whole, these findings provide evidence that mental imagery, as perceived and re-enacted via the visual system, is an essential component of the sensorimotor simulations underlying nouns and their sundry attributes.

It is important to note that, although gesture and speech convey different information, they arise from the same sensorimotor simulations (Hostetter & Alibali, 2008). Thus, mental imagery is just as essential—if not more so—to gesture processing as it is to speech processing. Indeed, as discussed in the previous section, gesture's iconicity stems from the isomorphism between its visual form and motion and that of its referents. Iconic gesture is unique among communicative signals—including gestures—because it brings together mental imagery and embodied action in a visible, tangible means, providing a particularly powerful means of communication. Several studies provide insight into how mental imagery and embodied action work together to facilitate communication through iconic gesture. One study of individual differences underlying gesture indicates that visuospatial skill was correlated with iconic gesture production, indicating that individuals with more robust mental imagery produce more iconic gestures than individuals with less robust mental imagery (Hostetter & Alibali, 2007). In another study (Hostetter & Alibali, 2010), individuals who had traced a path along a dot pattern were more likely to gesture iconically when describing the path's trajectory than individuals who had seen the path without tracing it, indicating that meaningful embodied action enriched the visual representations underlying gesture. Along similar lines, two other studies (Cook & Tanenhaus, 2009; Goldin-Meadow & Beilock, 2010) provide evidence that performing physical actions during a task increases the tendency to gesture iconically, providing further evidence that embodied action, in conjunction with highly visual stimuli, increases iconic gesturing. Together, these findings confirm that mental imagery and embodied action underlie the sensorimotor simulations underlying iconic gesture, contributing to its unique form and communicative potential.

Mental imagery plays a particularly important role in second language (L2) word learning because it provides a means by which learners can associate arbitrary phonetic forms with their meanings, creating rich, multimodal lexical representations.

This has been demonstrated via the keyword method (Atkinson, 1975), a successful method of L2 word learning that relies on association of words with mental imagery. In the keyword method, learners are instructed to choose a word from their native language that is phonologically similar to the target L2 word, and to formulate a mental image of the referents of these words interacting. For example, for the Spanish word *chico* (boy), a learner might choose the similar-sounding English word chick, and then could imagine a boy holding a chick. This method has been shown to be a more effective strategy for L2 word learning than verbal association of target words and translations (Atkinson & Raugh, 1975; Levin, McCormick, Miller, Berry, & Pressley, 1982; Raugh & Atkinson, 1975; Van Hell & Mahn, 1997). There is also evidence that, particularly in the early stages of L2 learning, the meanings of target words are learned more effectively via physical images, which cue mental imagery through isomorphism, than via verbal translation (Carpenter & Olson, 2011; Chun & Plass, 1996). These findings confirm that mental imagery enhances L2 word learning by adults, and indicates that word learning can be enhanced by encouraging L2 learners to consciously link novel words with mental images of their referents.

Although the role of mental imagery in children's word learning is less clear, it appears that children also associate the phonological forms of words with mental images in order to learn them effectively. Across languages, the meanings of imageable words are learned before the meanings of less imageable words, despite that many of the most frequently-encountered words are closed class words such as articles, which are low in imeageability (Bird, Franklin, & Howard, 2001; Ma, Golinkoff, Hirsh-Pasek, McDonough, & Tardif, 2009; Rinaldi, Barca, & Burani, 2004). Although there is some cross-linguistic variability in whether nouns are learned earlier than verbs (Gentner, 1982; Tardif, 1996), across languages, both nouns and verbs are learned before other types of words, such as modifiers and articles, that are less imageable. Nevertheless, imageability predicts the age at which words are learned above and beyond word class, indicating that the ease with which words can be associated with representative mental images is more primary to children's word learning than syntactic constraints, such as word class (McDonough, Song, Hirsh-Pasek, Golinkoff, & Lannon, 2010). All of these findings suggest that, like adults, children associate the phonetic forms of words with mental images of their referents, enhancing their ability to learn words by creating rich, multimodal representations.

Given that sign language is more iconic than spoken language, and that sign language production relies on the body to a greater extent than spoken language production, it follows that mental imagery may be even more essential to sign language processing than spoken language processing. Indeed, there is empirical evidence that this is the case. Compared with hearing English speakers, Deaf signers are more adept at confabulating and manipulating complex mental images (Emmorey, Kosslyn, & Bellugi, 1993), and show greater right hemisphere engagement when processing visuospatial stimuli (Bosworth & Dobkins, 2002; Emmorey & Kosslyn, 1996). These results suggest that, as a result of using the body to communicate via sign language, Deaf signers develop more robust mental imagery than hearing speakers. While it is unclear whether hearing speakers process highly imageable words using only the left hemisphere or both the right and left hemispheres (Fiebach & Friederici, 2004; Sabsevitz, Medler, Seidenberg, & Binder, 2005), there is evidence that Deaf signers rely mainly on the right hemisphere when processing highly imageable signs (Emmorey & Corina, 1993). Considered as a whole, these findings suggest that mental imagery plays a crucial role in sign language processing by allowing sign language users to associate signs with the visual attributes of their referents, undergirding the isomorphism between iconic signs and their referents.

How Mental Imagery and Embodied Action Contribute to Word Learning

Given that mental imagery and embodied action combine to play an important role in fluent language comprehension, it stands to reason that they play a similar—if not greater—role in language learning. Indeed, a substantive body of research has examined the contributions of these processes to various aspects of language learning through the means of iconic gesture, which visibly incorporates mental imagery through its iconicity, as well as embodied action through its articulation. This work has examined the effect of iconic gesture on various aspects of both first and second language acquisition, including phonology (Iverson & Thelen, 1999; Kelly & Lee, 2011), fluent speech (Iverson & Goldin-Meadow, 2005; Sueyoshi & Hardison, 2005), and language transfer (Stam, 2006; Yoshioka & Kellerman, 2006). Arguably, the aspect of language learning to which iconic gesture—and, by proxy, mental imagery and embodied action—contribute the most is word learning. Words are the most conceptually basic units of language and are learned before anything else, so they have the greatest potential to be influenced by enriched mental representations.

Several lines of research provide evidence that iconic gesture facilitates word learning in both first and second languages. In early childhood, it has been shown that, when taught iconic gestures representing entities (baby signs), children can express themselves via these gestures before they can say the corresponding words (Acredolo & Goodwyn, 1988), since gross motor control of the limbs precedes fine motor control of the vocal tract. Moreover, even without instruction, children commonly produce iconic gestures representing referents before they produce words representing them, and they produce gesture/word combinations before they produce two word combinations (Iverson & Goldin-Meadow, 2005). Both of these findings indicate that speech and gesture are based on the same mental simulations, but that children can express themselves more effectively via gesture before they can do so via speech. School-aged children also learn novel words more effectively when they are presented with representative iconic gestures than with representative images, particularly when they re-enact the gestures themselves (Tellier, 2008). Taken together, all of these findings provide evidence that, through the means of iconic gesture, both mental imagery and embodied action contribute to children's ability to learn words in their first language.

Research has also provided evidence that representative body movement—including iconic gesture—helps adults learn lexical items in a novel second language. In the Total Physical Response Method (Asher, 1969), an L2 vocabulary teaching method, the instructor teaches words by saying them while demonstrating their meanings using the body, and learners re-enact the instructor's actions using their own bodies to demonstrate comprehension of the words' meanings. This method has been shown to be particularly effective for teaching L2 vocabulary, with adult students sometimes reaching the attainment levels of children (Asher, Kusudo, & Torre, 1974; Asher & Price, 1967; Kunihira & Asher, 1965). There is also evidence that representative iconic hand gestures enhance word learning in adult L2 learners, similar to children. In an early study on this topic, undergraduate first semester French students recalled the meanings of more figurative expressions presented with iconic gestures representing their meanings in addition to English glosses (e.g., placing two index fingers next to one another, pointing down for *Comme les deux doigts de la main* [like two peas in a pod]) than expressions presented with English glosses alone, without gesture (Allen, 1995). In a more recent study (Kelly, McDevitt, & Esch, 2009), undergraduates recalled the meanings of French verbs better when they had been presented with representative iconic gestures in addition to glosses (e.g., raising and tipping hand in front of lips for *nomu* [to drink]) than when they'd been presented with glosses alone, but worse when they had been presented with non-representative iconic gestures (e.g., placing hands together and opening them outward for *nomu* [to drink]). All of this research provides evidence that representative embodied action enhances L2 word learning through its visual iconicity, which allows learners to pair the phonological forms of words with mental imagery and embodied action representing their referents.

In addition to research showing that iconic gesture facilitates word learning, research comparing spoken and sign language acquisition also provides evidence of the contributions of mental imagery and embodied action to language acquisition. A pioneering longitudinal study (Bonvillian, Orlansky, & Novack, 1983) demonstrated that deaf, signing children of Deaf parents reached major developmental milestones, including first sign and first sign combination, an average of 2-3 mos. prior to established norms for children learning to speak. A subsequent review of the research on this topic indicated that this sign advantage is particularly robust for lexical item learning, but that it disappears for more complex aspects of language, such as syntax (Meier & Newport, 1990). Taken together with research showing that iconic gesture facilitates the acquisition of novel spoken words, these findings indicate that mental imagery and embodied action play pivotal roles in early lexical acquisition in signed as well as spoken language.

Phylogenetic evidence indicates that mental imagery and embodied action contributed to the evolution of language through the means of iconic gesture, similar to its contribution to ontogenetic language development, as discussed above. Several prominent researchers (Corballis, 2003; Tomasello, 2008) have argued that humans communicated via the manual modality (i.e., through iconic gesture) before they did so via the oral modality (i.e., through speech). This argument is predicated on gesture's greater iconicity, which was conducive to early communication, but which was later supplanted by speech due to its physical characteristics (less effortful, travels better over long distances, etc.). Like the ontogenetic evidence demonstrating that gestures and signs can be learned prior to speech, this phylogenetic evidence indicates that iconic gesture's incorporation of mental imagery and embodied action allowed humans to begin to communicate with one another, paving the way for spoken language evolution.

The Present Work

Although much research has demonstrated indirectly that mental imagery and embodied action contribute to language acquisition, to date, no published work has attempted to disambiguate their influences. Through the means of carefully designed and selected stimuli, the present work sought to examine the influences of imagery and action on lexical learning in both signed and spoken languages. In particular, the present study focused on word learning in adult second language (L2) learners, because our primary interest was to better understand how imagery and action contribute to language processing in adults. The present study focused on lexical learning due to the previous evidence that imagery and action contribute to word (and sign) learning, and due to the primacy of lexical learning within language acquisition. Both word and sign learning were examined to examine the extent to which imagery and action underpin language acquisition and processing, regardless of modality. Finally, both active and passive versions of the stimuli were used to examine the extent to which active engagement in imagery and/or action is necessary to facilitate lexical acquisition.

Based on the research discussed above, it was predicted that iconic gesture viewing and enactment would result in the highest number of L2 words and signs recalled, given that it incorporates both mental imagery and embodied action, creating multimodal traces comprising visuospatial, verbal, and motoric codes. It was also anticipated that mental imagery would result in the next highest number of L2 words and signs recalled, due to the combination of visuospatial and verbal codes, and that meaningless embodied action and text/speech only would result in the lowest number of words recalled. Furthermore, it was expected that active learning conditions would allow participants to recall more words and signs than passive conditions, due to its greater engagement of the sensorimotor system through iconic gesture enactment and active mental imagery formation. Finally, it was predicted that iconic signs would be recalled more accurately than non-iconic signs, due to the isomorphism between their form and articulation and the affordances of their referents. Overall, it was anticipated that both mental imagery and embodied action would contribute to word and sign learning, and that greater engagement of the sensorimotor system through these processes would result in superior lexical learning.

Preliminary Norming Study: A Representative Sample of Words and Iconic Gestures

Many previous studies of the effect of iconic gesture on L2 word learning (Kelly & Lee, 2011; Kelly et al., 2009; Tellier, 2008) are unclear about how target words and their corresponding gestures were selected. Consequently, it is unclear to what extent target words differ in dimensions that affect word learning, such as imageability, concreteness, and meaningfulness. Moreover, it is unclear whether the gestures used in these studies are representative of iconic gestures that participants associate with target words outside of these studies, so the ecological validity of their findings is questionable. The current study sought to overcome these limitations through the means of a preliminary norming study in which participants rated words on relevant dimensions of meaning and produced gestures representing the words' meanings. Based on the data collected in this norming study, words from three word classes (abstract words, concrete nouns, and concrete verbs) with the most consistent ratings and iconic gestures were selected for inclusion in the five principal studies comprising this project.

Method

Participants. Participants were undergraduate students (n = 29) unfamiliar with Hungarian who did not participate in any of the five main studies described subsequently. In return for their participation, participants received partial course credit.

Materials. Materials consisted of 60 English words that were glosses of potential Hungarian target words (see Table 1). These words fell into three classes: abstract words, concrete nouns, and concrete verbs.

Procedure. Participants were given an inventory containing all 60 words in randomized order, and were asked to rate each word on the basis of three factors, imageability, concreteness, and meaningfulness, using a scale ranging from 1 (Not imageable / concrete / meaningful at all) to 7 (Completely imageable / concrete / meaningful). After they had completed this inventory, participants were given a list of all 60 words presented in a different randomized order, and were asked to say each word while producing a gesture representing the meaning of the word. Participants were asked to produce a gesture for each and every word, and were video recorded while saying the words and producing gestures.

Results. Quantitative results for the rating and gesture data can be seen in Table 1. As can be seen, there was a great deal of variation between both the ratings and gestures that participants produced for each word, both between and within word class categories, indicating that participants' semantic conceptualizations of the words varied considerably. On the basis of this data, four words (boldfaced in Table 1) were selected from each category for inclusion in Study 5. Specifically, the four words chosen from each category had the most consistent gestures (i.e., the highest number of gestures with a similar form) and the most consistent ratings (i.e., ratings closest to the average across word class categories) across semantic dimensions.

Study 1: The Role of Passive Mental Imagery and Embodied Action on L2 Word Learning

The objective of the first study was to investigate whether viewing iconic gesture could facilitate the learning of novel L2 words via mental imagery and embodied manual action. While previous studies (Kelly et al., 2009; Tellier, 2005) have indicated that viewing iconic gestures concurrently with novel L2 words can enhance the acquisition of their meanings, the results of these studies have not addressed the cognitive mechanisms responsible for this effect. The current study seeks to determine whether mental imagery and embodied action are responsible for the facilitatory effect of gesture, as well as to disambiguate the relative contributions of these factors, by comparing the degree to which L2 word learning is enhanced by iconic gesture to its enhancement via other stimuli eliciting mental imagery, embodied action, or neither.

In order to systematically test the effects of mental imagery and embodied action on L2 word learning, iconic gesture (+ mental imagery, + embodied action) was compared to physical imagery (+ mental imagery, - embodied action), beat gesture (- mental imagery, + embodied action), and text (- mental imagery, - embodied action). Images, like iconic gestures, cue mental imagery through isomorphism between their visuospatial features and those of their referent; however, unlike iconic gestures, images (particularly of nouns) do not imply embodied action. Conversely, beat gestures, like iconic gestures, convey embodied action, but unlike iconic gesture, beat gestures do not convey the visuospatial features of referents.

Text was used as the baseline condition because it does not convey mental imagery or embodied action through its form, but it nonetheless represents the meaning of language visually. Based on the results of previous research (Kelly et al., 2009; Tellier, 2008), it was expected that iconic gesture would facilitate L2 word learning to the greatest extent (due to its combination of embodied action and mental imagery), followed by images (due to isomorphism of form), followed by beat gestures and text (due to a lack of direct correspondences between their haptic and visual forms and those of the referents).

To explore whether the effects of mental imagery and embodied action differ by word class, we selected target words of 3 types: concrete nouns, concrete verbs, and abstract words. It should be noted that this component of the study was exploratory, given that no work has directly examined the effects of mental imagery and embodied action (or iconic gesture) on the acquisition of L2 words of different classes. However, given research suggesting that nouns are learned before verbs (Gentner, 1982), and that iconic gesture is predicated upon embodied action (Hostetter & Alibali, 2008), it was predicted that iconic gesture would result in superior learning of concrete verbs, relative to concrete nouns and abstract words. Such an interaction between word class and learning condition would demonstrate that embodied action and mental imagery exert a greater influence on verb acquisition than on noun and abstract concept acquisition, indicating that the role of embodiment in language learning varies in accordance with lexical semantic features.

Method

Participants. Participants were undergraduate students (n = 26; mean age: 20.25; 15 females) unfamiliar with the target L2 (Hungarian). Hungarian was chosen as the target L2 because it is typologically unrelated to English and most other widely spoken languages. In return for their participation, participants received partial course credit.

Materials. Novel vocabulary consists of 12 Hungarian words (see Table 2) selected for inclusion in the study based on the data from the norming study described previously. Videos of gestures were created by filming a native speaker of Hungarian and English reproducing gestures produced by the greatest number of participants in the norming study. In general, these gestures represented the affordances of the word referents that they represented. For concrete nouns, such as *kulcs* (key) and *sepru* (broom), gestures represented actions typically performed with each object (e.g., turning one hand as if using a key to unlock a door, placing two hands above one another and moving them back and forth as if sweeping with a broom). For concrete verbs, such as *löni* (to shoot) and *mászni* (to climb), gestures typically represented the meanings of the verbs directly (e.g., forming the index finger and thumb into a gun-like shape and mimicking shooting, mimicking climbing by moving two clenched hands upwards in tandem). For abstract words, such as *orom* (joy) and *trefa* (joke), gestures typically represented a physical state associated with the concept (e.g., smiling to represent happiness associated with joy, laughing to represent the physical reaction to joking). Thus, regardless of which word class gestures were associated with, they represented affordances via embodied action.

Images were line drawings representing the English glosses of target words, excerpted from the International Picture Naming Project (Szekely et al., 2004). Sound files of the pronunciation of each word and gloss were excerpted from the iconic gesture condition and played during presentation of referent images and text of words and glosses; however, the beat gesture condition was played with its own soundtrack so that any verbal emphasis inherent in this condition was preserved. A post-experimental questionnaire was also created and administered to experimental participants to investigate other factors affecting performance on the experimental task.

Procedure. Study 1 consisted of a learning phase and a test phase separated by a 5-minute interval, and lasted about 30 minutes. In each trial of the three blocks of the learning phase, a Hungarian word was presented as speech for 2000 ms., and after a 1000 ms. interstimulus interval, the corresponding English word was presented for 2000 ms. For words (of both languages) randomly assigned to condition (1), iconic gesture, participants viewed video of a gesture that represented their meaning; for words in condition (2), beat gesture, participants viewed video of simple emphasizing gestures made in time to speech prosody that does not convey their meaning; for words in condition (3), image, participants viewed images representing the words' meaning; and for words in condition (4), text, participants viewed words presented as text.

The test phase of this study consisted of a single block in which each Hungarian word from the learning block was presented as speech for 2000 ms. Participants responded by saying the corresponding English word or "skip" if they could not remember it. Participants completed the test phase at three intervals following the learning phase—five minutes, one week, and one month—to examine how learning conditions affected long-term L2 word recall

Results. L2 word recall was quantified by scoring responses using a binary coding scheme (1 = correct, 0 = incorrect/skipped), and by converting scores into proportion correct for each participant and condition (to control for unscorable responses due to factors such as unintelligibility or technical errors in running the recall task). Proportional scores were submitted to repeated measures ANOVAs, in which participant and word were used as fixed factors.

This analysis revealed a main effect of recall interval, $F_{pp}(2, 51) = 12.16$, $p < .001$, $\eta_p^2 = .36$; $F_{word}(2, 18) = 6.42$, $p = .008$, $\eta_p^2 = .42$. Bonferroni-corrected post hoc analyses showed that participants recalled more L2 words after five minutes than they recalled after one week ($p_{pp} = .03$; $p_{word} =.02$) and one month ($p_{pp} = .001$; $p_{word} = .06$). There was also a main effect of learning condition, $F_{pp}(3, 77) = 7.70$, $p < .001$, $\eta_p^2 = .26$; $F_{word}(3, 27) = 7.04$, $p = .001$, $\eta_p^2 = .44$, see Figure 2. Post-hoc analyses showed that participants learned more words via text than they did via beat gesture ($p_{pp} =$

.002; p_{word} = .008), iconic gesture (by participant; p_{pp} = .05; p_{word} = .34) and images (p_{pp} = .08; p_{word} = .05). However, the interval by condition interaction failed to reach significance. Finally, when word class (abstract word, concrete noun, concrete verb) was entered into the by word analysis as a between-participants factor, it failed to reach significance, as well.

Discussion. The results of this study demonstrate that L2 words are recalled more accurately after short (five minute) than after longer (one week and one month) intervals. This finding is uncontroversial, given that prior research on semantic memory has established that recall for L2 information (including vocabulary) decays over time (Bahrick, 1984). Surprisingly, however, the results also demonstrate that L2 words learned via text are recalled more accurately than L2 words learned via iconic gesture, beat gesture, or images, with no differences in recall accuracy for words learned via the latter three conditions. This finding is inconsistent with the prediction that words learned via iconic gesture would be recalled best, followed by words learned via mental imagery, followed by words learned via text. However, this finding does not necessarily support generative theories of L2 acquisition, given that it may be caused by a number of factors, including stimuli that are not sufficiently evocative of mental imagery and embodied action. In fact, the advantage of text over iconic gesture, beat gesture, and imagery cannot be explained by extant theories of L2 acquisition. The most plausible explanation for this textual advantage is transfer-appropriate processing (Morris, Bransford, & Franks, 1977), which posits that the presence of similar cues (text) in both the text learning condition and test trials caused superior recall for L2 words learned via text, creating this illusory and theoretically inconsistent effect.

Study 2: Learning-Test Similarity Control

The objective of this study was to determine whether the advantage of text observed in Study 1 was due to the presence of text in both the text learning condition and the test condition. This result may be an artifact of transfer-appropriate processing, which occurs when similar retrieval cues are present at encoding and retrieval (Morris et al., 1977). To determine whether transfer-appropriate processing was responsible for the facilitatory effect of text, the recall test was modified so that words were presented as speech only, without any text. If, under these conditions, participants still recalled more words presented as text than words presented via iconic gesture, beat gesture, or images, than it could be concluded that presenting words as text facilitates L2 word learning above and beyond these other stimuli even after controlling for transfer-appropriate processing effects. However, if, on the other hand, participants did not recall any more L2 words presented as text than L2 words presented as representational gesture, beat gesture, or speech, it could be concluded that the facilitatory effect of text observed in Study 1 was due to transfer-appropriate processing. Given a lack of conclusive reported evidence showing that L2 words learned as text are recalled better than L2 words presented via other modalities (e.g., images), it was predicted that no significant advantage of text would be found in Study 2.

Participants. Participants were undergraduate students (n = 26; mean age: 20.64; 14 females) unfamiliar with Hungarian who had not participated in Study 1. In return for their participation, participants received partial course credit.

Materials. Materials were identical to those used in Study 1.

Procedure. Learning conditions were identical to those used in Study 1. The recall test was also identical to that used in Study 1, except for that Hungarian words were presented as speech only, accompanied by a screen displaying the task instructions in lieu of the words.

Results. As in Experiment 1, L2 word learning was quantified proportionally and was submitted to repeated measures ANOVAs with participant and word as fixed factors. This analysis revealed a main effect of recall interval, $F_{pp}(2, 51) = 7.11$, $p = .003$, $\eta_p^2 = .31$; $F_{word}(2, 10) = 47.14$, $p < .001$, $\eta_p^2 = .90$, see Figure 3. Bonferroni-corrected post-hoc analyses showed that participants recalled more words after five minutes than they recalled after one week ($p_{pp} = .04$; $p_{word} = .001$) and one month ($p_{pp} = .05$; $p_{word} = .004$). However, learning condition failed to reach significance, as did the interval by condition interaction. Furthermore, when word class was entered into the by word analysis, it failed to reach significance, as well. These results indicate that the superior recall for words learned via text observed in Experiment 1 was due to the presence of text in both learning and test trials.

Discussion. Consistent with the main prediction, the results of this study failed to show that words learned via text were recalled any better than words learned via iconic gesture, beat gesture, or images when text was omitted from recall trials. Thus, this finding indicates that the advantage for words learned via text observed in Study 1 was caused by transfer-appropriate processing, due to the presence of text in both the text learning condition and recall trials.

The more surprising finding from Study 2 was that iconic gesture still failed to produce an advantage over beat gesture, images, and text for L2 word learning. As such, the results of Study 2 still fail to support the prediction that mental imagery and embodied action would facilitate word learning through iconic gesture. What might be the cause of this null effect for iconic gesture on L2 word learning, especially in light of the gestural advantage shown by several previous studies? One plausible explanation is that the referents of the iconic gestures used for some of the words may have been unclear. Thus, it is possible that, upon seeing these gestures while hearing unfamiliar L2 words, participants may have assumed that the words had different meanings than they actually did. If this is the case, hearing the English glosses of these words may have confused participants, unnecessarily adding to their cognitive load and cancelling any facilitatory effect that iconic gesture may have otherwise produced.

Study 3: Gesture-Meaning Mapping Control

The objective of this study is to explore whether the lack of advantage observed for representational gesture in Studies 1 and 2 was due to difficulties mapping gestures onto the words that they were intended to represent. In a norming study, representational gestures were selected by asking a group of native English speakers to make a gesture representing the meaning of each English gloss, and by choosing the most common gestures to represent the glosses and their corresponding Hungarian words. However, it is possible that the representational gestures chosen may logically map onto the meanings of more than one word (e.g., the gesture representing to climb may also map onto ladder), which may in turn increase participants' cognitive load when they hear the "correct" word, thereby negatively impacting performance on the L2 word learning task. Study 3 attempted to control for gesture-meaning mismatches by presenting English glosses prior to Hungarian words to ensure that participants associated representational gestures with their intended meanings. Additionally, the text learning condition was replaced with a speech only learning condition. This replacement was made because speech with no visual representation is a more accurate baseline condition than speech with text, given that all of the other conditions include speech, but none of them include text. In light of these changes, it was anticipated that words learned via iconic gesture would be recalled most accurately, followed by words learned via imagery, followed by words learned via beat gesture and speech. These predictions were based on the incorporation of embodied action and mental imagery into iconic gesture, and the simultaneous storage of verbal and visuospatial information into working and long-term memory.

Participants. Participants were undergraduate students (n = 27, mean age: 21.05; 14 females) unfamiliar with Hungarian who had not participated in Studies 1 or 2. In return for their participation, participants received partial course credit.

Materials. Materials were identical to those used in Studies 1 and 2.

Procedure. Learning conditions were similar to those used in Studies 1 and 2, with the following differences: English glosses were presented prior to Hungarian words; English and Hungarian words were presented as speech only (accompanied by a blank screen) in the baseline condition. The recall test condition was identical to that used in Study 2.

Results. As in Experiments 1 and 2, L2 word learning was quantified proportionally and was submitted to repeated measures ANOVAs with participant and word as fixed factors. This analysis revealed a main effect of recall interval, $F_{pp}(2, 53) = 24.81$, $p < .001$, $\eta_p^2 = .61$; $F_{word}(2, 14) = 36.59$, $p < .001$, $\eta_p^2 = .84$, see

Figure 4. Bonferroni-corrected post-hoc analyses showed that participants recalled more words after five minutes than they recalled after one week ($p_{pp} = .04$; $p_{word} = .002$) and one month ($p_{pp} = .04$; $p_{word} < .001$). However, learning condition failed to reach significance, as did the interval by condition interaction. Taken together with the findings of Experiments 1 and 2, these results confirm that, in the passive form, mental imagery and embodied action do not significantly enhance the acquisition of L2 words.

Several additional analyses were conducted to investigate the source of the null effect of condition. First, as in Experiments 1 and 2, word class was entered into the by word analysis, but failed to reach significance. Second, 19 additional participants who had not participated in Experiments 1, 2, 3, or 4 rated how well each iconic gesture represented its corresponding word on a scale of 1 (not at all representative) to 7 (totally representative; see Table 3). These ratings were averaged for each gesture, and the resulting average ratings were entered into the by word analysis as a covariate to determine whether gestures' representativeness affected their corresponding words' learnability. However, even when this factor was included as a covariate, learning condition failed to reach significance, as did the test interval by learning condition interaction. The results of these additional analyses indicate that, even when taking the variability between word classes and individual word-gesture mappings into account, iconic gesture and its components, mental imagery and embodied action, do not significantly facilitate L2 word learning.

Discussion. Considered in conjunction with the results of Studies 1 and 2, the results of Study 3 indicate that neither iconic gesture nor embodied action or mental imagery significantly facilitates L2 word learning. As such, this finding fails to support the prediction that iconic gesture conveys its referents transparently via a combination of embodied action and mental imagery, facilitating L2 word learning. Moreover, this finding fails to replicate the results of other similar studies, which have shown that representative iconic gestures enhance recall of lexical items (words and phrases) by novice L2 learners (Allen, 1995; Kelly et al., 2009). However, the null effect for iconic gesture viewing is actually consistent with the results of a study (Tellier, 2008) in which children's L2 word learning was facilitated via iconic gesture enactment, but not via iconic gesture viewing. Thus, it is possible that participants must actively engage in gesturing, meaningless embodied action, or mental imagery themselves while learning L2 words, so that their effects can be detected. If this is the case, then the predictions may simply need to be modified slightly, rather than overhauled in favor of an alternate theory, such as generativism.

Study 4: The Role of Active Mental Imagery and Embodied Action on L2 Word Learning

The objective of this study is to examine the effects of active embodied motion and mental imagery formation on L2 word learning. Past research (Tellier, 2008) has shown that, for young children, gesture viewing is insufficient to facilitate word learning; only gesture enactment can enhance lexical acquisition. Thus, it is possible that a facilitatory effect may emerge if participants enact iconic gestures while learning L2 words, even though it is not present when participants view gestures without enacting them during L2 word learning. Study 4 was designed to test this possibility by examining L2 word learning under the active equivalents of the 4 conditions used in Studies 1-3: gesture enactment (+ mental imagery, + embodied action), mental imagery formation (+ mental imagery, - embodied action), meaningless hand motion (- mental imagery, + embodied action), and verbal repetition (- mental imagery, - embodied action). It was hypothesized that gesture enactment would facilitate L2 word learning to the greatest extent, followed by mental imagery formation, followed by meaningless hand motion and verbal repetition. Similar to the results for Studies 1-3, this pattern of results was predicted due to the incorporation of both active embodied action and mental imagery in gesture enactment, and due to the usefulness of mental imagery in L2 word learning, in combination with verbal information (Atkinson & Raugh, 1975; Raugh & Atkinson, 1975). Thus, both active engagement in mental imagery and meaningful embodied action were expected to contribute to successful L2 word learning.

Participants. Participants were undergraduate students (n = 28, mean age: 20.67; 16 females) unfamiliar with Hungarian who had not participated in Studies 1-3. In return for their participation, participants received partial course credit.

Materials. Materials consisted of the iconic gesture videos and audio tracks of the 12 words taught in Studies 1-3.

Procedure. Like Studies 1-3, the learning phase of Study 4 consisted of 3 blocks of words varying by condition. In each learning trial, participants viewed a video of a Hungarian-English bilingual saying an English gloss while making a gesture that represented its meaning for approximately 2000 ms., and then, after a 1000 ms. interstimulus interval, viewed a video of the bilingual speaker saying the Hungarian equivalent of the gloss while making the same gesture for 2000 ms. After one repetition of these events in the same sequence, for words in condition (1), gesture enactment, participants were instructed to enact the gesture that they had viewed in the video while saying the English and Hungarian words aloud; for words in condition (2), mental imagery formation, participants were instructed to close their eyes and visualize the word's meaning while saying the English and Hungarian words aloud; for words in condition (3), meaningless hand motion, participants were instructed to make an X-shaped hand motion three times while saying the English and Hungarian words aloud; and for words in condition 4, participants were simply instructed to repeat the Hungarian and English words aloud. The testing phase was identical to that of Studies 2-3: Participants were presented with each Hungarian word as speech only, and were instructed to say the English meaning of each word aloud.

Results. As in Experiments 1-3, L2 word learning was quantified proportionally and was submitted to repeated measures ANOVAs with participant and word as fixed factors. This analysis revealed a main effect of recall interval, $F_{pp}(2, 56) = 12.90$, $p < .001$, $\eta_p^2 = .52$; $F_{word}(2, 20) = 10.37$, $p = .001$, $\eta_p^2 = .51$, see Figure 5. Bonferroni-corrected post hoc analyses revealed that participants recalled more words after five minutes than they recalled after one week ($p_{pp} = .05$; $p_{word} = .02$) or one month ($p_{pp} = .03$; $p_{word} = .01$). However, learning condition failed to reach significance, as did the interval by condition interaction. These results suggest that enactment does not seem to result in L2 word learning that is any more effective than word learning via mental imagery, arbitrary hand motion, or verbal repetition.

As in Experiment 3, two additional analyses were conducted to determine the impact of differences in stimuli (words and gestures) on the main results described above. In the first of these additional analyses, word class was entered into the by word analysis as a between-participants factor, but failed to reach significance. In the second of these analyses, the gesture representativeness ratings described above were entered into the by word analysis as a covariate. However, even when taking this factor into consideration, learning condition failed to reach significance, as did the condition by test interval interaction. As in Experiment 3, the results of these additional analyses indicate that variation in word class and gesture representativeness was not responsible for the failure of iconic gesture and its components, mental imagery and embodied action, to enhance L2 word learning.

Data from Studies 3 and 4 were combined and submitted to additional analyses to compare the effects of active and passive engagement of mental imagery and embodied action on L2 word learning. These analyses revealed a marginal effect of engagement type (active or passive) by participant, $F(1, 55) = 3.28$, $p = .08$, $\eta_p^2 = .11$, but failed to show this effect by word. Additionally, these analyses confirmed that more words were recalled after five minutes than after one week or one month, $F_{pp}(2, 110) = 34.50$, $p < .001$, $\eta_p^2 = .55$; $F_{word}(2, 34) = 39.53$, $p < .001$, $\eta_p^2 = .70$, and that more words were learned via mental imagery than via speech only (by participant), $F_{pp}(3, 165) = 4.31$, $p = .007$, $\eta_p^2 = .13$; $F_{word}(3, 51) = 1.16$, $p = .33$, $\eta_p^2 = .06$.

Discussion. The results of Study 4 provide some evidence that mental imagery facilitates L2 word learning. This finding is consistent with the results of studies showing that active engagement in mental imagery representative of L2 words' meanings enhances recall for novel L2 words (Atkinson, 1975; Atkinson & Raugh, 1975; Raugh & Atkinson, 1975). However, the results fail to show that enactment of iconic gestures representing referents during L2 word learning enhances recall. This finding is inconsistent with the results of prior research showing that gesture enactment enhances L2 word recall among children (Tellier, 2008), as well as with the Gesture-Simulation Bootstrapping Hypothesis's prediction that iconic gesture would facilitate L2 word learning. Although the cause of this null effect for gesture enactment is unclear, it is possible that embodied action may not play as large of a role in word learning by adults as by children, given that adults' motor system has already developed, and that adults may be inhibited by the cultural expectation that gesturing during word learning is abnormal.

Together, the results of Studies 1-4 suggest that active methods of word learning may be more effective than passive methods. More specifically, they suggest that, although the roles of mental imagery and embodied action in L2 word learning may not be noticeable via passive stimuli, the effect of mental imagery on L2 word learning is noticeable when participants actively engage in it. These findings suggest that the predictions of the Gesture-Simulation Bootstrapping Hypothesis are at least partially correct, but that this hypothesis could be improved to reflect the importance of mental imagery over embodied action in L2 word learning.

Study 5: The Role of Mental Imagery and Embodied Action on L2 Word Learning

The objective of this study was to examine the effects of mental imagery and embodied action on L2 sign learning. Although Studies 1-4 failed to show that mental imagery and embodied action significantly enhance spoken L2 word learning, it is possible that they may facilitate L2 sign acquisition. Paralleling the design of Study 4, this study examined the influence of mental imagery and embodied action on sign acquisition through active engagement of the learner's mental imagery and embodied action via carefully selected learning conditions. These conditions were sign enactment (+ mental imagery, + embodied action), mental imagery formation (+

mental imagery, - embodied action), meaningless hand motion (- mental imagery, + embodied action), and sign comprehension (- mental imagery, - embodied action). As in Study 4, it was predicted that sign enactment would result in the greatest number of signs correctly recalled, followed by mental imagery formation, and then meaningless hand motion and sign comprehension.

Due to its employment of sign language, Study 5 also offered an alternative method of testing the influence of mental imagery and embodied action on sign acquisition: sign iconicity. In this study, signs were classified into 3 categories: Iconic, metaphorical, and arbitrary (see Figure 6). Iconic signs depicted their referent holistically or metonymically (e.g., pantomiming hammering for hammer); metaphorical signs represented the source domain of the conceptual metaphor structuring their referent (e.g., cupped hands moving forward three times, as if conveying an entity of information from the signer to the listener, for to teach); and arbitrary signs bore no structural resemblance to their referent (e.g., two fingers from both hands taping one another repeatedly for name). The distinctions between these sign types were supported by empirical data collected from a separate group of participants unfamiliar with ASL (O'Brien, 1999), ensuring that they were applicable to the target population of the current study. Based on the results of this previous study, it was predicted that concrete and metaphorical signs would be recalled with greater accuracy than arbitrary signs, due to the isomorphism between their visual form and that of their referents.

Participants. Participants were undergraduate students (n = 29, mean age: 21.15; 18 females) with normal hearing who were unfamiliar with ASL. In return for their participation, participants received partial course credit.

Materials. Materials in this study consisted of silent videos of a native signer of ASL producing 20 signs, and audio clips of synthesized speech conveying the English meanings of each sign. As can be seen in Figure 6, signs were classified on the basis of their iconicity, comprising concrete nouns and verbs as well as abstract concepts.

Procedure. Like Studies 1-4, this study consisted of three sessions, the first of which included both a learning and test phase, and the second and third of which included only a test phase. In learning trials, participants were presented with video of a randomly selected sign (~2500 ms.), and after a 1000 ms. interstimulus interval, were presented simultaneously with the corresponding English gloss as text and audio (2500 ms.). After one additional repetition of this sequence of events, participants performed one of four actions. For words presented in the enactment condition, participants enacted the sign with their own hands; as such, this condition included both mental imagery and embodied action. For words presented in the imagery condition, participants closed their eyes and visualized the sign's referent in their mind's eye without moving their hands; as such, this condition included mental imagery, but not embodied action. For words presented in the motion condition, participants made an X-shaped motion with their dominant hand three times; as such, this condition included embodied action but not mental imagery. For words presented in the comprehension condition, the learning sequence was repeated one additional time, and participants were not explicitly told to do anything; as such, this condition did not include either mental imagery or embodied action. Within each experimental session, each sign was randomly assigned to one of these four conditions in a within-participants design, such that five signs were presented in each condition for each participant. The learning phase consisted of 3 blocks comprising 20 trials apiece (one for each sign), yielding a total of 60 learning trials altogether.

Following the learning phase in the first session, participants were given a 5-minute break, and then completed the test phase. In test trials, upon being presented with English glosses as text and audio, participants were asked to produce the corresponding ASL sign. Participants were instructed to try to recall the sign as best they could, but were told that they could say "skip" to move on if they could not recall a sign. During test trials, participants' signing was recorded by a video camera set approximately 45° to the left of their central viewing point. The test phase consisted of one block of 20 trials (one for each sign). Overall, the first experimental session lasted about 30 minutes.

To examine how long-term recall of signs varied by condition, participants returned to the lab for two follow-up sessions held one week and four weeks after the first session. In each of these sessions, participants completed the test phase in the manner described above. Each of the follow-up sessions lasted approximately 10 min. apiece.

Results. Similar to Studies 1-4, sign recall was quantified using a binary coding scheme (1 = correct; 0 = incorrect/skipped). Total number of signs recalled correctly for each participant and condition were converted into proportions to control for unscorable responses caused by technical errors in running the recall task (which accounted for less than 5% of the data). For a sign to be coded as correct, it must have been performed using the same hand (dominant/non- dominant, as specified per participant on a post-experimental questionnaire), and must have had the same hand shape and movements as the correct ASL sign, as modeled by the signer.

To address the question of whether learning condition affects sign recall, proportional data were submitted to repeated measures ANOVAs, using participant and sign as fixed factors. These analyses revealed significant main effects of learning condition, $F_{pp}(3, 87) = 7.16$, $p < .001$, $\eta_p^2 = .29$; $F_{sign}(3, 45) = 14.07$, $p < .001$, $\eta_p^2 = .48$, and recall interval, $F_{pp}(1, 29) = 10.99$, $p = .004$, $\eta_p^2 = .38$; $F_{sign}(1, 15) = 18.16$, $p = .001$, $\eta_p^2 = .55$, but failed to reveal a significant condition-by-interval interaction, $F_{pp} > 1$; $F_{sign} > 1$; see Figure 7. Bonferroni-corrected post-hoc analyses showed greater recall accuracy for signs learned via enactment than via mental imagery ($p = .04$), motion ($p = .06$), and comprehension ($p = .05$), as well as greater recall accuracy after an interval of 5 minutes than 1 week ($p < .01$) and 4 weeks ($p < .001$). These results indicate that enactment facilitates the acquisition of novel signs by hearing adult learners unfamiliar with sign language across both short and long learning-test intervals.

To address the question of whether iconicity affects the learning and recall of ASL signs, sign type (iconic, metaphorical, arbitrary) was entered into a repeated measures ANOVA, using sign as a fixed factor. This analysis failed to reveal a main effect of sign category on recall, $F(1, 17) = 1.13$, $p = .35$, $\eta_p^2 = .12$; see Table 1. This result indicates that, similar to deaf children, hearing adult learners do not benefit significantly from iconicity when learning novel signs.

Discussion. The results of this study revealed that enactment facilitated sign learning more effectively than visualization of sign referents, performance of meaningless hand movements, or simple sign comprehension. However, the results failed to demonstrate that iconic signs are learned more effectively than arbitrary signs. Considered as a whole, these results suggest that enactment enhances ASL sign recall and production in hearing learners by actively engaging both mental imagery and embodied action. In particular, the enactment advantage during sign learning indicates that the motor system plays a key role in sign acquisition. Of note, only meaningful motion (i.e., sign enactment)—not arbitrary motion (i.e., X-shaped motions)—enhanced sign acquisition, indicating that the enactment advantage is due to isomorphism between the visual signs and referents. This indicates that embodied action is only an effective sign learning aid insofar as it is based on mental imagery.

Unfortunately, the results of the current study do not provide insight into why adult L2 learners fail to benefit from the iconicity inherent in some signs, and in sign language in general. One possible explanation is that, L2 sign learners go through a developmental stage in the initial stages of language learning in which they are unable to associate the visuospatial properties of iconic signs with the affordances of their referents. When learning their first set of signs, adults unfamiliar with ASL are unable to relate their semantic and phonological properties to other similar signs, which may negate their ability to recognize iconicity. Alternatively, hearing adults' experience with spoken languages, in which iconicity is sparse, may lead them to assume that language is not iconic, causing them to ignore any physical correspondences between signs and their referents. Finally, the novelty of processing language in the visuospatial modality may place a heavy cognitive load on adult L2 learners unfamiliar with sign language, negating any benefits that iconicity may have bestowed. In any case, future research is necessary to test between these possibilities and to clarify the cause of this null effect.

General Discussion

As stated in the introduction, it is undeniable that language processing is profoundly multimodal, given the multitude of informative paralinguistic cues available to listeners in naturalistic conversational settings. Indeed, several lines of research have shown that children rely on these cues to infer the meanings of unfamiliar utterances (Iverson & Goldin-Meadow, 2005; Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989), suggesting that these cues allow language learners to map linguistic forms to their meanings through their iconicity. However, few studies have systematically examined the cognitive processes subserving the mappings between iconicity of hand movements and word referents. The current work attempted to fill this gap by examining the impacts of embodied action and mental imagery on word learning by adult L2 learners of Hungarian and American Sign Language (ASL). It was predicted that embodied action and mental imagery would both significantly enhance learners' acquisition of novel L2 words and signs, given that both are components of iconicity, which maps transparently onto referents' affordances. The findings of the current work were mixed, with the results of Studies 1-4 failing to support these predictions, and the findings of Study 5 providing some support for them. Taken together, the results of Studies 1-5 provide limited evidence supporting the roles of mental imagery and embodied action in L2 word learning, suggesting that verbal association may be more essential to word learning than these processes.

The current work attempted to overcome the limitations of other studies of the effect of manual iconicity on lexical learning in several ways. First, iconic gestures and ASL signs used in this work were selected on the basis of empirical data collected from populations similar to those who completed Studies 1-5 to ensure ecological validity. Second, in Studies 1-4, unlike some other studies (Kelly et al., 2009), only stimuli that were either representative of the meanings of target words or neutral were used, given that gestures and pictures that are unrepresentative of accompanying speech generally do not appear in natural conversation and instructional materials. Third, recall was tested at intervals of 5 min., 1 week, and 4 weeks to determine whether the results generalized to long-term memory of words. Fourth, this work included several studies that were similar to one another (Studies 1-4), as well as to a previous study (Kelly et al., 2009) to ensure reliability of the results through replication. Finally, to provide a more comprehensive view of the roles of mental imagery and embodied action on lexical learning, the current study examined the influences of these factors on ASL sign learning in addition to spoken word learning. By controlling for all of these factors while extending previous research in a novel direction, the current work represents a more rigorous approach than previous work has offered, ensuring reliability and validity of the results.

The current work focused on the effects of mental imagery and embodied action in L2 word learning by adults for two reasons. First, several independent lines of research have indicated that mental imagery and embodied action play important roles in word learning by children, whereas the evidence of the impact of these factors on word learning by adults is more sparse. Nevertheless, two studies (Kelly & Lee, 2011; Kelly et al., 2009) have shown that iconic gesture can facilitate L2 word learning by adults, which suggests that these factors also play an important role in adult lexical acquisition. Second, the magnitude of the impact of mental imagery and embodied action on L2 word learning in adults should be indicative of the role that these processes play in native language processing, particularly in conversational settings, where paralinguistic cues incorporating them are available. Although L2 word learning differs qualitatively from native language conversation, both involve mapping of concepts to linguistic forms via the same representations and cognitive processes. The results of the current work suggest that, while mental imagery and embodied action may subserve language processing, they are neither necessary nor sufficient to support it.

The Roles of Mental Imagery and Embodied Action in L2 Word Learning

Studies 1-4 examined the independent roles and interactions that passive and active engagement in mental imagery and embodied action play in spoken L2 word learning. The lack of differences in word learning as a function of condition suggest either that the conditions used in the current study do not adequately test mental imagery and embodiment, or that these factors may not play an integral role in L2 word learning by adults. Given that adults are particularly effective word learners (Snow & Hoefnagel-Höhle, 1978), it is possible that participants of Studies 1-4 were using alternative methods to associate target words with their meanings, such as phonological association or generation. One reason why the variation in recall between the learning conditions of Studies 1-4 may have been more modest than that observed in previous studies is that stimuli accompanying words in Studies 1-4 were either consistent with or neutral to the meanings of the target words. This greater similarity between conditions may explain the lack of significant differences between conditions, as well as why the results of the current study failed to replicate those of previous work (e.g., Kelly et al., 2009; Tellier, 2008).

Surprisingly, the results failed to show that meaningful embodied action, in the form of representative iconic gesture, facilitated L2 word learning and recall more effectively than mental imagery. This finding is inconsistent with several previous studies showing superior recall of L2 words learned via viewing or enactment of iconic gesture (Kelly et al., 2009; Tellier, 2008) or enactment of meaningful motion using the body (Asher, 1969; Asher et al., 1974). There are several possibilities why iconic gesture may have failed to facilitate L2 word learning in the current research. One possibility is that the gestures chosen to represent target words may not have been sufficiently iconic, and thus, may not have been as imagistic and meaningful as gestures used in other studies. However, if this were the case, a significant effect should have emerged when iconicity ratings were entered into the analysis as a covariate, as they were in Studies 3 and 4. Thus, it appears that a lack of gesture iconicity is not the main cause of the null effect of iconic gesture. An alternative possibility is that the learning phase may have been too brief to allow for associations between gestures and target words to be formed, so participants may have relied on alternate means of association, such as phonological linking. Future research should test this possibility by examining the acquisition of L2 words during a more extended period under similar learning conditions, and by comparing the results to those of Studies 1-4.

In addition to addressing the roles of mental imagery and embodied action in L2 word learning, the results of studies 1-4 also address the issue of whether these factors must be active or passive to facilitate L2 word learning. Given the null effects observed in Studies 1-3 and the results of previous studies suggesting that only active embodied action enhances L2 word learning (Asher et al., 1974; Tellier, 2008), Study 4 was conducted as a methodological corollary of Studies 1-3 that employed active versions of the stimuli used in Studies 1-3. In contrast to the results of the previous studies cited above, the results of Study 4 failed to reveal positive effects of either mental imagery or embodied action. Notably, the null effects of Studies 3 and 4 persisted even when gesture iconicity was taken into account through a covariate analysis. These null effects indicate that, even in their active forms, mental imagery and embodied action do not contribute significantly to L2 word learning. There are several reasons why effects of active mental imagery and embodied action were observed in previous studies, but not in the current study. First, the current work was conducted with adults, who may rely less on mental imagery and embodied action when learning words than children. Second, the current work included English glosses of target words, which may have further mitigated the impact of these factors on L2 word learning. Third, in the current work, target words were presented in a decontextualized manner, rather than in a broader narrative context. In order to systematically determine whether these methodological differences were responsible for the differences in results between Study 4 and previous related studies, future work should address these methodological inconsistencies.

Overall, the results of Studies 1-4 suggest that the roles of mental imagery and embodied action by adults are peripheral in L2 word learning. More specifically, the results indicate that the viewing of iconic gesture, images, and text during L2 word learning result in comparable recall of target words across both long and short learning-test intervals. The results also demonstrate that L2 words are recalled more accurately over short (5 min.) learning-test intervals than longer intervals (1 week, 1 month). Finally, secondary analyses indicated that gesture iconicity was not responsible for the results, and that active elicitation of mental imagery and embodied action did not produce a stronger effect than passive elicitation of these processes. Taken together, these findings fail to replicate the results of work showing that representative iconic gesture viewing and enactment enhance L2 word learning (Kelly et al., 2009; Tellier, 2008), suggesting that the word learning techniques of adult L2 learners are already sufficiently effective, and that mental imagery and embodied action have a negligible impact on them.

The Roles of Mental Imagery and Embodied Action in L2 Sign Learning

Study 5 investigated the roles of iconicity and enactment on the acquisition of ASL signs by hearing adult L2 learners. The results revealed that enactment facilitated sign learning more effectively than visualization of sign referents, performance of meaningless hand movements, or simple sign comprehension. However, the results failed to demonstrate that iconic signs are learned more effectively than arbitrary signs. As such, the results of Study 5 support the prediction that mental imagery and embodied action would enhance the acquisition of ASL signs through enactment, but failed to support the prediction that iconic and metaphorical signs would be recalled better than arbitrary signs. Taken together, the results of

Study 5 suggest that enactment enhances ASL sign recall and production in hearing learners through the creation of motorically rich lexical traces, but that iconicity does not significantly strengthen associations between signs and referents.

The finding that enactment of ASL signs during learning enhances sign recall indicates that the motor system plays a key role in L2 lexical acquisition, particularly for sign language. Of note, only meaningful motion (i.e., sign enactment)—not arbitrary motion (i.e., X-shaped motions)—enhanced sign acquisition. This result is consistent with embodied theories of cognition, which maintain that the mental representations underlying language derive from meaningful interactions between the body and the world (Barsalou, 1999). It is also consistent with work showing that meaningless repetitive motion can disrupt the formation of visuospatial representations (Vandierendonck, Kemps, Fastame, & Szmalec, 2004). The finding that enactment is more effective at promoting sign learning than visualization of referents via mental imagery indicates that meaningful engagement of the motor system results in richer, more robust mental representations of signs than imagery alone. This finding indicates that meaningful engagement of the motor system through sign enactment increases the chance that novel signs will be retrieved successfully, particularly by inexperienced learners, over and above the association cues that mental imagery alone provides when paired with signs.

The facilitatory effect of enactment on ASL sign learning by hearing English speakers observed in Study 5 raises the question of why a similar effect wasn't observed in Study 4 for spoken Hungarian words. There are several possible reasons for this seeming discrepancy. The most obvious reason concerns the modality of the language. The production of sign language engages the hands and arms, requiring more motor involvement than the production of spoken words. Given that hand movements are a necessary part of signing, enactment may facilitate sign acquisition because it activates the motoric aspects of sign representations. On the other hand, while gesture may accompany speech, it is not an integral part of speech production. Thus, the motoric representations that gesture enactment activates are peripheral to lexical representations. Repetition of words, on the other hand, may activate motor representations of speech articulation, which are integral to lexical representations. Thus, it is possible that verbal repetition may have produced greater benefits than gesture enactment in Study 4, which could explain the marginal effect of active over passive engagement. Another possible reason for the difference between the results of studies 4 and 5 may be differences in the tasks used to assess L2 lexical learning. In study 4, participants produced the English translations of Hungarian words, whereas in study 5, participants produced ASL signs when prompted with their corresponding English words. Given that forward translation (L1 to L2) is more difficult than backwards translation (L2 to L1), it is possible that the facilitatory effects of enactment may have been more apparent in study 5, which uses forward translation as its assessment, than study 4, which uses backwards translation as its assessment.

Unfortunately, the results of the current study do not provide insight into why adult L2 learners fail to benefit from the iconicity inherent in some signs, and in sign language in general. One possible explanation is that, like children, adults go through a developmental stage in the initial stages of language learning in which they are unable to associate the visuospatial properties of iconic signs with the affordances of their referents. Although adults are generally familiar with the affordances of common objects, it is possible that this inability to associate them with their corresponding signs derives from insufficient linguistic context, rather than from insufficient domain-general knowledge, which has been proposed to explain children's insensitivity to iconicity. When learning their first set of signs, adults unfamiliar with ASL are unable to relate their semantic and phonological properties to other similar signs, which may negate their ability to recognize iconicity. Alternatively, hearing adults' experience with spoken languages, in which iconicity is sparse, may lead them to assume that language is not iconic, causing them to ignore any physical correspondences between signs and their referents. Finally, the novelty of processing language in the visuospatial modality may place a substantial cognitive load on adult L2 learners unfamiliar with sign language, negating any benefits that iconicity may have bestowed. Needless to say, future research is necessary to test between these possibilities and to clarify the cause of this null effect.

Overall, the results of Study 5 demonstrate that adult L2 learners can take advantage of enactment, but not iconicity, to facilitate their acquisition of sign language. More specifically, the results indicate that enactment results in robust, detailed representations of signs in learners' minds, which allow them to recall and produce signs more accurately than other methods that involve mental imagery or embodied action alone. However, the results also demonstrate that arbitrary signs are recalled with similar accuracy as iconic and metaphorical signs, suggesting that isomorphism between the forms of signs and the affordances of referents does not strengthen sign-referent associations. Taken together, these results partially support the claim that both mental imagery and embodied action contribute to the acquisition of ASL signs, but are not fully consistent with the assertions of the Perceptual Symbol Systems Hypothesis in its strong form (Barsalou, 1999). Thus, similar to the results of Studies 1-4, the results of Study 5 suggest that, while mental imagery and embodied action contribute to lexical learning by adult L2 learners, they are neither necessary nor sufficient for word learning to occur.

Conclusions

Considered as a whole, the results of Studies 1-5 provide limited support for embodied theories of language processing (Barsalou, 1999; Gibbs, 2006), because they provide mixed evidence that mental imagery and embodied action are involved in L2 word learning. It should be noted, however, that the results of the current work do not support strong versions of embodied language processing, which predict that iconicity and mental imagery consistent with referents should greatly enhance lexical learning by activating relevant embodied experiences. On the other hand, the results fail to explicitly support generative theories of language processing (Chomsky, 1965; Pinker, 1984), which predict that mental imagery and embodied action should not facilitate L2 lexical learning at all. Although the null results of Studies 1-4 suggest that participants may have relied on techniques that do not include mental imagery or embodied action, such as verbal association, to map L2 words to their meanings, it is equally possible that they were using mental imagery to make these associations. Moreover, the facilitatory effect of sign enactment in Study 5 is explicitly inconsistent with generative theories of language processing, which predict that involvement of the motor system should not affect L2 sign learning at all. Thus, while the results are not completely consistent with either class of language processing theories, they are more compatible with the claims of embodied theories than generative theories.

The results of Studies 1-5 reveal important information about the roles of mental imagery and embodied action in L2 lexical learning in particular. First, the contrast between the results of Studies 1-4 and Study 5 suggest that mental imagery and embodied action may play a more central role in sign acquisition—at least, among hearing speakers—than they do in spoken word learning. This may be the case because hearing L2 learners are less familiar with sign language than spoken language, and, as a result, are more likely to rely on mental imagery and embodied action than verbal associations when learning signs than learning words. To determine the plausibility of this explanation, Study 5 should be conducted with Deaf individuals unfamiliar with ASL, in order to determine whether they demonstrate a different pattern of results. Second, the results of Studies 1-4 suggest that the enhancing effect of iconic gesture on L2 word learning observed in previous studies (Asher et al., 1974; Kelly et al., 2009; Tellier, 2005) may be dependent upon a strict set of conditions to emerge. In particular, these conditions may include competing stimuli that are inconsistent with word referents, involvement of the entire body rather than just the hands, and/or participants who are children rather than adults. Thus, the results of Study 5 suggest that the facilitatory effect of iconic gesture on L2 word learning demonstrated in previous studies is relatively constrained and brittle, and does not easily generalize to differing conditions.

Aside from illustrating the roles of mental imagery and embodied action in lexical learning, the results of Studies 1-5 also provide insight into the roles that these processes play in other aspects of L2 acquisition, as well as language processing in general. The observation that iconic gesture, imagery, beat gesture, and text are equally effective as methods of word learning suggests that all of these types of stimuli can be components of lexical representations, and that they are active to some degree during both first and second language processing. Thus, paradoxically, the null results of Studies 1-4 support the conceptual basis of this work on the multimodality of language and communication. On the other hand, the results of Study 5 indicate that, in the initial stages of learning, relevant motoric and imagistic components of ASL signs are more salient within representations than verbal associations between signs and their referents. This indicates that, at least for hearing learners, sign representations may be more firmly grounded in the embodied experience of signing during the initial stages of learning than spoken lexical representations are based on the embodied characteristics of their referents. Thus, for this population, the act of signing itself may enhance not only communication, but sign recall itself within the initial stages of sign language acquisition.

Overall, the results of Studies 1-5 provide mixed evidence that the body participates in lexical learning through the means of mental imagery and embodied action. As such, they demonstrate that, in certain circumstances, L2 lexical learning relies on the body, in addition to the mind, indicating that L2 acquisition is an embodied process. While future research should further clarify the conditions under which iconic gesture, as well as other paralinguistic signals, facilitate L2 word learning, the current study represents a substantive first step in the investigation of the role of the body in L2 word learning, as well as L2 acquisition and language processing more generally.

Table 1. English words considered for use in Studies 1-4, with ratings and number of similar gestures collected in pre-experimental norming study (boldfaced words selected for use in Studies 1-4).

Word	Class	Concrete	Imageable	Meaningful	# of Same Gestures
Bird	Noun	5.37 (2.00)	6.28 (1.67)	4.10 (1.93)	5
Glasses	Noun	4.64 (2.27)	5.67 (1.99)	3.87 (2.10)	3
Glove	Noun	5.08 (2.08)	6.11 (1.65)	2.89 (1.89)	10
Key	**Noun**	**5.59 (1.87)**	**6.14 (1.75)**	**4.34 (2.00)**	**10**
Priest	Noun	5.35 (1.74)	6.07 (1.38)	5.00 (1.54)	6
Hammer	**Noun**	**5.58 (1.93)**	**6.07 (1.71)**	**3.35 (1.92)**	**9**
Knife	Noun	5.65 (1.94)	6.17 (1.63)	3.86 (1.94)	9
Shark	Noun	5.19 (1.96)	6.07 (1.46)	3.90 (1.74)	2
Bell	Noun	5.67 (1.59)	6.03 (1.48)	4.24 (1.96)	9
Book	Noun	6.00 (1.49)	6.45 (1.18)	5.28 (1.79)	8
Broom	Noun	5.54 (1.75)	5.97 (1.80)	3.07 (1.77)	9
Floor	Noun	5.81 (1.47)	6.00 (1.67)	3.83 (2.05)	7
Flag	Noun	6.00 (1.49)	6.10 (1.61)	5.69 (1.56)	5
Shoulder	Noun	5.48 (1.65)	5.93 (1.51)	4.14 (1.81)	5
Stone	Noun	5.84 (1.70)	6.30 (1.27)	3.85 (1.92)	8
Well	Noun	5.13 (1.98)	5.73 (1.61)	3.77 (1.90)	2
Toy	Noun	5.85 (1.49)	5.83 (1.54)	4.79 (1.66)	0
Apple	Noun	5.81 (1.55)	6.34 (1.32)	4.59 (1.96)	9
Brush	**Noun**	**5.56 (1.53)**	**5.90 (1.44)**	**3.62 (1.84)**	**10**
Watch	**Noun**	**5.70 (1.46)**	**6.07 (1.44)**	**4.97 (1.66)**	**10**
To shoot	**Verb**	**4.38 (2.08)**	**5.36 (1.89)**	**4.39 (1.66)**	**7**
To smoke	Verb	4.22 (2.14)	5.69 (1.73)	3.97 (2.13)	10
To write	Verb	4.85 (2.05)	6.00 (1.54)	5.90 (1.05)	10
To deliver	Verb	4.22 (2.04)	5.41 (1.55)	4.38 (2.04)	10
To fight	Verb	4.52 (1.97)	5.52 (1.62)	5.07 (1.93)	3
To climb	**Verb**	**4.19 (1.74)**	**5.50 (1.48)**	**4.25 (1.86)**	**4**
To catch	Verb	4.31 (1.70)	4.94 (1.63)	4.17 (1.89)	2
To fall	Verb	4.08 (1.94)	5.61 (1.79)	4.82 (1.98)	3
To sew	**Verb**	**4.46 (2.00)**	**5.66 (1.40)**	**4.11 (1.79)**	**5**
To shout	Verb	4.44 (1.97)	6.03 (1.57)	4.69 (1.89)	5
To fly	Verb	4.19 (1.98)	5.72 (1.44)	4.83 (2.07)	9
To run	Verb	4.59 (2.15)	6.10 (1.40)	4.76 (1.77)	8
To throw	Verb	3.63 (1.84)	4.62 (1.83)	3.54 (1.75)	4

To hit	**Verb**	**4.52 (2.03)**	**6.00 (1.44)**	**4.28 (1.85)**	**7**
To rise	Verb	4.67 (1.94)	5.41 (1.68)	4.00 (1.89)	8
To sleep	Verb	5.04 (2.03)	6.48 (1.27)	5.93 (0.96)	10
To drink	Verb	4.85 (2.11)	6.00 (1.58)	5.03 (1.61)	10
To burn	Verb	4.52 (1.95)	6.17 (1.39)	4.45 (1.80)	3
To awaken	Verb	4.89 (2.01)	6.00 (1.39)	4.97 (1.57)	6
To open	Verb	4.67 (2.02)	5.66 (1.72)	4.17 (2.04)	6
To believe	Abstract	4.41 (2.26)	5.24 (1.92)	6.03 (1.15)	2
To think	Abstract	4.63 (2.31)	5.38 (2.08)	6.07 (1.07)	6
Love	Abstract	4.33 (2.15)	5.66 (1.67)	6.34 (1.14)	4
Beauty	Abstract	3.88 (1.95)	5.64 (1.50)	5.39 (1.45)	0
Boredom	Abstract	3.85 (2.05)	4.90 (2.01)	3.97 (1.68)	4
Knowledge	Abstract	3.96 (2.08)	4.52 (2.21)	5.69 (1.39)	0
Challenge	Abstract	4.31 (1.92)	5.11 (1.68)	5.28 (1.27)	0
Death	Abstract	5.07 (2.20)	6.38 (0.90)	6.21 (1.21)	2
Wish	Abstract	4.26 (2.33)	5.03 (1.99)	5.72 (1.33)	0
Hope	Abstract	3.93 (2.20)	4.62 (2.18)	5.93 (1.36)	2
Illness	**Abstract**	**4.96 (2.07)**	**5.55 (1.48)**	**5.24 (1.62)**	**5**
Joy	**Abstract**	**4.44 (2.33)**	**6.00 (1.51)**	**6.14 (1.62)**	**6**
Sports	**Abstract**	**5.07 (1.86)**	**6.24 (1.41)**	**5.00 (1.79)**	**5**
Poverty	Abstract	5.11 (1.60)	6.38 (0.86)	6.07 (1.00)	0
Suspicion	Abstract	4.08 (2.15)	4.32 (2.21)	4.64 (1.54)	4
Joke	**Abstract**	**4.19 (2.15)**	**4.82 (2.14)**	**5.29 (1.58)**	**6**
Strength	Abstract	4.84 (2.24)	5.43 (5.17)	5.86 (1.35)	0
To save	Abstract	4.44 (1.97)	5.17 (1.49)	5.55 (1.48)	0
Health	Abstract	4.81 (2.04)	5.39 (1.73)	6.00 (1.47)	0
To achieve	Abstract	4.44 (2.06)	5.33 (1.44)	5.93 (1.47)	0

Table 2. Hungarian words used in Studies 1-4, with English glosses, by word class.

Hungarian word	**English gloss**	**Word class**
Betegség	Illness	Abstract
Öröm	Joy	Abstract
Tréfa	Joke	Abstract
Unott	Bored	Abstract
Kalapács	Hammer	Noun
Kulcs	Key	Noun
Óra	Watch	Noun
Seprű	Broom	Noun
Löni	To shoot	Verb
Mászni	To climb	Verb
Megütni	To hit	Verb
Varrni	To sew	Verb

Table 3. Gesture representativeness ratings for English glosses of Hungarian words used in Studies 1-4 collected in post-experimental norming study.

English word	**Word class**	**Representativeness Rating**
Illness	Abstract	5.89 (1.29)
Joy	Abstract	6.05 (1.03)
Joke	Abstract	6.00 (1.00)
Bored	Abstract	6.58 (0.69)
Hammer	Noun	5.97 (1.27)
Key	Noun	5.95 (1.22)
Watch	Noun	6.11 (0.74)
Broom	Noun	4.95 (1.22)
To shoot	Verb	3.89 (1.52)
To climb	Verb	5.16 (1.54)
To hit	Verb	5.68 (0.89)
To sew	Verb	6.53 (1.02)

Figure 1. Kendon's continuum, as characterized by McNeill (2005).

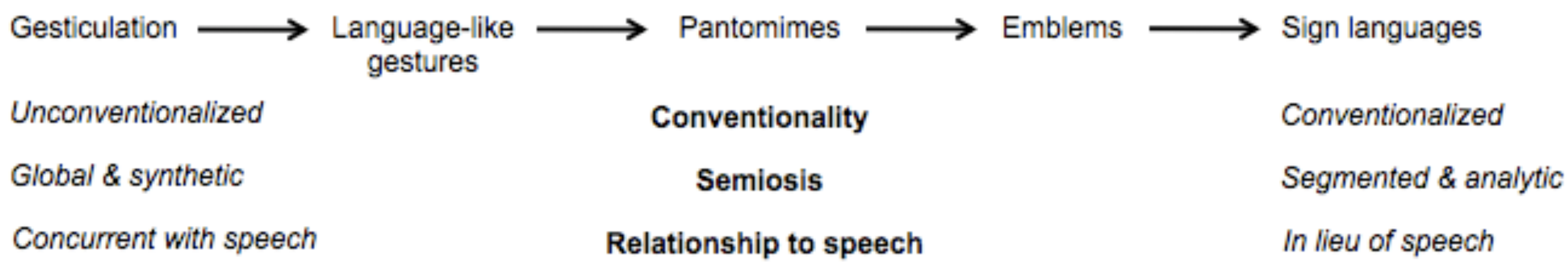


Figure 2. Percent of word meanings recalled by condition and recall interval for Expt. 1 (error bars represent *SEM*).

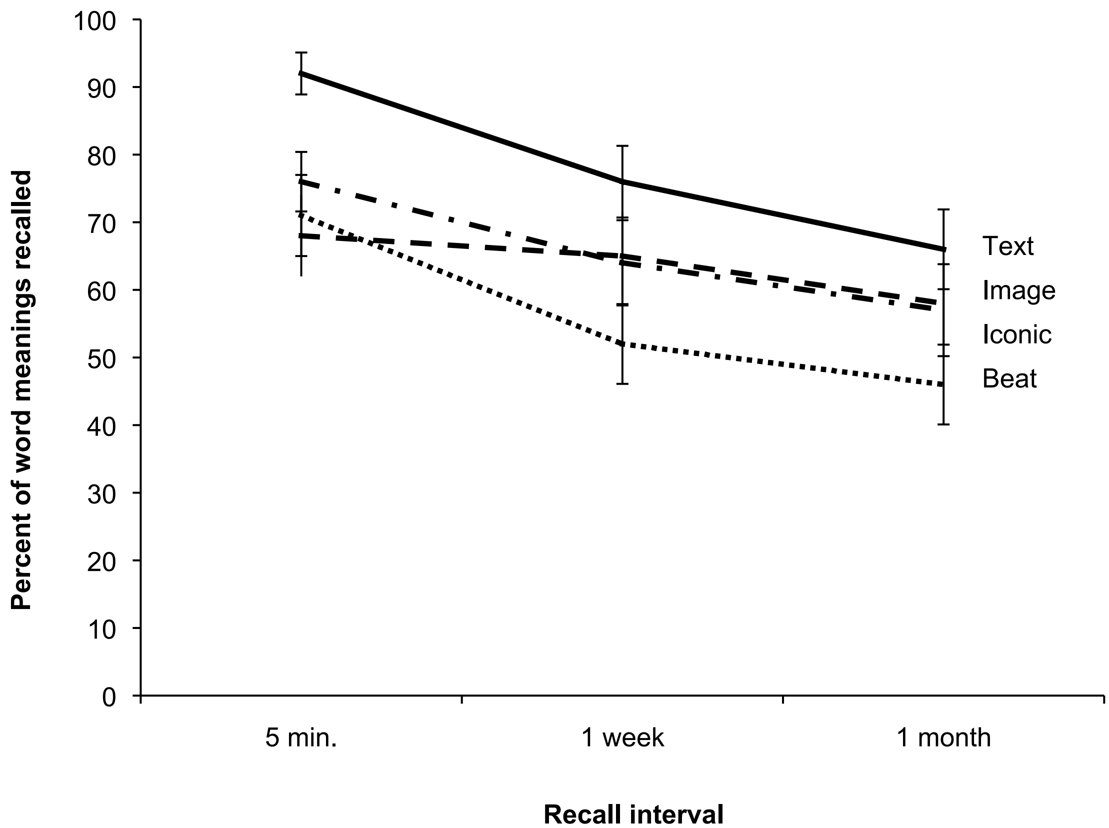


Figure 3. Percent of word meanings recalled by condition and recall interval for Expt. 2 (error bars represent *SEM*).

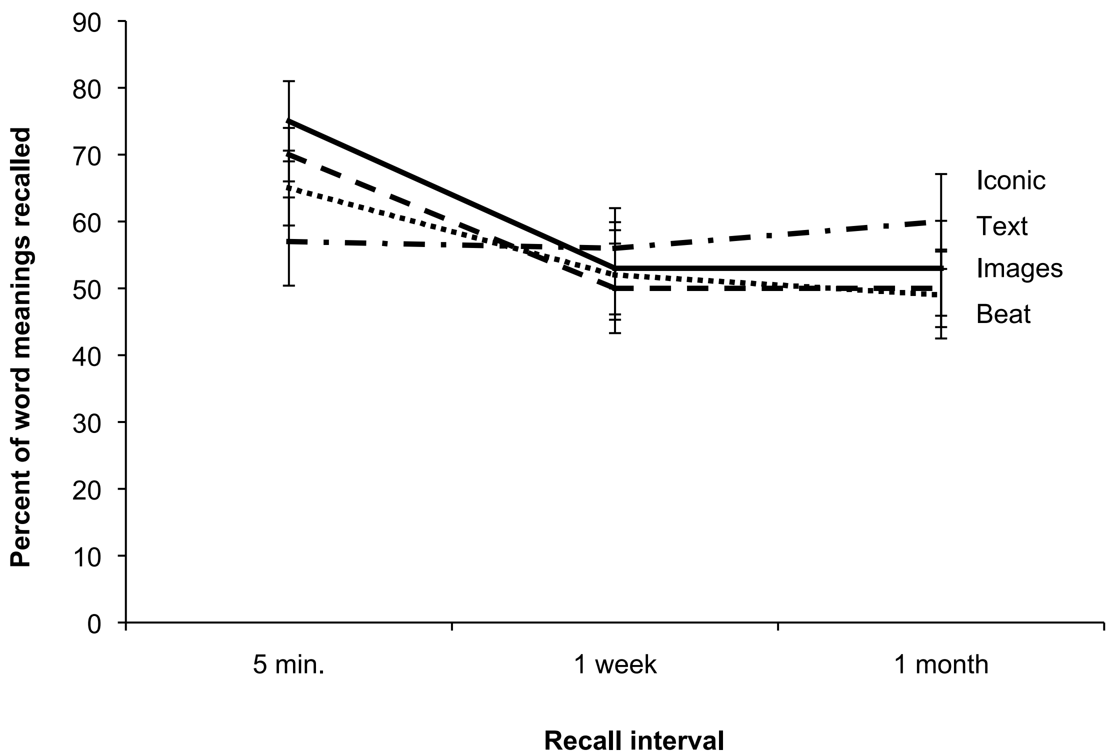


Figure 4. Percent of word meanings recalled by condition and recall interval for Expt. 3 (error bars represent *SEM*).

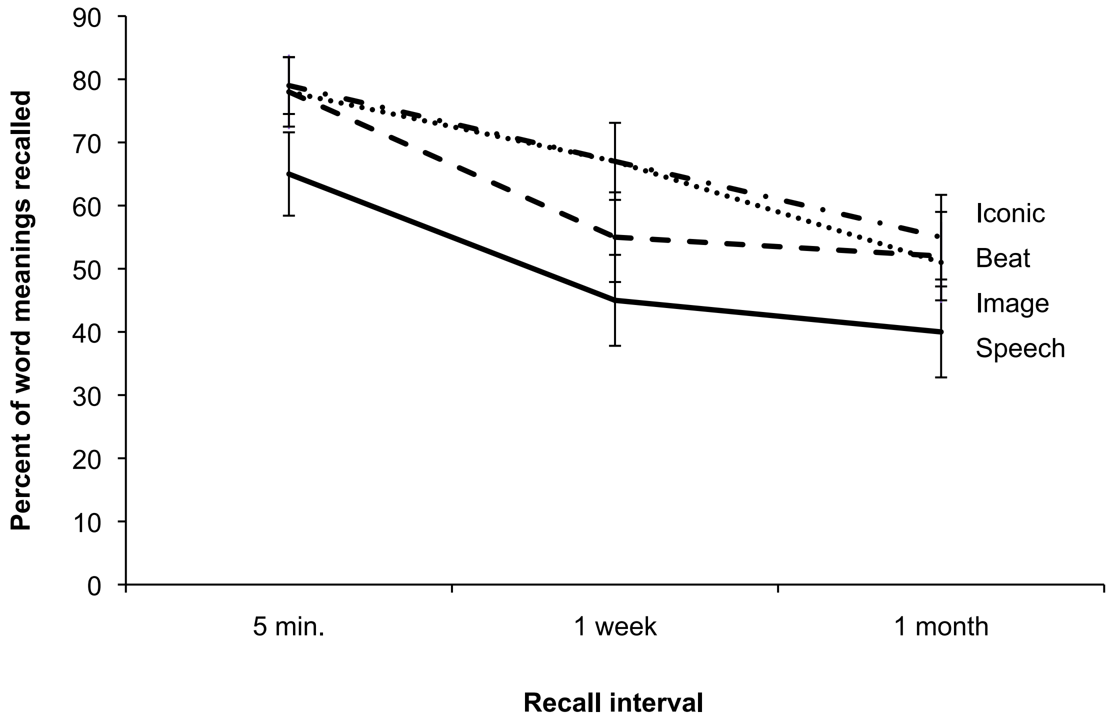


Figure 5. Percent of word meanings recalled by condition and recall interval for Expt. 4 (error bars represent *SEM*).

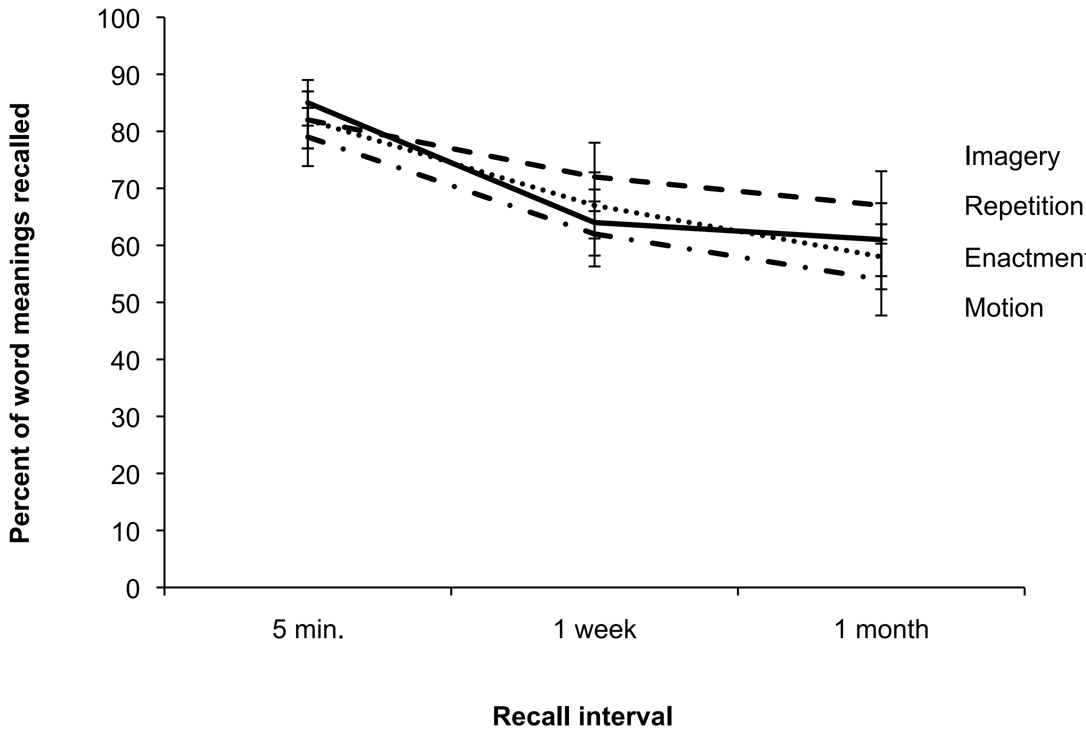
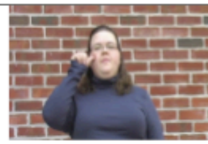
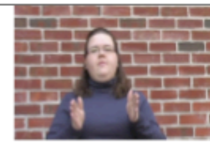
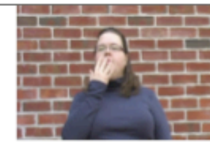
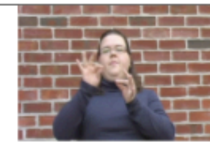
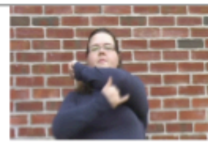
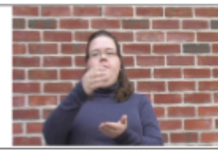
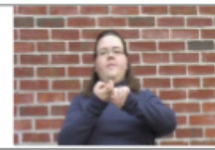
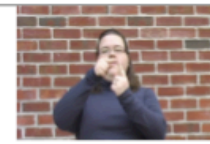
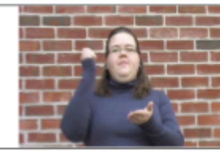
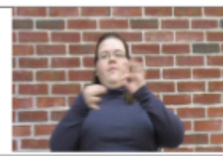
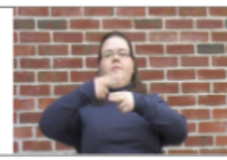
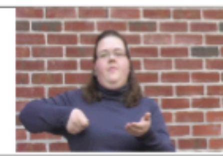
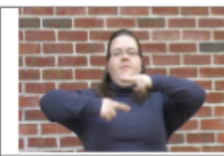


Figure 6. ASL signs used in Study 5, with English glosses, organized by type.

ASL Sign					
English	**Adult**	**To answer**	**Box**	**Color**	**To count**
Sign Type	Arbitrary	Metaphorical	Iconic	Arbitrary	Metaphorical

ASL Sign					
English	**Country**	**Cup**	**Friend**	**Goal (objective)**	**Hammer**
Sign Type	Arbitrary	Iconic	Metaphorical	Metaphorical	Iconic

ASL Sign					
English	**To help**	**Message**	**Name**	**Pool (billiards)**	**To roll**
Sign Type	Metaphorical	Metaphorical	Arbitrary	Iconic	Iconic

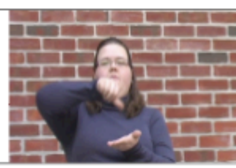
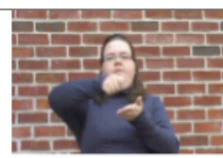
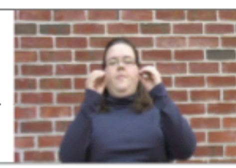
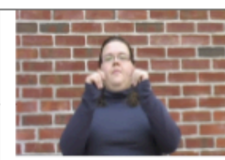
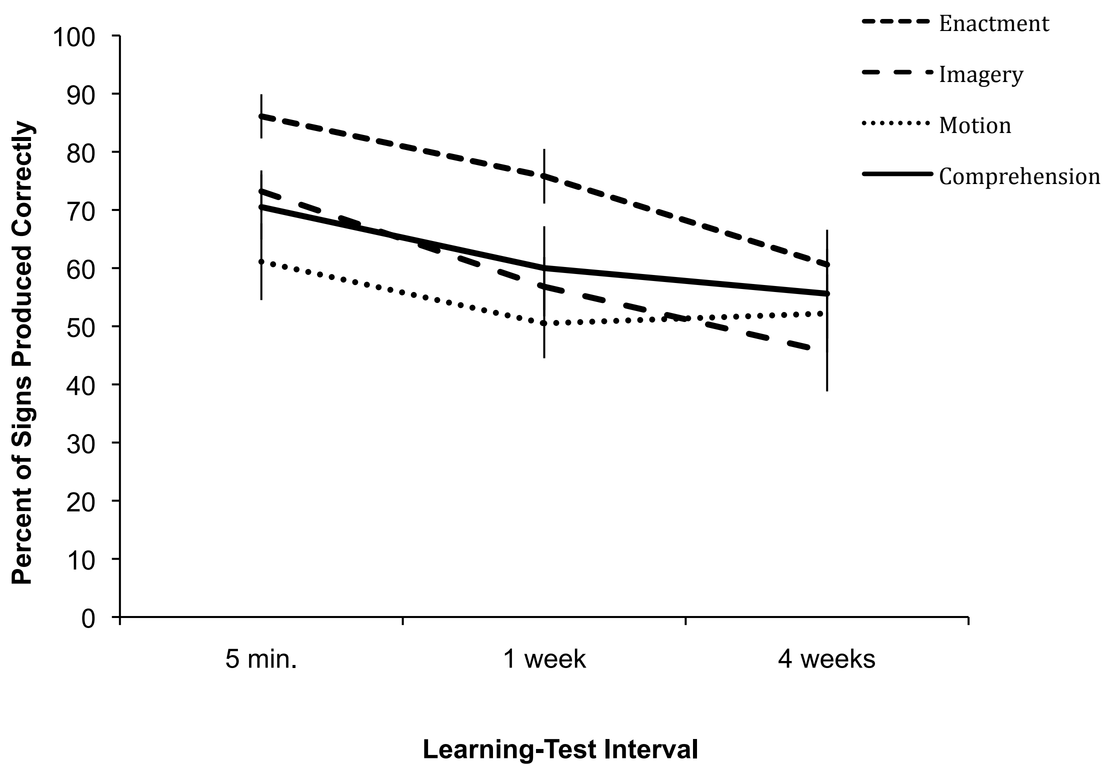
ASL Sign					
English	**Sauce**	**Saw**	**To teach**	**Team**	**To twist**
Sign Type	Iconic	Iconic	Metaphorical	Metaphorical	Iconic

Figure 7. Percent of signs produced correctly by learning condition and recall interval (error bars represent *SEM*).



References

Acredolo, L., & Goodwyn, S. (1988). Symbolic gesturing in normal infants. *Child Development*, *59*(2), 450–466.

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, *44*(2). doi:10.1006/jmla.2000.2752

Allen, L. Q. (1995). The effects of emblematic gestures on the development and access of mental representations of French expressions. *The Modern Language Journal*, *79*(4), 521–529.

Asher, J. J. (1969). The total physical response approach to second language learning. *The Modern Language Journal*, *53*(1), 3–17.

Asher, J. J., Kusudo, J. A., & Torre, R. de la. (1974). Learning a second language through commands: The second field test. *The Modern Language Journal*, *58*(1/2), 24–32.

Asher, J. J., & Price, B. S. (1967). The learning strategy of the Total Physical Response: Some age differences. *Child Development*, *38*, 1219–1227.

Atkinson, R. C. (1975). Mnemotechnics in second-language learning. *American Psychologist*, *30*(8), 821–828. doi:10.1037/h0077029

Atkinson, R. C., & Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *Journal of Experimental Psychology: Human Learning and Memory*, *1*, 126–133. doi:10.1037/0278-7393.1.2.126

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. doi:10.1126/science.1736359

Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General*, *113*(1), 1–29. doi:10.1037/0096-3445.113.1.1

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(04), 577–660. doi:null

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*(1), 617–645. doi:10.1146/annurev.psych.59.103006.093639

Bavelas, J. B., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, *58*(2), 495–520. doi:10.1016/j.jml.2007.02.004

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods*, *33*(1), 73–79. doi:10.3758/BF03195349

Bonvillian, J. D., Orlansky, M. D., & Novack, L. L. (1983). Developmental Milestones: Sign Language Acquisition and Motor Development. *Child Development*, *54*(6), 1435–1445. doi:10.2307/1129806

Bosworth, R. G., & Dobkins, K. R. (2002). Visual Field Asymmetries for Motion Processing in Deaf and Hearing Signers. *Brain and Cognition*, *49*(1), 170–181. doi:10.1006/brcg.2001.1498

Campbell, R., Martin, P., & White, T. (1992). Forced choice recognition of sign in novice learners of British Sign Language. *Applied Linguistics*, *13*(2), 185–201.

Carpenter, S. K., & Olson, K. M. (2011). Are pictures good for learning new vocabulary in a foreign language? Only if you think they are not. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi:10.1037/a0024828

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. The MIT press.

Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *Modern Language Journal*, *80*(2), 183–198.

Cohen, R. L., & Bryant, S. (1991). The role of duration in memory and metamemory of enacted instructions (SPTs). *Psychological research*, *53*(3), 183–187.

Cohen, R. L., & Otterbein, N. (1992). The mnemonic effect of speech gestures: Pantomimic and non-pantomimic gestures compared. *European Journal of Cognitive Psychology*, *4*(2), 113. doi:10.1080/09541449208406246

Cook, S. W., & Tanenhaus, M. K. (2009). Embodied communication: Speakers' gestures affect listeners' actions. *Cognition*, *113*(1), 98–104. doi:10.1016/j.cognition.2009.06.006

Corballis, M. C. (2003). *From hand to mouth: The origins of language*. Princeton University Press.

Dunn, L. M., & Dunn, L. (1997). Peabody picture vocabulary test-III (PPVT-III). *Circle Pines, MN: American Guidance Service*.

Emmorey, K., & Corina, D. (1993). Hemispheric specialization for ASL signs and english words: Differences between imageable and abstract forms. *Neuropsychologia*, *31*(7), 645–653. doi:10.1016/0028-3932(93)90136-N

Emmorey, K., & Kosslyn, S. M. (1996). Enhanced image generation abilities in deaf signers: A right hemisphere effect. *Brain and cognition*, *32*(1), 28–44.

Emmorey, K., Kosslyn, S. M., & Bellugi, U. (1993). Visual imagery and visual-spatial language: Enhanced imagery abilities in deaf and hearing ASL signers. *Cognition*, *46*(2), 139–181. doi:10.1016/0010-0277(93)90017-P

Fiebach, C. J., & Friederici, A. D. (2004). Processing concrete words: fMRI evidence against a specific right-hemisphere involvement. *Neuropsychologia*, *42*(1), 62–70. doi:10.1016/S0028-3932(03)00145-3

Fodor, J. A. (1983). *The Modularity of Mind*. A Bradford Book.

Frick-Horbury, D. (2002). The use of hand gestures as self-generated cues for recall of verbally associated targets. *The American Journal of Psychology*, *115*(1), 1–20. doi:10.2307/1423671

Frick-Horbury, D., & Guttentag, R. E. (1998). The effects of restricting hand gesture production on lexical retrieval and free recall. *The American Journal of Psychology*, *111*(1), 43–62.

Frishberg, N. (1975). Arbitrariness and Iconicity: Historical Change in American Sign Language. *Language*, *51*(3), 696–719. doi:10.2307/412894

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *Language*, *2*(1), 301–334.

Gibbs, R. W. (2006). *Embodiment and cognitive science*. Cambridge University Press.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, *26*(1), 69–83. doi:10.1016/0749-596X(87)90063-5

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

Goldin-Meadow, S., & Beilock, S. L. (2010). Action's influence on thought: The case of gesture. *Perspectives on Psychological Science*, *5*(6), 664–674.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining Math: Gesturing Lightens the Load. *Psychological Science*, *12*(6), 516 –522. doi:10.1111/1467-9280.00395

Hostetter, A. B., & Alibali, M. W. (2007). Raise your hand if youre spatial: Relations between verbal and spatial skills and gesture production. *Gesture*, *7*(1), 73–95.

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

Hostetter, A. B., & Alibali, M. W. (2010). Language, gesture, action! A test of the Gesture as Simulated Action framework. *Journal of Memory and Language*, *63*(2), 245–257. doi:10.1016/j.jml.2010.04.003

Iverson, J. M., & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science*, *16*(5), 367 –371. doi:10.1111/j.0956-7976.2005.01542.x

Iverson, J. M., & Thelen, E. (1999). Hand, mouth and brain. The dynamic emergence of speech and gesture. *Consciousness and Cognition*, *6*, 19–40.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition*, *94*(3), B79–B89. doi:10.1016/j.cognition.2004.06.005

Kelly, S. D., & Lee, A. L. (2011). When actions speak too much louder than words: Hand gestures disrupt word learning when phonetic demands are high. *Language and Cognitive Processes*.

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

Kendon, A. (1980). The relationship of verbal and nonverbal communication. *The relationship of verbal and nonverbal communication*. The Hague: Mouton. Retrieved from http://books.google.com/books?hl=en&lr=&id=Tbg98xnbIYwC&oi=fnd&pg=PA207&dq=gesticulation&ots=ptY2xSiPoH&sig=tqCsM2MZM1nT8yWF5-w3l0BCkck#v=onepage&q=gesticulation&f=false

Kunihira, S., & Asher, J. J. (1965). The stategy of the total physical response: An application to learning Japanese. *IRAL - International Review of Applied Linguistics in Language Teaching*, *3*(4), 277–290. doi:10.1515/iral.1965.3.4.277

Levin, J. R., McCormick, C. B., Miller, G. E., Berry, J. K., & Pressley, M. (1982). Mnemonic versus nonmnemonic vocabulary-learning strategies for children. *American Educational Research Journal*, *19*(1), 121 –136. doi:10.3102/00028312019001121

Ma, W., Golinkoff, R. M., Hirsh-Pasek, K., McDonough, C., & Tardif, T. (2009). Imageability predicts the age of acquisition of verbs in Chinese children. *Journal of child language*, *36*(02), 405–423.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, *270*(5233), 102.

Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, *379*(6566), 649–652.

Matlock, T. (2004). Fictive motion as cognitive simulation. *Memory & Cognition*, *32*(8), 1389–1400.

McDonough, C., Song, L., Hirsh-Pasek, K., Golinkoff, R. M., & Lannon, R. (2010). An image is worth a thousand words: why nouns tend to dominate verbs in early word learning. *Developmental Science*, *14*(2), 181–189. doi:10.1111/j.1467-7687.2010.00968.x

McNeill, D. (1992). *Hand and mind*. University of Chicago Press.

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

Meier, R. P., & Newport, E. L. (1990). Out of the hands of babes: On a possible sign advantage in language acquisition. *Language*, 1–23.

Miller, M. S. (1987). Sign iconicity: Single-sign receptive vocabulary skills of nonsigning hearing preschoolers. *Journal of Communication Disorders*, *20*(5), 359–365. doi:10.1016/0021-9924(87)90024-4

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 519–533. doi:16/S0022-5371(77)80016-9

Nelson, D. G. K., Hirsh-Pasek, K., Jusczyk, P. W., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of child Language*, *16*(1), 55–68.

O'Brien, J. (1999). Metaphoricity in the Signs of American Sign Language. *Metaphor and Symbol*, *14*(3), 159–177. doi:10.1207/S15327868MS140301

Orlansky, M. D., & Bonvillian, J. D. (1984). The Role of Iconicity in Early Sign Language Acquisition. *J Speech Hear Disord*, *49*(3), 287–292.

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford University Press US.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science*, *14*(2), 119–124. doi:10.1111/1467-9280.t01-1-01429

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2004). Sensorimotor simulations underlie conceptual representations: Modality-specific effects of prior activation. *Psychonomic Bulletin & Review*, *11*(1), 164.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

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

Rinaldi, P., Barca, L., & Burani, C. (2004). A database for semantic, grammatical, and frequency properties of the first words acquired by Italian children. *Behavior Research Methods*, *36*(3), 525–530. doi:10.3758/BF03195599

Sabsevitz, D. S., Medler, D. A., Seidenberg, M., & Binder, J. R. (2005). Modulation of the semantic system by word imageability. *NeuroImage*, *27*(1), 188–200. doi:10.1016/j.neuroimage.2005.04.012

Saygin, A. P., McCullough, S., Alac, M., & Emmorey, K. (2009). Modulation of BOLD Response in Motion-sensitive Lateral Temporal Cortex by Real and Fictive Motion Sentences. *Journal of Cognitive Neuroscience*, *22*(11), 2480–2490. doi:10.1162/jocn.2009.21388

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, *305*(5691), 1779.

Snow, C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child development*, 1114–1128.

Solomon, K., & Barsalou, L. (2004). Perceptual simulation in property verification. *Memory & Cognition*, *32*(2), 244–259. doi:10.3758/BF03196856

Spence, C., Nicholls, M. E. R., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, *63*(2), 330–336.

Spivey, M. J. (2007). *The continuity of mind*. Oxford, UK: Oxford University Press.

Stam, G. (2006). Thinking for speaking about motion: L1 and L2 speech and gesture. *International Review of Applied Linguistics in Language Teaching (IRAL)*, *44*(2), 145–171. doi:10.1515/IRAL.2006.006

Steklis, H. D., & Harnad, S. R. (1976). From hand to mouth: Some critical stages in the evolution of language. *Annals of the New York Academy of Sciences*, *280*(1), 445–455.

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*(4), 661–699. doi:10.1111/j.0023-8333.2005.00320.x

Svensson, T., & Nilsson, L. G. (1989). The relationship between recognition and cued recall in memory of enacted and nonenacted information. *Psychological research*, *51*(4), 194–200.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C. C., et al. (2004). A new on-line resource for psycholinguistic studies.

Journal of Memory and Language, *51*(2), 247–250. doi:10.1016/j.jml.2004.03.002

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology*, *32*(3), 492.

Tellier, M. (2005). *How do teacher's gestures help young children in second language acquisition?* Text presented at the International Society of Gesture Studies, Lyon, France. Retrieved from http://gesture-lyon2005.ens-lyon.fr/article.php3?id_article=253

Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture*, *8*, 219–235.

Thompson, R. L., Vinson, D. P., & Vigliocco, G. (2009). The link between form and meaning in American Sign Language: Lexical processing effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 550.

Tomasello, M. (2008). *Origins of human communication* (Vol. 2008). The MIT Press.

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

Vandierendonck, A., Kemps, E., Fastame, M. C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology*, *95*(1), 57–79. doi:10.1348/000712604322779460

Vinson, D. P., Cormier, K., Denmark, T., Schembri, A., & Vigliocco, G. (2008). The British Sign Language (BSL) norms for age of acquisition, familiarity, and iconicity. *Behavior research methods*, *40*(4), 1079–1087.

Yoshioka, K., & Kellerman, E. (2006). Gestural introduction of ground reference in L2 narrative discourse. *IRAL - International Review of Applied Linguistics in Language Teaching*, *44*(2), 173–195. doi:10.1515/IRAL.2006.007