**Title**
How are speech and gesture related?

**Permalink**
https://escholarship.org/uc/item/35b912q8

**Journal**
Proceedings of the Annual Meeting of the Cognitive Science Society, 24(24)

**ISSN**
1069-7977

**Authors**
Yoshida, Hanako
Smith, Linda B
Ping, Raedy M
et al.

**Publication Date**
2002

Peer reviewed

# How are Speech and Gesture Related?

**Hanako Yoshida (hayoshid@indiana.edu)**
**Linda B. Smith (smith4@indiana.edu)**
**Raedy M. Ping (rping@indiana.edu)**
**Elizabeth L. Davis (elldavis@indiana.edu)**
Psychology Department, Indiana University, 1101 E. 10th Street
Bloomington, IN 47405-7007 USA

## Abstract

People gesture when they speak. Despite considerable attention from a variety of disciplines, the precise nature of the relation between gesture, speech and thought has remained elusive. The research reported here considers two very different hypotheses about the fundamental relationship. By one account, gesture is a consequence of physiological arousal. By another account, gesture use reflects more cognitive processes and is strongly linked to mental representation. This paper seeks the mechanism underlying the link between gesture and speech by showing that both mental representation and physiological arousal are reflected in our gesture use.

## Introduction

Communication is the activity of transmitting information about things in the world. In everyday life, we refer to things by employing a number of communicative tools. Language, gestures, facial expressions, and non-linguistic vocalizations are often considered communicative tools. Interestingly, we integrate multiple modalities in our communication and indeed, past research suggests a tight link between modalities, and particularly between speech and gesture.

This study is motivated by two approaches to the study of gesture. McNeill (1992), Kita (2000) and McNeill & Duncan (2000) consider the relationship between the specific forms of gesture and speech, specifically between iconic gesture and speech. An iconic gesture is one in which the speed, motion, or shape of the gesturing hand resembles the meaning being conveyed. McNeil and colleagues consider the emergence of iconicity in gesture as natural, and due to the embodied and imagistic nature of thought. Another approach, however, considers gesture as a consequence not of meaning, but of arousal. Schwarts & Black, (1996) and Iverson & Thelen (1999) particularly suggest that gestures can be explained as an overflow of one's excitement, which is physiologically produced by speech. The idea here is that speech is a motor program, and as the speaker is aroused—by topic or by the very act of speaking—that energy overflows and is evident in other bodily movements.

This present research is concerned with one flashpoint in these accounts: the idea that gesture use fluctuates. At present, the variability in gesture use and in kind of gesturers is not well explained. Yet this very variability should be the key to underlying mechanism. We demonstrate the bi-directional relationship between gesture, thoughts and arousal by confirming that both mental representation and arousal are reflected in and influence iconic gesture use. We experimentally manipulate speech rate as a means of increasing the "motor overflow" from speech. Although speech rate is at best an indirect marker of arousal, we believe it is a good starting point for a mechanistic explanation of variability in gesture use.

## Iconicity

Communicative tools in general range from conventional forms that are more arbitrary, to less conventional forms that are more iconic. An example of the conventional forms in traditional verbal languages would be the word "dog" which can hold the meaning only when both senders and receivers are knowledgeable about the rule indicating the label—'a dog' means a dog. This form can also be seen in written language when one spells out letters, d, o, g to refer to a dog. Again, receivers can make sense out of the particular combination of letters only by knowing the convention, that the order and the combination of letters are signifying a dog. People also refer iconically. For example, in verbal language use, one might imitate a sound of a dog to refer to a dog (e.g., "woof, woof!"), and in American Sign Language the word for a dog consists of a pat on one's leg and a snap of one's fingers as if calling a dog.

People also often gesture iconically (Kita, 2000). For example, a speaker might make a circle with one's hands to describe shape and/or the size of a plate to which the person refers. Indeed, even nonverbal primates may gesture to convey meaning. For example, Tanner and Byrne (1996) recently reported that several observed gestures of gorillas are iconic. Another example that highlights both natural emergence and ease of iconic gesture use is an observation made by Goldin-Meadow and Feldman (1977). They observed

communication formation of linguistically deprived deaf children who were unable to acquire oral language naturally and who were not exposed to a standard manual language. They reported that children spontaneously refer to things by using iconic gestures. All these results suggest a tight and natural relationship between gesture and thought. If this is so, then experimental manipulations of the content of thought should have direct and measurable effects on gesture.

We show experimental evidence for this connection in Experiment 1. Participants were asked to read and re-tell the story to their children. A story was about either a carp, which was expected to induce more gesture that engages a hand held in a vertical position or a stingray, which was expected to induce more gesture that engages a hand held in a horizontal position. If mental representations are reflected in our gesture use, then we should expect these different gesture patterns across participants who read the same story about the two different fishes. Moreover, Experiment 1 also serves as baseline measure of gesture production with a normal speech rate. This baseline was used to manipulate participants' speech rate in Experiment 2.

## Gesture as motor overflow

The second idea that gestures result from an overflow in the motor program for speech is not one that has received much empirical attention, nor one that has been well specified theoretically. However, this idea suggests that amount of motor activity in speaking should have direct influences on amount of gestures, and this should be thought so independently of the content of speech. Accordingly, to test this idea, in Experiment 2, we entrained speech to either a slow or fast metronome. Do gestures decrease when people speak at slower than normal rates? Do gestures increase given the faster rate of speaking? They should if gestures reflect motor overflow. Finally, does rate of speech influence rate of gesture independent of content? An affirmative answer would suggest two independent driving forces behind gesture: the meaningful content of thought and arousal.

# Experiment 1

## Method

**Participants** Twenty mothers and their children whose ages ranged from 26 to 61 months participated in the study.
**Procedure** Participants were assigned to either the Carp condition or the Stingray condition. The difference in the manipulation between the conditions was a kind of fish used as a hero of the story. Parents were presented with a sheet of paper, which contained either a carp story or a stingray story and were asked to read and retell the story from memory to their child. Participants were also asked to have the child tell the story back to them. While participants were retelling a story, the aspects of their story telling were recorded by a video camera for a later coding. The story used in the study is provided in Table 1.

## Result

Each parent's session was coded for several variables: all gestures, two target hand positions and speech rates. All variables were measured as rates, that is, production per minute. In order to code for these variables, the video was played in slow motion; the exact times that a hand movement began and ended were coded. Two target hand positions were counted. The "Carp-swimming" gesture consisted of holding the hand flat palm perpendicular to the horizontal table surface and the "Stingray-swimming" gesture was defined as a gesture in which the hand was flat palm down parallel to the horizontal surface of the table. These gestures were coded as either corresponding or non-corresponding gestures depending on whether the hand position corresponds to the fish in the assigned story (e.g., a carp gesture given a carp story was corresponding gesture and a stingray gesture given a carp story was a non-corresponding gesture.) All gestures were independently scored by two observers.

Participants produced gestures (target gestures plus all other gestures) at a rate of 4.1 gestures per a minute. However, as shown in Figure 1, participants produced more target gestures that corresponded to the story than gestures that did not correspond. In addition, storytellers held their hands in a horizontal position more often when telling the story about the stingray than when telling one about a carp. Finally, the average speech rate was 141.0 words per a minute.

# Experiment 2

## Method

**Participants** Twenty mothers and their children whose ages ranged from 23 to 60 months participated in the study.
**Procedure** The procedure was the same as that of Experiment 1 with an exception. Participants were assigned to one of four conditions: Carp condition, fast metronome, Carp condition with slow metronome, Stingray condition with fast metronome and Stingray condition with slow metronome. The speeds of the metronome were selected based on the average speech rate observed in Experiment 1. Since the normal speech rate in Experiment 1 was 140 words per a minute, we selected 150 beats per a minute for the Fast condition and 80 beats per a minute for the Slow condition. We were conservative in attempting to entrain speech to a

faster rate in this first entraining experiment in an effort to maximize our success in entraining speech. The parent storytellers were not told about the purpose of the metronomes, but were instead told to ignore them. The task of individual participants was identical to that of Experiment 2.

## Table

Table 1: Story used for the study: The Adventures of (Carl the Carp / Steve the Stingray)

(Carl the Carp / Steve the Stingray) was a happy, friendly (carp with a round, silver body, big, round eyes, and a flowing, fan-like tail / stingray with a flat, gray body, beady black eyes, and a long, whip-like tail) who loved swimming with his other (carp / stingray) friends. They especially liked swimming into the depths of the water and returning to the surface. One day, when (Carl the Carp / Steve the Stingray) and his friends were enjoying another adventure in the depths of the water, they sensed some danger. They were eating lunch when, all of a sudden, there was a giant shadow looming over them. (Carl the Carp / Steve the Stingray) turned around, but it was too late! He found himself inside the belly of a big fish. He thought to himself, "This is the end of (Carl the Carp / Steve the Stingray). I'll never be able to swim around with my (carp / stingray) friends and have fun again." Just as he was thinking this, he heard a loud rumbling and was pushed out from the fish's stomach. His friends, who had been hiding in the rocks, swam cautiously towards him to make sure he was okay. "(Carl / Steve)," one of his friends said, "that big fish just burped you up! Are you all right?" "Yes," a relieved (Carl the Carp / Steve the Stingray) said, "I'm fine. Let's finish our adventure." Then, (Carl the Carp / Steve the Stingray) and all of his (carp / stingray) friends swam happily off for another fun adventure.

## Results

First, we did successfully manipulate speech rate, $t = -2.44$, $p < .05$. The average number of words per minute was 145.5 in the Fast condition and 120.0 in the Slow condition. Second, as can be seen in Figure 2, participants produced more gestures per a minute in the Fast condition than they did in the Slow condition. This indicates that gesture production is tied to speech rate.

As shown in Figures 3 and 4, content effects are seen in the both Fast and Slow conditions, but the effect is reliable only in the Fast condition (for Fast condition; $t =2.198$, $p< .05$, and for Slow condition; $t =1.405$, $p=$

.18.) Participants in both Fast and Slow conditions produced more gestures that corresponded to the story that they retold than gestures that did not correspond. The fact that speech entrained to a slow metronome, 80 beats per a minute, decreases gesture use and weakens the content effect on iconic gestures suggests a direct link between arousal and content.

## Discussion

The two experiments yield 3 main results. First, the content of thought directly influences hand position. Thinking about carps leads one to holds one's hand perpendicular to the tabletop. Thinking about stingrays leads to hands held in a horizontal position over the table. By experimentally manipulating the content of thought, one manipulates hand position. This is consistent with the proposal that gestures iconically represent meaning. Second, speech and gesture are readily entrained to a rhythmic beat. And, faster rates of speech lead to more gestures and slow rates of speech lead to fewer gestures. This is consistent with the idea of gestures emerging as a consequence of arousal. Amount of energy appears to be is a driving force in gesture. Third, arousal (or rate of speech) appears linked to the frequency of iconic gestures. This suggests, contrary to the motor overflow hypothesis, that the relevant activation concerns meaningful content and not just the speech motor plan. High activation leads to more iconically related gestures as if high arousal leads to more highly activated meanings that emerge in both spoken words and in meaning related gestures.
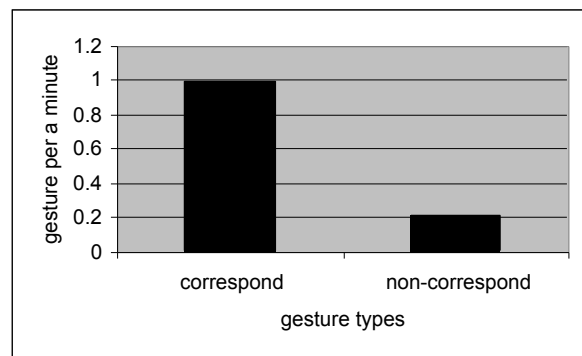
## Figures



Figure 1: correspond and non-correspond gesture
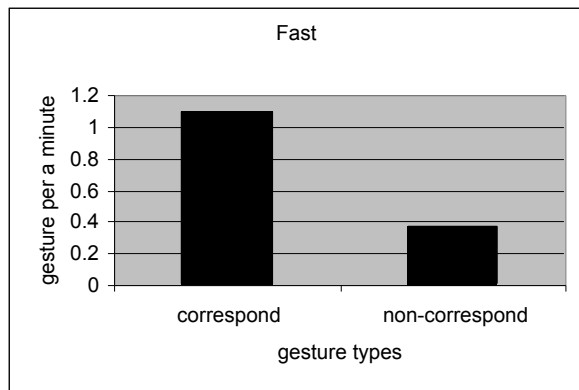
Figure 2: all gestures per minute.



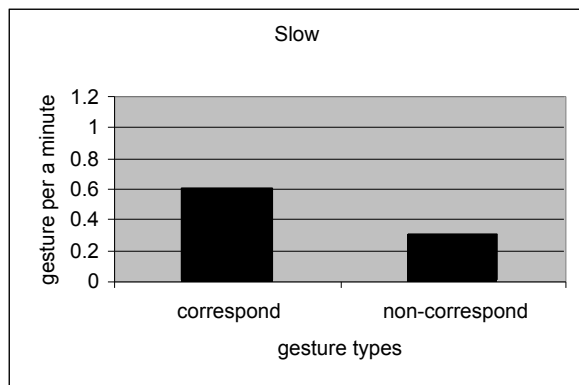Figure 3: correspond and non-correspond gesture in Fast condition



Figure 4: correspond and non-correspond gesture in Slow condition

**References**

Goldin-Meadow, Susan & Feldman, Heidi (1977) The development of language-like communication without a language model. Science. 197(4301): 401-403. US: American Assn. for the Advancement of Science.

Iverson, J.M. & Thelen, E. (1999). Hand, mouth and brain: The dynamic emergence of speech and gesture. Journal of Consciousness Studies, 6 (11-12), 19-40.

Kita S. (2000). Why Do People Gesture? Cognitive Studies, 7 (1), 9-21

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

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

McNeill D. & Duncan, S. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), Language and Gesture: Window into Thought and Action, 141-161. Cambridge: Cambridge University Press, in press.

Schwartz, D. L., & Black, J. B. (1996) Shutting between depictive models and abstract rules: Induction and fallback. Cognitive Science, 20, 457-498

Tanner, J. E. & Byrne, R. W. (1996) Representation of action through iconic gesture in a captive Lowland gorilla. Current Anthropology, 37, 162-173