COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 35 (2011) 171–183

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/j.1551-6709.2010.01141.x

Iconic Gestures Prime Words

De-Fu Yap, Wing-Chee So, Ju-Min Melvin Yap, Ying-Quan Tan, Ruo-Li Serene Teoh

Department of Psychology, National University of Singapore

Received 30 August 2009; received in revised form 22 June 2010; accepted 22 June 2010

Abstract

Using a cross-modal semantic priming paradigm, both experiments of the present study investigated the link between the mental representations of iconic gestures and words. Two groups of the participants performed a primed lexical decision task where they had to discriminate between visually presented words and nonwords (e.g., *flirp*). Word targets (e.g., *bird*) were preceded by video clips depicting either semantically related (e.g., pair of hands flapping) or semantically unrelated (e.g., drawing a square with both hands) gestures. The duration of gestures was on average 3,500 ms in Experiment 1 but only 1,000 ms in Experiment 2. Significant priming effects were observed in both experiments, with faster response latencies for related gesture–word pairs than unrelated pairs. These results are consistent with the idea of interactions between the gestural and lexical representational systems, such that mere exposure to iconic gestures facilitates the recognition of semantically related words.

Keywords: Gesture; Semantic priming; Lexical decision; Lexical processing; Semantic representation

1. Introduction

Speakers from all cultural and linguistic backgrounds move their hands and arms when they talk (Feyereisen & de Lannoy, 1991). For example, a speaker opens her palm and moves her hand forward, while saying *I gave him a present*. Such hand and arm movements are collectively referred to as a gesture (McNeill, 1992). Even though there are various linguistic and structural differences between gestures and speech, previous findings generally support the idea that gestures and speech are tightly integrated temporally, semantically,

Correspondence should be sent to So Wing Chee, Department of Psychology, National University of Singapore, BLK AS4, 9 Arts Link, Singapore 117572. E-mail: psyswc@nus.edu.sg

and pragmatically (Kita, 2000; McNeill, 1992; So, Kita, & Goldin-Meadow, 2009). Temporally, Morrel-Samuels and Krauss (1992) have shown that gestures are initiated either just before or simultaneously with their lexical affiliate (i.e., the word whose retrieval the gesture facilitates), indicating that gestural and language systems are linked during communication (see also Mayberry & Jaques, 2000). Semantically, the meaning of co-expressed gestures goes hand in hand with the meaning of accompanying speech (e.g., Kita & Ozyurek, 2003; Ozyurek, Kita, Allen, Furman, & Brown, 2005). Pragmatically, gesture packages preverbal spatio-motoric messages into units that are suitable for speaking (Kita, 2000).

This research will further explore the integrated link between gestures and language by examining the interplay between gestures and the lexical processing system. Specifically, we examine whether the presentation of a gesture (e.g., a two-hands flapping gesture) would activate a semantically related word (e.g., ''bird'' or ''flying''). Although there are many different types of gestures, we focus on *iconic* gestures, because they express meaning that is related to the semantic content of the speech they accompany (Krauss, Chen, & Chawla, 1996; McNeill, 2005) and resemble the physical properties and movement of objects or actions being described in speech (McNeill, 2005). We are primarily interested in whether the presentation of an iconic gesture (e.g., a pair of hands flapping) would prime a semantically related word (e.g., *bird* or *flying*), using a cross-modal semantic priming paradigm.

Semantic priming refers to the facilitation in the cognitive processing of information after recent exposure to related information (Neely, 1991), and it is most often studied using primed visual word recognition paradigms (see McNamara, 2005; Neely, 1991, for excellent reviews). In these tasks, the participants are presented with a context word (e.g., *cat*) that is followed by a target letter string which is either semantically related (e.g., *dog*) or unrelated (e.g., *table*). In the primed speeded naming task, participants read the target letter string aloud, whereas in the primed lexical decision task, they have to decide via a button press whether the letter string forms a real word or nonword (e.g., *flirp*). In general, words preceded by semantically related primes are responded to faster than words preceded by semantically unrelated primes. This effect, which is extremely robust, is known as the *semantic priming effect*.

In our experiments, we use a cross-modal priming paradigm where context words are replaced with gestures, with the aim of testing the hypothesis that gestures prime semantically related words. There are strong theoretical motivations for this hypothesis. For example, Krauss (1998) proposed that our memorial representations are encoded in multiple formats or levels, including gestures and their lexical affiliates. McNeill (1985) further argued that gesture and language should be viewed as a *single* system within a unified conceptual framework. In other words, language and gestures, despite being represented with different formats,[1] can be seen as elements of a single tightly integrated process.

Interestingly, there is already some provocative evidence in the literature that gestures do prime words. To our knowledge, Bernardis et al. represents the first study demonstrating gesture–word priming effects. In their study, participants were visually presented with gestural primes, followed by either related or unrelated lexical targets, and they were then asked to name the lexical targets. On average, each gesture clip lasted for 3,672 ms (ranging from 2,320 to 4,680 ms). Bernardis et al. then measured the time taken to name each lexical target and compared this against a ''neutral'' baseline latency, which was estimated using a *separate* group of participants who simply had to name aloud the lexical targets (i.e., gestural primes were not presented). The use of a baseline allows researchers to estimate *facilitation* (faster latencies for related compared with neutral condition) and *inhibition* (slower latencies for unrelated compared with neutral condition) effects, which are supposed to respectively map onto the fast automatic and slower expectancy-based aspects of priming (see Neely, 1977; Posner & Snyder, 1975). Bernardis and colleagues reported a significant inhibition effect, that is, participants took longer to name a word when it was unrelated to the gesture, compared with the neutral baseline. Intriguingly, they did *not* find a facilitation effect, that is, there was no significant difference between the related and neutral conditions. On the basis of the nonreliable facilitation effect, they concluded that ''the meaning of iconic gestures did not prime the same-meaning words'' (p. 1125). However, when the participants were instructed to read the target words *and* form mental images of those words, a facilitation effect was found (Bernardis & Caramelli, 2009).[2]

Yet the conclusion drawn by Bernardis, Salillas, and Caramelli (2008) should be interpreted with some caution for *two* main reasons. First, each gesture was displayed for more than 3,000 ms. This might allow sufficient time for the participants to actually name or label the gestures. Naming of gestures would, in turn, activate the semantically related words that might serve as lexical primes. Henceforth, the facilitation effect in the naming task could be attributed to the lexical primes instead of the gesture primes. In order to discourage participants from naming the gestures, the duration of each gesture clip should be substantially shorter than 3,000 ms.

Second, Bernadis et al.'s conclusion rests exclusively on the nonsignificant facilitation effect. As discussed earlier, the magnitude of the facilitation effect reflects the difference between the related and neutral conditions, and it is therefore modulated by the specific neutral baseline selected. While the neutral baseline putatively allows researchers to disentangle the automatic and controlled influences of priming, there is evidence that the use of neutral baselines is associated with a number of problems. Most important, Jonides and Mack (1984) have convincingly demonstrated that neutral baselines can artifactually overestimate *or* underestimate facilitation and inhibition effects, depending on the specific neutral prime (e.g., BLANK or XXXXX) selected. Related to this, Forster (1981) pointed out that evidence for small (or null) facilitation effects is necessarily ambiguous, because facilitation effects can be underestimated by one's choice for the neutral condition.

One might contend that Bernardis et al. have neatly circumvented the problem of using a neutral prime by estimating the neutral condition from an independent sample that saw *only* the target words (without the context of the gesture primes). However, in order to validly estimate facilitation and inhibition, it is necessary to insure that the neutral and cuing conditions are ''*identical* with respect to *all* processing consequences of the cue except the specific preparatory effect elicited by the informative cue'' (Jonides & Mack, 1984, p. 33). Clearly, the processing demands associated with naming isolated words versus naming words preceded by a gesture are not identical. For instance, it is likely that there are differences in *target* encoding times for the neutral and cuing conditions. Specifically, primes are informative and take time to process. To the extent that related primes are not *fully* encoded by the time the target appears, this adds additional processing time to target processing in related trials. Given that this additional processing is not relevant to trials in the neutral condition, this leads to an underestimation of facilitation effects.

Given that finding an appropriate neutral baseline (i.e., one that is equated on all dimensions with the prime conditions) is virtually impossible, Jonides and Mack (1984) strongly recommended that the neutral condition should be excluded. Importantly, they also pointed out that if a researcher is primarily interested in whether a related condition facilitates performance, it is sufficient to examine the *overall* priming effect (Neely, 1977), defined as the difference between the unrelated and related conditions. Interestingly, Bernardis et al. (2008) *did* find a significant overall priming effect. Specifically, targets preceded by a related gesture, compared with targets preceded by an unrelated gesture, were named 39 ms faster, $p < .001$ (p. 1118). In summary, our reexamination of the Bernardis et al.'s (2008) results indicates that contrary to their claim, their basic findings are actually *consistent* with the idea that iconic gestures prime semantically related words.

Our present study aimed to reexamine the cross-modal gesture–word priming effect, addressing the two major limitations in Bernardis et al.'s (2008) study as discussed above. We conducted two experiments and investigated whether the gestures would prime the responses to semantically related words. In both experiments, we measured the overall priming effect. In Experiment 1, the average duration of gesture clip was 3,500 ms to replicate the overall priming effect found in the Bernardis et al.'s study. In Experiment 2, the duration of each gesture clip was shortened to 1,000 ms in order to minimize the likelihood that participants strategically recode the gestural primes into verbal labels and use the latter to prime the lexical targets. Different from Bernardis et al.'s study, we used a different word recognition paradigm-primed lexical decision task. Doing this is critical because different word recognition tasks produce effects specific to that task (see Balota & Chumbley, 1984; Yap & Balota, 2007), and it is important to establish effects across different tasks (Grainger & Jacobs, 1996). In addition, we used a different set of gesture primes and targets for this study than Bernardis et al. in both of our experiments, which will help provide converging evidence for the generality of the effect.

2. Experiment 1

2.1. Method

2.1.1. Participants

Sixty-three undergraduates from the Research Participant Program at the National University of Singapore participated in this experiment for course credit. They were all native English speakers with normal or corrected-to-normal vision.

2.1.2. Materials

2.1.2.1. Gesture primes: To select the gestures that conveyed meaningful semantic information for this study and pair them up with semantically related and nonrelated words, a separate group of 45 English-speaking undergraduates from the National University of

Singapore were presented with 80 silent videotaped gestures, each lasting for 3–4 s on a computer screen in a speechless context. They were given 7 s[3] to write a single word that best described the meaning of a gesture. As speech was not available, the participants derived meaning from gestures according to their physical forms and movements.

We determined that a gesture had a common interpretation if its meaning was agreed upon by 70% of the participants (see also Goh, Suárez, Yap, & Tan, 2009). Our findings showed that 40 of the gestures had consistency rates of 70% or above. Reliability was assessed by a second coder who analyzed all the responses and calculated the consistency rate of each gesture. These gestures served as gesture primes in our study. The interrater reliability was .946. See Appendix for a list of gestures and their associated meanings given by the participants.

2.1.2.2. Lexical targets: The lexical targets consisted of words and nonwords. In terms of words, we used the gesture meanings derived from the participants and matched the meaning with the words listed in Nelson, McEvoy, and Schreiber's (1998) free association norms. We then selected the strongest associate in the Nelson et al.'s norms for each gesture. For example, consider the following gesture, whereby the index and middle fingers of both hands form versus above the head, and the fingers are flexing and unflexing. Most participants classified this gesture as *rabbit*. The strongest associate of *rabbit* is *bunny*, and we thus selected *bunny* to be the lexical target for the *rabbit* gestural prime. The Appendix provides the prime–target associative strength for all our stimuli ($M = 0.27$). The nonwords were matched to the semantically related words from the Nelson et al. (1998) norms with respect to (a) word length, (b) number of orthographic neighbors, and (c) number of syllables according to the ARC Nonword Database (Rastle, Harrington, & Coltheart, 2002) and the English Lexicon Project[4] (Balota et al., 2007). Please see Table 1 for summary statistics for the lexical targets across the three conditions (see Appendix for the list of nonwords).

Each participant was presented with 10 related and 10 unrelated prime–target pairs, and 20 nonwords that were preceded by gesture primes. The primes in the three conditions (related, unrelated, and nonword) were counterbalanced across participants such that each prime had an equal chance of appearing in each of the three conditions. The order of presentation was randomized anew for each participant.

2.1.3. Procedure

The experiment was run using E-PRIME 2.0 (Schneider, Eschmann, & Zuccolotto, 2002). The gesture video clips and words were presented one at a time at the center of the computer

Table 1
Stimulus characteristics of the words used in Experiment 1

Factor	Minimum	Maximum	*M* (*SD*)
Length	3	8	4.65 (1.42)
Orthographic neighbors	0	20	6.65 (5.89)
Syllables	1	3	1.23 (0.480)

screen against a white background. Words were presented in black lowercase letters. Each trial began with a fixation stimulus (+) appearing in the center of the screen for 1,000 ms followed by a blank screen for 200 ms. After the blank screen, the gesture prime video clip appeared for 3,000–4,000 ms, which was followed by another blank screen for 200 ms. The blank screen was replaced by the lexical target, which remained on the screen for 3,000 ms or until the participants indicated on the keyboard their lexical decision.

The participants were seated in front of a 17-inch monitor of a computer and instructed to make a lexical decision (i.e., ''/'' for a word and ''z'' for a nonword) on the keyboard as quickly and accurately as possible. They received 10 practice trials prior to the experiment. A blank screen that served also as intertrial interval followed a correct response for 1,000 ms. An ''Incorrect'' display would be presented for 1,000 ms above the fixation point for incorrect responses. Both the accuracy rates and response latencies to the hundredth millisecond were recorded.

2.2. Results and discussion

Accuracy was almost at ceiling across all three conditions. However, a few of the participants had noticeably low accuracy rates in some conditions. A box-plot analysis based on mean accuracy rates collapsed across all three conditions was conducted and five participants who had mean accuracy rates below 2 *SD*s of the sample mean were removed. Errors (3.1% across all three conditions) and response latencies faster than 200 ms or slower than 3,000 ms were excluded from the analyses. Response latencies more than 2.5 *SD*s above or below each participant in each condition were also excluded from the analyses. These criteria removed a further 1.2% of the responses.

The dependent variables examined were the response latency as well as the accuracy rates. Participants responded faster and more accurately in the related condition, mean reaction time = 641 ms (*SD* = 143 ms) and mean accuracy = 99% (*SD* = 3.1%), when compared with the unrelated condition, mean reaction time = 670 ms (*SD* = 179 ms) and mean accuracy = 96% (*SD* = 6.7%). The mean response latency in the related condition was significantly faster than that in the unrelated condition (i.e., 29 ms); the effect was significant by participants, $t_p(57) = 2.16$, $p = .035$, Cohen's $d = .18$, and by items, $t_i(39) = 2.10$, $p = .042$, Cohen's $d = .26$. The effect in accuracy (3%) was also significant by participants, $t_p(57) = 3.47$, $p = .001$, Cohen's $d = .57$, and by items, $t_i(39) = 3.56$, $p = .001$, Cohen's $d = .84$.

Overall, we found that iconic gestures facilitated the recognition of a semantically related word such that the words preceded by semantically related gestures were responded to faster than words preceded by semantically unrelated gestures. However, the duration of each gesture clip in our study was unnecessarily long (on average 3,500 ms), leading to long stimulus onset asynchronies (SOAs). Hence, one might contend that participants might have had sufficient time to ''label'' each gesture prime in their mind, and such unspoken label might activate the lexical target. In other words, the iconic gesture activates its lexical referent, and the lexical referent, in turn, activates or facilitates the retrieval of the semantically related lexical target. The obvious solution is to shorten the SOAs to minimize strategic influences. Thus, in Experiment 2, the SOA was reduced by shortening the duration of the gesture clip from 3,500 to 1,000 ms.

3. Experiment 2

3.1. Methodology

3.1.1. Participants

Seventy-six undergraduates from the Research Participant Program at the National University of Singapore participated in this experiment for course credit. They were all native English speakers with normal or corrected-to-normal vision.

3.1.2. Materials

3.1.2.1. Gesture primes: We used the same gesture clips in Experiment 1. However, we truncated the gesture clips such that each clip lasted for only 1,000 ms. According to McNeill (1992, 2005), the production of a gesture undergoes at least three phases: preparation, stroke, and retraction. Consider the gesture for *bird*. When producing this gesture, a speaker raises both hands in the preparation phase, flaps both hands in the stroke phase, and relaxes both hands in the retraction phrase. The stroke carries the imagistic content that conveys the semantic meaning of a gesture (in this example, *bird*). Therefore, when shortening the duration of each gesture, we retained the stroke phases only. For each of the 80 gestures, the stroke phase lasted for 1,000 ms.

In order to examine whether the semantic meanings of gestures were preserved after cropping, we presented the 80 gestures to a separate group of 37 English-speaking undergraduates from the National University of Singapore and asked them to describe their meanings in one word. Specifically, we explored whether the 40 gestures used in Experiment 1 elicited the same high consistency rates when we shortened the duration from 3,000 to 1,000 ms.

Same as Experiment 1, we determined that a gesture had a common interpretation if its meaning was agreed upon by 70% of the participants. Our findings showed that 42 of the gestures had consistency rates of 70% or above. Reliability was assessed by a second coder who analyzed all the responses and calculated the consistency rate of each gesture. The interrater reliability was .932. Of the 42 gestures, all the 40 gestures that had common interpretation in Experiment 1 obtained consistency rates of 70% or above in Experiment 2. More important, they all had the *same* interpretation in both Experiment 1 and Experiment 2.[5] These 40 gestures served as gesture primes in Experiment 2.

3.1.2.2. Lexical targets: Since the shortened gesture primes were interpreted the same way, we used the same set of lexical targets (words and nonwords) presented in Experiment 1. As before, each participant was presented with 10 related and 10 unrelated prime–target pairs, and 20 nonwords that were preceded by gesture primes in Experiment 2. The primes in the three conditions (related, unrelated, and nonword) were counterbalanced across participants such that each prime had an equal chance of appearing in each of the three conditions. The order of presentation was randomized anew for each participant.

3.1.3. Procedure

The procedure was the same as in Experiment 1.

3.2. Results and discussion

As in Experiment 1, accuracy was almost at ceiling across all three conditions. However, a few participants had noticeably low accuracy rates in some conditions. A box-plot analysis based on mean accuracy rates collapsed across all three conditions was conducted and two participants who had mean accuracy rates below 2 *SD*s of the sample mean were removed. Errors (4.93% across all three conditions) and response latencies faster than 200 ms or slower than 3,000 ms were excluded from the analyses. Response latencies more than 2.5 *SD*s above or below each participant in each condition were also excluded from the analyses. These criteria removed a further 1.75% of the responses.

Participants responded faster and more accurately in the related condition, mean reaction time = 725 ms (*SD* = 181 ms) and mean accuracy = 98% (*SD* = 4.9%), as compared to the unrelated condition, mean reaction time = 757 ms (*SD* = 187 ms) and mean accuracy = 96% (*SD* = 7.1%). The mean response latency in the related condition was significantly faster than that in the unrelated condition (i.e., 32 ms); the effect was significant by participants, $t_p(73) = 2.42$, $p = .018$, Cohen's $d = .17$, but not by items, $t(39) = 1.21$, $p = .30$, Cohen's $d = .27$. The effect in accuracy (2%) was also significant by participants, $t(73) = 2.28$, $p = .026$, Cohen's $d = .32$, but not by items, $t_i(39) = 1.74$, $p = .09$, Cohen's $d = .68$.

Unlike Experiment 1, the priming effect by items was not significant for reaction times in Experiment 2.[6] To check whether the nonsignificant item effect in Experiment 2 was due to violations of nonparametric assumptions, we also conducted nonparametric sign tests based on the number of items showing priming in Experiment 1 and Experiment 2. The nonparametric sign tests (one-tailed) yielded the same results as the parametric tests. Specifically, 28 of 40 items showed priming in Experiment 1 ($p = .008$) while only 25 of 40 items in Experiment 2 showed priming ($p = .07$).[7] Thus, fewer items showed priming in Experiment 2, compared with Experiment 1. The results suggest that, even though the priming effect was reliable by participants in both experiments, priming was actually stronger in Experiment 1 than in Experiment 2. This raises the possibility that the priming effect in Experiment 1 was partially mediated by some type of strategic verbal recoding. When that strategy was minimized in Experiment 2 due to the shorter SOA, priming was consequently less strong.

4. General discussion

In both experiments, we demonstrated that iconic gestures prime responses to semantically related word targets, using a lexical decision task. We found the semantic link between iconic gestures and words that are semantically related to them. Our findings are inconsistent with the *claims* by Bernardis et al. (2008), but they are compatible with their empirical *findings*. Collectively, the two studies indicate that the cross-modal gesture–word priming effect is robust and generalizes to different lexical processing tasks (i.e., lexical decision and speeded naming).

Our findings indicate that gestures and words share a tight link in our mental representation. For example, a fist moving up and down near the mouth can refer to the action of brushing teeth. Likewise, two hands with V-shaped index and middle fingers above the head refer to *bunny*. Evidence for a gesture–word semantic link is also found in embodied cognition research. Based on the theory of embodied cognition, knowledge can be represented by reenactments of sensory, motor, and introspective states (Bower, 1981; Fodor, 1975; Fodor & Pylyshyn, 1988; Potter, 1979). As suggested by McNeill (1992, p.155), gestures are representations of thought or so-called material carriers of thinking. Recently, there is empirical evidence supporting the notion that gesture as a form of reenactments that embodies mental representation, such that production of a particular gesture would activate congruent mental representation (Casasanto & Dijkastra, 2010; Dijkstra, Kaschak, & Zwaan, 2007; Lakoff & Johnson, 1999). In Casasanto and Dijkastra's (2010) study, they asked a group of participants to retrieve positive and negative autobiographic memory while moving their hands upwards and downwards. Interestingly, the activation of positive autobiographical memory was faster when the participants moved their hands upwards than when they moved their hands downwards. Likewise, the activation of negative memory was faster when hands moved downwards than when they moved upwards. Their findings suggested that positive memory is embodied with upward hand movement while negative memory with downward hand movement.

Given the semantic link between iconic gestures and words, how might one explain the cross-modal semantic priming of words from iconic gestures? At this point, there is no clear model that provides a mechanism for cross-modal priming from gestures to words. Based on the spreading activation framework (Collins & Loftus, 1975), a canonical model of semantic priming, semantic memory is made up of a network of interconnected internal representations or nodes. Activation spreads from a node or concept to semantically related concepts, and prior activation of concepts facilitates their subsequent retrieval. However, gestures are not represented in the standard spreading activation network. Interestingly, Krauss (1998) has proposed that our semantic representation might contain multiple formats, such as gestures and words. We suggest that Krauss' ideas can be incorporated into an embellished spreading activation framework, such that nodes or concepts in the semantic network are represented in the forms of both gestures and their lexical affiliates. There is empirical evidence showing that an activation of a concept in one format (e.g., gesture) would activate related concepts in another format (e.g., words) during the process of speech production (Morrel-Samuels & Krauss, 1992). Along with this line of research, we found that mere exposure of iconic gesture would therefore activate semantically related words.

However, there are three limitations in our study. First, due to various constraints, the pool of useable gesture primes was very limited. The modest number of stimuli might have lowered the statistical power of our item analyses. Second, our gestures vary in their iconicity that might modulate their ability to prime words. Some iconic gestures possess a more direct and transparent relationship with their meaning (Beattie & Shovelton, 2002). For example, the *drive* gesture (two fists alternately moving up and down) is more iconic than the *saw* gesture (open palm moving repeatedly back and forth). The gestures with a higher degree of iconicity should be easier to recognize. Due to the limited number of gestures stimuli in our study, we are unable to explore the influence of iconicity on gestural priming. However, this is an interesting issue that merits further investigation.

Third, even though we have shortened SOA to 1,000 ms, we still cannot entirely rule out the possibility that the participants strategically recoded the gesture primes to lexical primes, which in turn activated the lexical target. Although our findings demonstrated that priming was reliable in both experiments, the participants also responded faster when the duration of gesture was on average 3,500 ms than when it was 1,000 ms. Why did participants take longer to respond when the SOA was 1,000 ms? One possibility is that with a shorter SOA, the participants might not have fully processed the gesture before the word target was presented. Thus, additional gesture processing might have taken place, delaying their lexical decision. When SOA was longer (e.g., 3,500 ms), the participants might have already fully processed the gestures before lexical targets were presented. It is also worth noting that priming was weaker (but still reliable) in the short SOA condition, suggesting that the priming seen in Experiment 1 may be partially mediated by strategic verbal recoding of the gestures.

To conclude, our study used a cross-modal semantic priming paradigm to show a tight semantic link between iconic gestures and words. Specifically, the mere exposure of iconic gestures activates semantically related words.

Notes

1. Although the former is linear-segmented, the latter is global-synthetic and instantaneous imagery.
2. According to Bernardis and Caramelli (2009), forming a mental image of word would activate visuospatial information that was also represented in the semantically related gesture, hence producing a facilitation effect.
3. We had conducted tests on 10 pilot participants and realized that 7 s was the optimal time for a participant to quickly process a gesture and describe it in a single word.
4. The English Lexicon Database was used to match nonwords to words that have more than one syllable.
5. Two gestures that were considered as ''meaningless'' (i.e., their meaning was agreed upon by <70% of the participants) in Study 1 were considered as meaningful in Study 2. However, our aim of Study 2 was to confirm our findings in Study 1. Thus, we did not include these two gestures in Study 2.
6. Note that our priming effect might differ from that found in Bernardis and Caramelli (2009). The participants formed mental images of target words in their study, but they were not instructed to do so in our study.

7. Note that 12 items showed the priming effect in Experiment 1 only and 9 items showed the priming effect in Experiment 2 only.

Acknowledgments

The research described in this manuscript was supported by R581000074133 to WCS at National University of Singapore. We gratefully acknowledge Dr. Jeff Zacks, Dr. Jen Coane, Dr. Paolo Bernadis, Dr. Niccoletta Caramelli, Dr. Al Au, Dr. Tan Seok Hui, and an anonymous reviewer for their valuable inputs and thoughtful comments on a previous version of this work.

References

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 340–357.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*, 445–459.

Beattie, G., & Shovelton, H. (2002). What properties of talk are associated with the generation of spontaneous iconic hand gestures? *British Journal of Social Psychology*, *41*, 403–417.

Bernardis, P., & Caramelli, N. (2009). Meaning in words, gestures, and mental images. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st annual conference of the Cognitive Science Society* (pp. 1693–1697). Austin, TX: Cognitive Science Society.

Bernardis, P., Salillas, E., & Caramelli, N. (2008). Behavioural and neuropsychological evidence of semantic interaction between iconic gestures and words. *Cognitive Neuropsychology*, *25*, 1114–1128.

Bower, G. H. (1981). Mood and memory. *American Psychologist*, *36*, 129–148.

Casasanto, D., & Dijkastra, K. (2010). Motor action and emotional memory. *Cognition*, *115*, 179–185.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

Dijkstra, K., Kaschak, M. P., & Zwaan, R. A. (2007). Body posture facilitates retrieval of autobiographical memories. *Cognition*, *102*, 139–149.

Feyereisen, P., & de Lannoy, J. D. (1991). *Gestures and speech: Psychological investigations*. New York: Cambridge University Press.

Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3–71.

Forster, K. I. (1981). Priming and the effects of sentence and lexical contexts on naming time: Evidence for autonomous lexical processing. *Quarterly Journal of Experimental Psychology*, *33*, 465–495.

Goh, W. D., Suárez, L., Yap, M. J., & Tan, S. H. (2009). Distributional analyses in auditory lexical decision: Neighbourhood density and word-frequency effects. *Psychonomic Bulletin & Review*, *16*, 882–887.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *103*, 518–565.

Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin*, *96*, 29–44.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge, England: Cambridge University Press.

Kita, S., & Ozyurek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal? Evidence for and interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*, 16–32.

Krauss, R. M. (1998). Why do we gesture when we speak? *Current Directions in Psychological Science*, *7*, 54–60.

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 389–450). San Diego: Academic Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. Chicago: University of ChicagoPress.

Mayberry, R., & Jaques, J. (2000). Gesture production during stuttered speech: Insights into the nature of gesture-speech integration. In D. McNeill (Ed.) *Language and gesture* (pp. 199–214). Cambridge, England: Cambridge University Press.

McNamara, T. (2005). *Semantic priming: Perspectives from memory and word recognition*. New York: Psychological Press.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, *92*, 350–371.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

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

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *18*, 615–623.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*, 226–254.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 236–264). Hillsdale, NJ: Erlbaum.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. Available at: http://www.usf.edu/FreeAssociation/. Accessed on September 17, 2010.

Ozyurek, A., Kita, S., Allen, S., Furman, R., & Brown, A. (2005). How does linguistic framing of events influence co-speech gestures? *Gesture*, *5*, 219–240.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Potter, M. C. (1979). Mundane symbolism: The relations among objects, names, and ideas. In N. Smith & M. Franklin (Eds.), *Symbolic functioning in childhood* (pp. 41–65). Hillsdale, NJ: Erlbaum.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology*, *55A*, 1339–1362.

Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools, Inc.

So, W. C., Kita, S., & Goldin-Meadow, S. (2009). Using the hands to identify who does what to whom: Speech and gesture go hand-in-hand. *Cognitive Science*, *33*, 115–125.

Yap, M. J., & Balota, D. A. (2007). Additive and interactive effects on response time distributions in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *33*, 274–296.

Appendix

Gesture Meaning	Lexical Targets (Nelson et al., 1998)	Prime–Target Associative Strength (Nelson et al.,1998)	Nonwords (ARC; ELP)
baby	child	0.2	youns
ball	round	0.15	sorms
brush	hair	0.2	salm
cap	hat	0.06	cug
carry	hold	0.02	lown
circle	square[a]	0.47	blurds
climb	mountain	0.09	sprounge
cold	hot	0.67	tib
comb	brush	0.16	vonks
curve	straight	0.08	phlieves
cut	blood	0.02	skent
cycle	bike	—[b]	mout
down	up	0.84	vu
drive	car	0.12	dar
fat	skinny	0.41	aggous
fill	empty[a]	—[b]	chuzz
fly	bird	0.32	vuch
full	empty[a]	0.61	slolt
grab	take	0.03	jole
guitar	string	0.06	terked
hot	cold	0.41	boke
open	close	0.44	nubes
photo	picture	0.06	iddings
pray	God	—[b]	gos
push	pull	0.59	pome
rabbit	bunny	0.73	trief
read	book	0.42	teap
rectangle	square[a]	—[b]	blurds
run	walk	0.46	joth
slap	hit	0.02	cug
stir	mix	0.19	zix
swim	water	0.05	aiped
think	brain	0.23	polfs
throw	ball	0.07	sogs
tie	neck	0.07	tesh
tiny	small	0.08	crosh
triangle	square[a]	0.04	quescs
up	down	0.58	datt
walk	run	0.49	ven
Write	Read	0.33	dunt

[a]The repeated stimuli were not presented in the same block within participants after counterbalancing.

[b]No values were reported for these stimuli because they were selected to fit the additional cues or information conveyed by the gestures and were therefore not found in Nelson et al.'s (1998) norms.