

## **The Role of Gestures in Spatial Working Memory and Speech**

**Ezequiel Morsella and Robert M. Krauss**

*Columbia University*

### **Abstract**

Co-speech gestures traditionally have been considered communicative, but they may also serve other functions. For example, hand-arm movements seem to facilitate both spatial working memory and speech production. It has been proposed that gestures facilitate speech indirectly by sustaining spatial representations in working memory. Alternatively, gestures may affect speech production directly, by activating embodied semantic representations involved in lexical search. Consistent with the first hypothesis, we found participants gestured more when describing visual objects from memory and when describing objects that were difficult to remember and encode verbally. However, they also gestured when describing a visually-accessible object, and gesture-restriction produced dysfluent speech even when spatial memory was untaxed, suggesting that gestures can directly affect both spatial memory and lexical retrieval.

This is a prepublication copy of a paper that has been accepted for publication in *The American Journal of Psychology*

## The Role of Gestures in Spatial Working Memory and Speech

Traditionally, the hand-arm movements that often accompany speech have been ascribed a communicative role. “As the tongue speaketh to the ear, so the gesture speaketh to the eye” is the way the 18th century naturalist Sir Francis Bacon put it (Bacon, 1891). According to Sapir, gestures constitute “...an elaborate and secret code that is written nowhere, known to none, and understood by all” (Sapir, 1949, p. 556). More recently, however, researchers have begun to explore other functions such motor activity might serve.

For example, people often make a variety of movements as they perform cognitive tasks. An attempt to imagine the shape of an inverted “S” might be accompanied by averted gaze, furrowed brow, and gesticulation. Movements have been found to accompany mental arithmetic (e.g., Graham, 1999), memory processes (e.g., Glenberg, Schroeder, & Robertson, 1998), and visual imagery (e.g., Spivey & Geng, 2001). Changes in gaze direction, sweeping movements of the arms, and an elaborate medley of hand and finger movements frequently occur in conversation (e.g., Beattie, 1980; Krauss, Chen, & Chawla, 1996).

Why do people perform these actions? One possibility is that movements play a functional role in cognition, a view that has a long history in psychology. Nearly 75 years ago, Washburn (1928) contended that “...the motor innervations underlying the consciousness of effort are not mere accompaniments of directed thought, but an essential cause of directed thought” (p. 105). Forty years later, Hebb (1968) espoused a similar notion about the role of eye movements during mental imagery. Although the evidence is far from definitive, there is evidence that motor processes may play a facilitative role for at least some cognitive tasks

(e.g., Beattie, 1980; Glenberg et al., 1998; Laeng & Teodorescu, 2002; Lawrence, Myerson, Oonk, & Abrams, 2001; Spivey & Geng, 2001).

In a recent paper, Wesp, Hesse, and Keutmann (2001) proposed that hand and arm gestures facilitate the maintenance of spatial representations in working memory. Drawing on Baddeley's (1986) working memory model, they hypothesized that repetitive motor or submotor activity (e.g., reafference from motor commands) can maintain spatial information in a working memory buffer, in much the same way as repetitive vocalizations (or subvocalizations) can maintain verbal information in the phonological loop. To test this hypothesis, Wesp et al. had participants describe a painting either from memory or with it visually present. They reasoned that if gestures serve to sustain spatial representations, they should occur more often when descriptions are made from memory than when the spatial information is visually available. Their results confirmed the hypothesis. Participants gestured roughly twice as often when describing the painting from memory. Working independently, but from similar theoretical assumptions, De Ruiter obtained comparable results (De Ruiter, 1998, Experiment 3).

According to Wesp et al. and De Ruiter, the fact that gestures aid spatial memory has misled some investigators (e.g., Krauss & Hadar, 1999; Morsella & Krauss, 1999; Rauscher, Krauss & Chen, 1996) to conclude that they can facilitate the retrieval of words with spatial semantic content. Gestures may affect speech, they contend, but the effect is an indirect one, mediated by spatial memory.

The general idea that co-speech gestures can facilitate speech production is not a new one (e.g., Butterworth & Hadar, 1989; De Laguna, 1927; Dobrogaev, 1929; Feyereisen & de Lannoy, 1991; Frick-Horbury & Guttentag, 1998; Krauss &

Hadar, 1999; Mead, 1934; Rose & Douglas, 2001; Werner & Kaplan, 1963).

However, the view that they preferentially facilitate the production of speech with spatial content rests largely on two recent findings: (1) Gestures are more likely to accompany lexical search for words with spatial semantics than words with other kinds of content (Morsella & Krauss, 1999; Rauscher et al., 1996). (2) Gesture-restriction selectively impairs the retrieval of speech with spatial content (Rauscher et al., 1996). Wesp et al. and De Ruiter would attribute these findings to the role gestures play in spatial memory. By maintaining the spatial concept that underlies the semantic of the to-be-selected word (and that will ultimately participate in lexical selection), they affect speech production indirectly.

In contrast, the *Gestural Feedback Model* (GFM) (Krauss & Morsella, 2002; Morsella, 2002) holds that gestures facilitate speech production more directly, by continually activating, through feedback from effectors or motor commands, the prelinguistic sensorimotor features that are part of the semantic representations of target words. Using gestures this way is helpful because purposefully-activated mental representations tend to be transient, and the process of activating them is effortful (Farah, 2000). As a result, it is difficult to hold them in mind for the lengthy intervals that often occur in lexical search (e.g., in a Tip-of-the-Tongue state).

A key difference between the GFM and the Wesp et al. and De Ruiter view is that, in the former, gestures also can activate and sustain *embodied* representations—that is, semantic representations that are grounded in bodily interactions with the world (Barsalou, 1999; Glenberg, 1997). These representations tend to be of tangible concepts, and there is evidence that gesturing tends to be associated with the retrieval of concrete rather than abstract

words (Morsella, 2002; Morsella & Krauss, 1999). The movements associated with such representations can be quite idiosyncratic, reflecting an individual's particular patterns of interactions with the world. They also may represent concepts functionally (reflecting how one interacts with an object) rather than spatially. For example, search for the word "button" might be accompanied by a tapping motion, rather than a finger outlining the object's shape.

De Ruiter (1998, Experiment 4) provides some evidence consistent with the hypothesis that gestures affect spatial memory and not speech production. In his study, participants described pictures that were either "hard" or "easy" to describe. Easy pictures consisted of well organized images, with the elements (e.g., circles, lines, and triangles) placed above, below, or lateral to the other elements, and with no diagonal lines. The elements of hard pictures were arranged randomly, and their lines were diagonal. Because the stimuli were visible as participants described them, De Ruiter assumes that spatial memory was not taxed. He argues that if gesturing facilitates speech production, more gesturing should accompany the descriptions of hard than easy pictures; if, on the other hand, gesturing aids spatial memory and plays no direct role in speech production, then the same amount of gesturing should accompany the descriptions of both kinds of pictures. De Ruiter found practically identical rates of gesturing in the hard and easy conditions, leading him to conclude that gestures facilitate spatial memory and not speech production. Of course, this conclusion rests on the assumption that his hard stimuli were more difficult to describe than the easy ones, an assumption that was never tested directly.<sup>1</sup>

Our study can be seen as an elaboration and extension of the Wesp et al. and De Ruiter experiments. In it, participants described visual objects that were

either present or absent. Our stimuli consisted of 40 visual objects that varied along several dimensions. Some of the stimuli were identifiable objects that were readily nameable (“Codable,” after Brown & Lenneberg, 1954). These codable objects can be readily labeled, e.g., a house, flower, and an ice cream cone. Other objects were nonsense figures that resembled no familiar objects (Non-codable). The non-codable stimuli varied in their complexity, from a simple squiggle to a complex medley of irregular lines and shapes. Like Wesp et al., we expected more gesturing when the stimulus was described from memory than when it was visually accessible. We also predicted less gesturing to accompany descriptions of Codable than the Non-codable images. But because we believe that gestural facilitation involves more than spatial representations, we also expected a substantial amount of gesturing to occur in the Present condition: Participants may gesture in order to retrieve words whose semantics are not spatial.

We also took the opportunity to address another question: Since gestures aid the recall process, what happens when participants are restricted from gesturing as they recall and describe images? If gestures only affect speech indirectly, by facilitating spatial memory, restricting gesturing should have no effect on speech production when the stimulus is visually accessible. However, because we believe that gestural facilitation involves more than the maintenance of spatial representations, we predict that restriction will have a deleterious effect on the description task regardless of whether the visual object is Present or Absent.

## Method

### *Overview*

We videotaped participants as they described objects that were either visually Present or Absent. Some of the pictures depicted codable objects; the others were abstract line drawings that varied in complexity. We also varied whether or not participants were allowed to move their arms. A mixed 2x2x2 design was employed with movement Restricted or Unrestricted (Restriction) and stimulus Present or Absent (Presence) as between-subjects factors, and Codable vs. Non-codable stimuli (Codability) as a within-subjects factor.

### *Participants*

Seventy-nine Columbia University students (44 male and 35 female) received course credit for their participation. Of the 79, 46 participants were in the Unrestricted condition, and 33 were in the Restricted condition.<sup>2</sup> Half of the Unrestricted group was randomly assigned to the Present condition and the other half to the Absent condition (23 in each of the Presence conditions). For those in the Restricted condition, 17 were randomly assigned to the Absent condition and 16 to the Present condition. Another group of 21 students rated the visual objects along several dimensions (see below). All participants were native English speakers.

### *Materials and Apparatus*

Forty green-on-black line drawings served as stimuli (see Figure 1). Twenty-eight were non-codable images based on figures used by Fussell and Krauss (1989a, 1989b), and twelve were line drawings of identifiable images:

armored tank, candle, clock, flower, guitar, hen, hot-air balloon, house, ice cream cone, pencil, television, and wrench.

The experiment utilized two rooms. One (the experimental room) housed the participant, and contained two video cameras, an intercom speaker, and a computer monitor for stimulus presentation. One camera was trained on the participant's face and torso, and the other was trained on the computer monitor. In the other room (the observation room) the experimenter monitored and recorded the events occurring in the experimental room. The signal from the video cameras in the experimental room were inputted to a Panasonic 3500 System Switcher (WJ-3500; Panasonic Company; Anaheim, CA) in the observation room. This produced a split screen image showing both the participant and the computer monitor that was recorded on a VCR.

#### *Procedure*

Participants were run individually. The experiment was described to them as a referential communication study. Participants received instructions via the program PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) and were told that their descriptions of forty images would be tape recorded and played, a week later, to another participant whose task it would be to identify the images from a larger selection. To reduce self-consciousness, participants were told that their face was being video-recorded to help decipher syllables that were unintelligible. A camera was trained on the participant's head and torso. Participants in the Restricted condition, wore dummy electrodes, and were told that we were trying to discover how arousing the description task was by measuring their Galvanic Skin Response. The electrodes were put on both

forearms, and participants were told that movement of the limbs could ruin the quality of the recordings.

On each trial, the stimulus appeared on the computer monitor for 10 s. At the end of the inspection period, a message appeared instructing the participant to describe the stimulus. Participants had up to 45 s to describe the stimulus, after which the display automatically timed out, and the screen read, “Your time has expired. Click the mouse to continue to the next trial.” Participants who completed their description before the 45 s timeout could continue on to the next trial by pressing a mouse key. Stimuli were presented in random order. In the Present condition, the stimulus picture remained on the screen during description period. In the Absent condition, the screen was blank after the inspection period. Participants were fully debriefed at the end of the session. When interviewed after debriefing, none suspected the true purpose of the study.

#### *Dependent Measures*

The primary dependent variable was gesture rate: the proportion of time during which participants gestured during the description phase of the trial. This was obtained by examining the video record and counting the number of picture frames that captured gesturing. The number of frames were converted to seconds (there are 30 frames per second) and divided by trial duration. This provided us with an estimate of gesture rate. We observed a variety of different kinds of gestures (for a review of gesture types, see Krauss et al., 1996). We coded lexical gestures following previously used criteria (Rauscher et al., 1996): As complex, spontaneous, non-rhythmic movements of the fingers, hands, and arms. We also coded *motor movements*—the brief, repetitive, rhythmic

movements that often accompany speech. These gestures are generally thought to be unrelated to the ideational content of the accompanying speech (see McClave, 1994), and may play a role in formulation of the speech prosody. Other gestures such as symbolic gestures (e.g., the “thumbs up” sign) and adaptors (e.g., scratching the nose and adjusting clothing) were coded and excluded from the analysis. Coding was done blind to experimental conditions. To obtain an estimate of inter-rater reliability for the coding of gestures, a randomly sampled subset (100 of the 790 trials) was coded by two additional judges who were blind to the hypotheses and experimental conditions of the study. Inter-rater reliability was high: the mean correlation coefficient among ratings made by the three judges was .87 ( $SD = 0.036$ ).

To learn more about the properties of stimuli associated with gesturing, we had another group of participants ( $n = 21$ ) rate the forty objects on complexity, describability, memorability, and verbal-codability using 6-point bipolar scales ranging from “not at all” to “very.” The dimensions were defined as follows. *Complexity*: How complex is the visual object? *Describability*: How easy would it be to describe this visual object? *Memorability*: How easy would it be to remember this visual object? *Verbal Codability*: How readily could this visual object be named? Participants were also given examples of visual objects that were judged by four raters to be representative of the high or low end of each of the four dimensions.<sup>3</sup>

## Results

### *Movement Analysis*

Lexical gesture rate (proportion of time gesturing during the description phase) was substantially higher in the stimulus Absent ( $M = .436$ ,  $SEM = .041$ ) than in the Present condition ( $M = .256$ ,  $SEM = .043$ ), and Non-codable images were accompanied by more gesturing than Codable images ( $M = .400$ ,  $SEM = .045$  and  $M = .292$ ,  $SEM = .042$ , respectively). The means are plotted in Figure 2. A 2x2 analysis of variance with Presence as a between-subjects factor and Codability as a within-subjects factor revealed reliable main effects for Presence,  $F(1, 44) = 4.965$ ,  $p = .031$  (partial eta squared = .10), and Codability,  $F(1, 44) = 23.395$ ,  $p < .001$  (partial eta squared = .35), with no interaction,  $F(1, 44) = .684$ ,  $p = .413$  (partial eta squared = .02). The same pattern of results is obtained in by-item analyses in which visual object is the unit of analysis. Stimuli associated with high rates of gesturing in the Present condition were associated with gesturing in the Absent condition ( $r = .64$ ,  $p < .001$ ).

Motor movements were rare, occurring about 1% of the time when the image was present, and 0.8% of the time when it was absent; the difference was not significant,  $F(1, 44) = .429$ ,  $p = .516$  (partial eta squared = .01). However, significantly more motor movements occurred with Codable ( $M = .016$ ,  $SEM = .005$ ) than with Non-codable images ( $M = 0.003$ ,  $SEM = .001$ ),  $F(1, 44) = 6.188$ ,  $p = .017$  (partial eta squared = .14). The same result was obtained in a by-item analysis. This result could be a simple consequence of the fact that, because motor movements and lexical gestures use the same limbs, the high rate of lexical

gesturing (in the Non-codable condition) precludes motor movements from occurring.

### *Speech Analysis*

We also looked at speech rate calculated in syllables per second (sps) as an index of fluency. Because the entire corpus of descriptions totaled nearly 20 hours of speech, we decided to calculate speech rate for the descriptions of the ten visual objects with the longest average trial duration ( $M = 37.366$  s,  $SEM = .310$  s), under the assumption that stable speech-rate effects are most likely to be found in longer descriptions. (Obviously, rates calculated from, say, three-word descriptions are unlikely to be reliable.) As expected, these 10 stimuli were non-codable visual images that were quite complex. We transcribed the selected trials *verbatim* and obtained syllabic rate by tallying the number of syllables divided by trial time. To obtain a measure of inter-rater reliability, a randomly sampled subset (100 of the 790 descriptions) were coded by two additional judges. Rates for the two judges were virtually identical; the mean correlation coefficient was 1.0, and the average difference between the counts of the raters was about one syllable per description.

We performed a 2x2 analysis of variance with Presence and Restriction as between-subjects factors.<sup>4</sup> Across all conditions, speech rate decreased significantly when participants' hand movements were restricted ( $M_{\text{Restricted}} = 2.583\text{sps}$ ,  $SEM = .113$ ) versus unrestricted ( $M_{\text{Unrestricted}} = 2.957\text{sps}$ ,  $SEM = .073$ ),  $F(1, 74) = 9.400$ ,  $p = .003$  (partial eta squared = .11). Interestingly, speakers spoke less rapidly when the visual object was present ( $M_{\text{Present}} = 2.693\text{sps}$ ,  $SEM = .094$  vs.  $M_{\text{Absent}} = 2.909\text{sps}$ ,  $SEM = .091$ ),  $F(1, 74) = 4.035$ ,  $p < .05$  (partial eta squared =

.05). Presence and Restriction did not interact significantly,  $F(1, 74) = 1.301$ ,  $p = .258$  (partial eta squared = .02). In the Unrestricted condition, mean speech rate for Absent was 3.012sps ( $SEM = .075$ ) and for Present was 2.902sps ( $SEM = .125$ ); in the Restricted condition, mean speech rate for Absent was 2.769sps ( $SEM = .187$ ) and for Present was 2.373sps ( $SEM = .094$ ). The means for the four treatments are plotted in Figure 3.

#### *Properties of the Visual Objects that Predict Gesturing*

Using general linear models, we examined how gestural activity was related to ratings of stimuli on complexity, describability, memorability and verbal-codability. In the Present condition, *memorability*, *describability*, and *verbal-codability* are strong predictors of lexical gesture rate ( $r_s = -.75, -.72$ , and  $-.78$ , respectively,  $p_s < 0.0001$ ). Interestingly, *complexity* predicts neither gesture rate nor trial duration ( $r_s < .10$ ). The same pattern of results is found in the Absent condition (see Table 1). Trial durations were significantly longer for objects that were rated low on *memorability*, *describability*, and *verbal-codability* (each  $r < -0.81$ ,  $p < 0.0001$ ).

A different pattern of results is obtained for motor movements. Motor movements rate is positively related to *memorability*, *describability*, and *verbal-codability* (for the Absent condition,  $r_s = .48, .46, .50$ ; for the Present condition,  $r_s = .42, .35, .44$ ;  $p_s < .05$ ), but, as with lexical gestures, not for complexity. Speech rate was uncorrelated with the four measures.

## **Discussion**

Along with Wesp et al. (2001) and De Ruiter (1998), we found that speakers gesture more when describing visual objects from memory than when

the objects are visually accessible. We also found that speakers gesture more when describing objects that tax spatial working memory, such as drawings that are difficult to remember and encode verbally.

With these findings in hand, we can confidently conclude that movements are prevalent when spatial working memory is taxed, and it is reasonable to assume that they facilitate the recall of spatial information. However, Wesp et al. and De Ruiter conclude that gesture's apparent facilitatory effect on speech production is mediated by its effects on spatial working memory. Although their studies support the conclusion that hand-arm movements can facilitate spatial memory, their data really do not bear directly on this issue. In our study, we found a substantial amount of gesturing when the visual object was present (see Figure 2), and this raises the possibility that more than the maintenance of spatial images may be involved. If gestures *only* facilitated spatial memory, participants would have no need to gesture when spatial memory was not involved. We suggest that they did so because gesturing facilitated the process of retrieving from lexical memory the words they needed to describe the stimuli. In this connection, it is relevant that participants gestured more when describing objects that were Non-codable than they did describing Codable objects.

In addition, we found that gesture-restriction decreased speech rate in both the Present and Absent conditions. According to the spatial memory hypothesis, restriction should have affected the description task only in the Absent condition, since spatial memory should not have been problematic when the object was visually available. In light of these findings, the hypothesis that gestures function solely to maintain spatial information seems less tenable. According to the GFM, restriction led to dysfluency because gestures normally

aid speech production by activating the sensorimotor features of semantic representations.<sup>5</sup>

In our experiment a visual object's describability predicted the rate of gesturing that accompanied its description in both the Present and Absent conditions. However, in two experiments, De Ruiter found no difference in the gesture rates that accompany descriptions of "easy" and "hard" pictures, and from this he concluded that describability has no effect on gesture rate. De Ruiter's "hard" and "easy" pictures differed only in the way their elements were arranged. We believe that his classification actually reflects stimulus *complexity*, not describability, a related but importantly different concept. Objects can be both very complex and readily describable. For example, a picture of a locomotive can be seen as a complex arrangements of geometric forms, but it is readily identified and described as "a locomotive." We found gesture rate to be inversely related to a visual object's memorability, describability, and verbal codability, but unrelated to its complexity. If we are correct that de Ruiter's pictures varied in complexity rather than describability, our findings are in harmony, for we, too, found that complexity did not predict gesture rate.

A proponent of the view that the function of gestural movement is primarily communicative (e.g., Beattie & Shovelton, 2000; Graham & Argyle, 1975; Kendon, 1994) might view the movements observed in our experiment as reflections of long-standing habits of using gestures communicatively. De Ruiter (1998) makes the argument explicitly.<sup>6</sup> However, this position is difficult to reconcile with the finding that speakers gestured most often when their memory was taxed. If the gestures we observed were communicatively intended, an

equal amount of gesturing should have occurred in the Absent and Present conditions.

The most reasonable conclusion one can draw from these data is that gestures probably serve multiple intrapersonal and interpersonal functions. For instance, Goldin-Meadow, Nusbaum, Kelly, and Wagner (2001) have found that some gestures decrease cognitive load during an explanation task, and Alibali, Kita, and Young (2000) have found gestures that facilitate the conceptual processes preceding language production. And there can be little doubt that in some instances gestures can and do function communicatively. That these movements may be accomplishing several things at once is a problem for theories that strive for simplicity, but theories of the role of gestures must allow for this possibility. We believe that the GFM provides a good account of one of the functions gestures serve, but a complete account of all of these functions is not yet within our grasp.

### References

- Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language & Cognitive Processes, 15*, 593 – 613.
- Bacon, F. (1891). *The advancement of learning, Book 2*. (4<sup>th</sup> ed.). London: Oxford University Press.
- Baddeley, A. D. (1986). *Working memory*. Oxford, England: Oxford University Press.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577 – 560.
- Beattie, G. W. (1980). A further investigation of the cognitive interference hypothesis of gaze patterns during conversation. *British Journal of Social Psychology, 20*, 243 – 248.
- Beattie, G. W. & Shovelton, H. (2000). Iconic hand gestures and the predictability of words in context in spontaneous speech. *British Journal of Psychology, 91*, 473 – 491.
- Brown, R. W., & Lenneberg, E. H. (1954). A study in language and cognition. *Journal of Abnormal & Social Psychology, 49*, 454 – 462.
- Butterworth, B., & Hadar, U. (1989). Gesture, speech and computational stages: A reply to McNeill. *Psychological Review, 96*, 168 – 174.
- Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, & Computers, 25*, 257 – 271.
- De Laguna, G. (1927). *Speech: Its function and development*. New Haven, CT: Yale University Press.
- De Ruiter, J. -P. (1998). *Gesture and speech production*. (Doctoral Dissertation, Max Planck Institute for Psycholinguistics, 1998).
- Dobrogaev, S. M. (1929). Ucnenie o reflekse v problemakh iazykovedeniia [Observations on reflexes and issues in language study]. *Iazykovedenie i Materializm, 105* – 173.

- Farah, M. J. (2000). The neural bases of mental imagery. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (2<sup>nd</sup> ed., pp. 965 – 974). Cambridge, MA: MIT Press.
- Feyereisen, P., & de Lannoy, J, -D. (1991) *Gesture and speech: Psychological investigations*. Cambridge, England: Cambridge University Press.
- Frick-Horbury, D., & Guttentag, R. E. (1998). The effects of restricting hand gesture production on lexical retrieval and free recall. *American Journal of Psychology*, 111, 43 – 62.
- Fussell, S., & Krauss, R. M. (1989a). The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of Experimental Social Psychology*, 25, 203 – 219.
- Fussell, S. R., & Krauss, R. M. (1989b). Understanding friends and strangers: The effects of audience design on message comprehension. *European Journal of Social Psychology*, 19, 509 – 526.
- Glenberg, A. M. (1997). What memory is for. *Behavioral & Brain Sciences*, 20, 1 – 55.
- Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition*, 26, 651 – 658.
- Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesture lightens the load. *Psychological Science*, 12, 516 – 522 .
- Graham, T. A. (1999). The role of gesture in children's learning to count. *Journal of Experimental Child Psychology*, 74, 333 – 355.
- Graham, J. A. and M. Argyle (1975). A cross-cultural study of the communication of extra-verbal meaning by gestures. 10, 57 – 67.
- Hebb, D. O. (1968). Concerning imagery. *Psychological Review*, 75, 466 – 477.
- Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction*, 27, 175 – 200.
- Krauss, R. M., Chen Y., Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? *Advances in Experimental Social Psychology*, 28, 389 – 450.

- Krauss, R.M., & Hadar, U. (1999). The role of speech-related arm/hand gestures in word retrieval. In R. Campbell & L. Messing (Eds.), *Gesture, speech, and sign* (93 – 116). Oxford: Oxford University Press.
- Krauss, R. M., & Morsella, E. (2002). *The gestural feedback model*. Unpublished manuscript, Columbia University.
- Laeng, B., & Teodorescu, D.-S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science*, 26, 207 – 231.
- Lawrence, B. M., Myerson, J., Oonk, H. M., & Abrams, R. A. (2001). The effects of eye and limb movements on working memory. *Memory*, 9, 433 – 444.
- McClave, E. (1994). Gestural beats: The rhythm hypothesis. *Journal of Psycholinguistic Research*, 23, 45 – 66.
- Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.
- Morsella, E. (2002). The motor components of semantic representation. (Doctoral dissertation, Columbia University, 2002). *Dissertation Abstracts International: Section B: the Sciences & Engineering*, 63 (4-B). (University Microfilms No. AAI3048195).
- Morsella E., & Krauss, R. M. (1999, October). *Electromyography of arm during the lexical retrieval of abstract and concrete words*. Poster presented at the Annual Convention of the Society for Psychophysiology, Granada, Spain. *Psychophysiology*, 36, S82 - S82.
- Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in the processing of speech. *Psychological Science*, 7, 226 – 231.
- Rose, M., & Douglas, J. (2001). The differential facilitatory effects of gesture and visualisation processes on object naming in aphasia. *Aphasiology*, 15, 997 – 990.
- Sapir, E. (1949). The unconscious patterning of behavior in society. In D. Mandelbaum (Ed.), *Selected writing of Edward Sapir in language, culture and personality* (pp. 544 – 559). Berkeley, CA: University of California Press.
- Spivey, M., & Geng, J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, 65, 235 – 241.

- Washburn, M. F. (1928). Emotion and thought: a motor theory of their relation. In C. Murchison (Ed.), *Feelings and emotions: The Wittenberg Symposium* (pp. 99 – 145). Massachusetts: Clark University Press.
- Werner, H., & Kaplan, B. (1963). *Symbol formation*. New York: Wiley.
- Wesp, R., Hesse, J., Keutmann, D., & Wheaton, K. (2001). Gestures maintain spatial imagery. *American Journal of Psychology*, 114, 591 – 600.

## **Acknowledgments**

We gratefully acknowledge the advice and comments of Lois Putnam, Robert Remez, Tory Higgins, James Magnuson, Michele Miozzo, Donna Frick-Horbury, Richard Wesp, and the assistance of Monica Ciapponi, Ari Dollid, Anna Marie Nelson, Anne Ribbers, Adam Waytz, and Jillian White.

Table 1

*Correlations Between Lexical Gesture Rate And Object Properties*


---

	<u>Presence Condition</u>	
	<u>Present</u>	<u>Absent</u>
Complexity	-.06	.09
Describability	-.72	-.67
Memorability	-.75	-.69
Verbal Codability	-.78	-.72

---

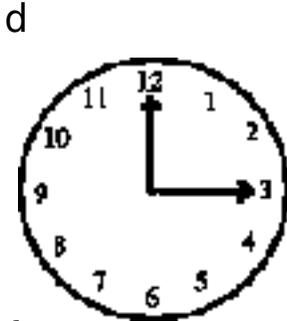
*Note:* Based on a by-item analysis with the 40 objects. All of the *rs* are significant ( $ps < .01$ ) except for those of *complexity*.

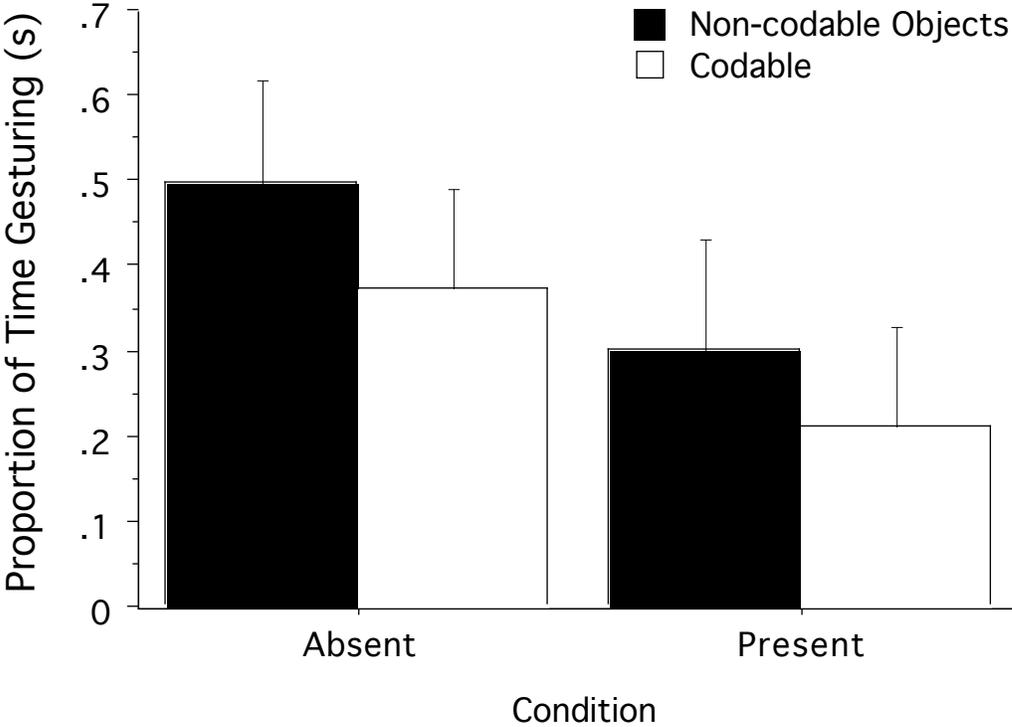
## Figure Captions

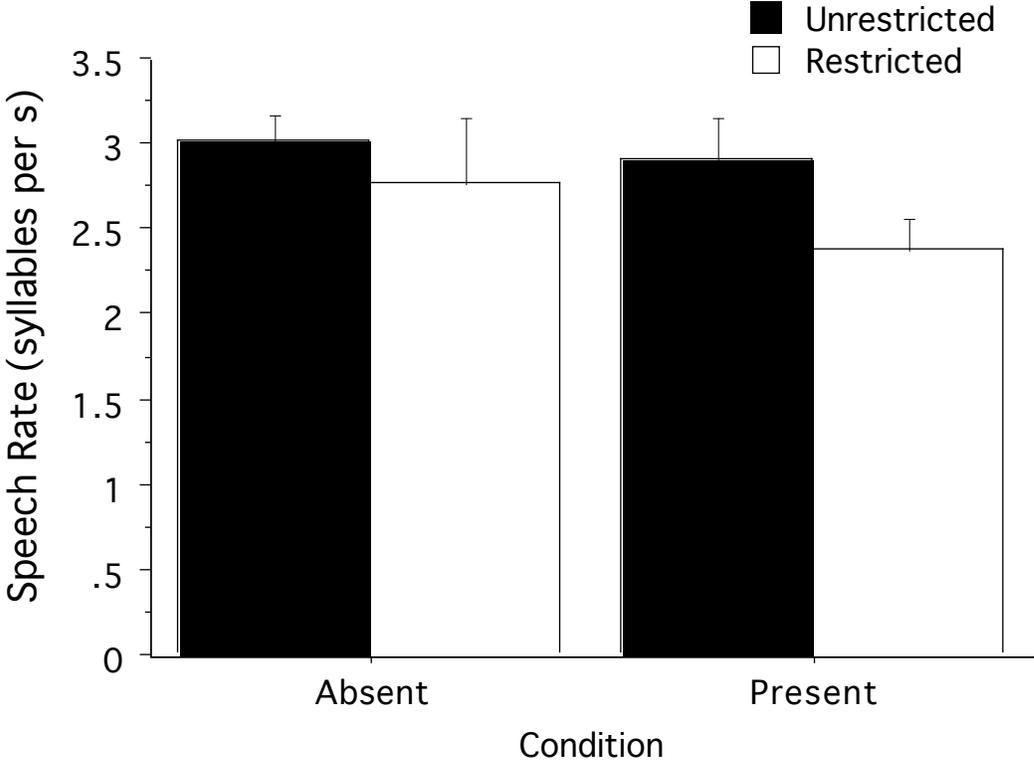
*Figure 1:* Sample of visual stimuli. Visual object *b*, *e*, and *f* are codable, and *a*, *c*, and *d* are non-codable objects.

*Figure 2:* Mean proportion of time (s) gesturing as a function of condition. Error bars indicate 95% confidence interval.

*Figure 3:* Mean speech rate (syllables per s) as a function of condition. Error bars indicate 95% confidence interval.







## Footnotes

---

<sup>1</sup> De Ruiter's Experiment 4 has a number of shortcomings, among them the fact that results for nearly 40% of the participants were discarded, that stimuli were classified as "hard" or "easy" on an *a priori* basis, and the problem (which de Ruiter acknowledges) of drawing a positive conclusion from a failure to reject the null hypothesis. However, our primary disagreement concerns the assumption that a lexical retrieval model would predict more gesturing with his "hard" stimuli than his "easy" stimuli. We will address this issue in the Discussion section.

<sup>2</sup> With the resources at hand, the practical problems involved in expediently setting up and removing the materials for the Restricted condition (e.g., dummy electrodes and their leads) made it unfeasible for us to randomly determine whether a given participant would be assigned to the Restricted or Unrestricted condition. Hence, these conditions were run in blocks—a series of consecutive participants was run in the Unrestricted condition and then another group was run in the Restricted condition.

<sup>3</sup> These objects were not used in the study.

<sup>4</sup> The data of one participant in the Restricted condition who claimed to be ill and was not speaking normally were removed from the analysis.

<sup>5</sup> An alternative explanation for the effects of gesture restriction on speech is that it is due to the dual-task nature of speaking while keeping one's hands still. While we cannot reject this possibility definitively, we think it is unlikely. One way to rule it out would be to show that restriction affects speech but not other, comparable cognitive tasks, but it is not clear what an appropriate control task would be. In their study Rausher et al., (1996) imposed movement restrictions on subjects' hands or feet, and found effects on the former but not the latter. It is significant that Rausher et al.

---

found effects of hand-arm restriction only on speech with spatial content. Speech rate (and dysfluency rate) for nonspatial speech was unaffected. If the effect is due to the added cognitive load imposed by the task of not moving one's hands, we would expect it to affect all speech and not to be limited to a particular kind of semantic content.

<sup>6</sup> "The fact that people gesture on the telephone is also not necessarily in conflict with the view that gestures are intended to be communicative. It is conceivable that people gesture on the telephone because they always gesture when they speak spontaneously—they simply cannot suppress it" (De Ruiter, 1998, p. 18).