

Cambridge University Press

978-1-107-13718-9 - Why we Gesture: The Surprising Role of Hand Movements in Communication

David McNeill

Excerpt

[More information](#)

---

*Part I*

Gesture-orchestrated speech

Cambridge University Press  
978-1-107-13718-9 - Why we Gesture: The Surprising Role of Hand Movements in Communication  
David McNeill  
Excerpt  
[More information](#)

---

Cambridge University Press

978-1-107-13718-9 - Why we Gesture: The Surprising Role of Hand Movements in Communication

David McNeill

Excerpt

[More information](#)

## 1 Why we gesture

---

Why do we gesture? Many would say it brings emphasis, energy and ornamentation to speech (which is assumed to be the core of what is taking place); in short, gesture is an “add-on.”<sup>1</sup> However, the evidence is against this. The lay view of gesture is that one “talks with one’s hands.” You can’t find a word, so you resort to gesture. Marianne Gullberg (2013) debunks this ancient idea. As she succinctly puts it, rather than gesture starting when words stop, gesture stops as well. So if, contrary to lay belief, we don’t “talk with our hands,” why do we gesture? This book offers an answer.

The reasons we gesture are more profound. Language itself is inseparable from gesture. While gestures enhance the material carriers of meaning, *the core is gesture and speech together*. They are bound more tightly than saying the gesture is an “add-on” or “ornament” implies. They are united as a matter of thought itself. Even if, for some reason, the hands are restrained and a gesture is not externalized, the imagery it embodies can still be present, hidden but integrated with speech (and may surface in some other part of the body, the feet for example).

As stated in the Preface, the purpose of the current book is to present the multifaceted hypothesis that *to orchestrate speech is why we gesture*. Gestures of course do not always occur. This is itself an aspect of gesture; there is a natural variation of gesture occurrence. Apart from forced suppressions (as in formal contexts), gestures fall on an elaboration continuum, their position an aspect of the gesture itself.<sup>2</sup> The degree of elaboration is the extent to which the gesture adds communicative “push” to the utterance, what Firbas (1971) called

<sup>1</sup> Kendon (2008), who also argues against the view.

<sup>2</sup> Forced suppressions only shift the gesture to some other part of the body – the feet or an overactive torso. We once taped an individual who had been highly recommended to us as an elaborate and vigorous gesturer. Somewhat maliciously, when asked to recount our cartoon stimulus, he sat on his hands yet unwittingly began to perform the gestures typical of the experiment with his feet insofar as anatomically possible—foot up for Sylvester’s ascent, other foot next to it for his ascent inside a pipe, etc. James Goss (personal communication) undertook a systematic study of foot gestures (never published).

## 4 Why we gesture

“communicative dynamism.” At one end, very elaborate gestures participate at the highest levels of communicative dynamism; at the other end, where communicative dynamism or “push” is minimal, elaboration reaches zero, ending with no motion at all; yet this is not a disappearance of the gesture and its imagery; it is the minimum of its concrete enactment. The reality is imagery with speech ranging over the entire continuum. It is visuoactional imagery, not a photo. Gesture imagery linked to speech is what natural selection chose, acting on gesture–speech units free to vary in elaboration. As communicative dynamism varies, the gesture–speech unit moves from elaborate movement to no movement at all. To speak of gesture–speech unity we include gestures at all levels of elaboration.

### 1.1 What is a “gesture”?

The term “gesture” covers a range of phenomena. We focus on one in particular: gesture in our sense is *the intrinsic imagery of language*. Language is inseparable from it. Inseparable, because gesture orchestrates speech; it and speech (and all the language forms speech includes) cannot be sundered. Such gestures are not exotic or rare. They are the ordinary gestures of daily speech and by far the most abundant of any kind of gesture.

A journalist’s cliché portrays gesture as pretense, fake action for show and not for substance—for example, “[t]here is also the suggestion of a [story] plot, or rather a gesture in the direction of a weave of narratives” or “a small gesture stirs giving on a big scale” (both *New York Times*). The cliché is worse than irrelevant. I do not say that it is meaningless, but if adopted it misleads or worse, positively interferes with understanding. In our discussion a gesture is the very fuel of language and thought. Moreover, language could not have evolved without it. Gesture and speech were “equi-primordial” (a term from Liesbet Quaghebeur, pers. comm.). Gesture and language are inseparable, and the journalist’s cliché hides this deep relationship.

I occasionally say “gesticulation” to designate the gestures that orchestrate speech, but the word is far from satisfactory. It conveys a picture of windmilling arms (the *Oxford English Dictionary* confirms), but the gestures we observe are nearly all small, confined to a space in front of the torso. When the ambiguity is harmless, I will say simply “gesture.” Speech-synchronized gestures (and not windmilling arms) are by far the most common form of gesture in narratives and conversations.

Adam Kendon (2004) placed gestures in the category of “actions that have the features of manifest deliberate expressiveness.” I adopt this definition with the qualification that speech-orchestrating gestures cannot be deliberate. Kendon may have meant by “deliberate” non-accidental, and with this I agree; but the word also conveys, “done for a purpose,” and with this I do not

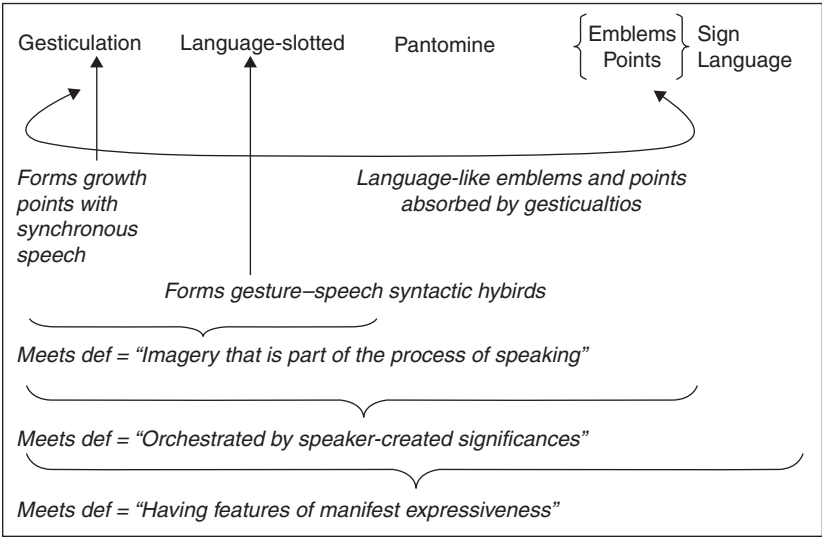


Figure 1.1 Gesture continuum. Cambridge University Press, reprinted with permission.

agree—orchestration of speech via a gesture is not the goal of any speaker in making a gesture. If the speaker intends anything it is to “communicate,” not to perform a speech-orchestrating gesture. Pantomime, which can be deliberate, is not able to orchestrate speech, as will become clear later.

1.2 A gesture continuum

Figure 1.1 shows a gesture continuum and how places along it relate to Kendon’s definition and to the others indicated in the figure.

1.3 Notation

In addition to semiotic properties, a gesture can be described in terms of movement. This mode is incorporated into our notation. One complete “manifestly expressive action” is what Kendon (1980) called a gesture phrase, a “gesture” in normal parlance. For the gesture illustrated in Figure 1.2,

(1.1) he goe[ss / **up** / **through** the pipe] this time #

the *gesture phrase* as a whole is marked by “[” and “]” (the “/” marks a silent hesitation of speech, the “#” an audible breath intake, and large font prosodic stress). A gesture phrase has up to five *phases*. Not all phases need be present,

6 Why we gesture



Figure 1.2 “New” gesture-action. Computer art by Fey Parrill, now on the faculty of Case Western University. Used with permission of University of Chicago Press.

but one or more *strokes*—the image-bearing phase—is obligatory; without a stroke a gesture is not said to have occurred. This is marked in boldface (“up through”). The *preparation* is the hand getting into position to make the stroke and is indicated by the span from the left bracket to the start of boldface (during “goes”). Preparation shows that the gesture, with all its significance, is coming into being—there is no reason the hands move into position and take on form than to perform a stroke yet to come. *Holds* are cessations of movement, either *prestroke* (the “I”), the hand frozen awaiting the stroke, or *poststroke* (the “the”), the hand frozen in the stroke’s ending position and hand shape after movement has ceased. Holds of either kind are indicated with underlining. They show the precise synchrony of stroke and orchestrated speech. *Retraction* is also an active phase, the gesture not simply abandoned but closing down (in this case, during “pipe”). Not atypically, the gesture phase in Figure 1.2 did not align with a syntactic constituent of the sentence, preparation beginning in the middle of “goes.” We will see later salient examples of disconnection and explain why an imperfect alignment of gesture and linguistic form will happen, but the Figure 1.2 example suggests even now the reason is that the gesture has orchestrated the speech—the gesture does not spring from the sentence constituent structure but instead this structure fits or tries to fit the gesture.

### 1.3.1 Storytelling

The gesture examples described in this book are spontaneous, unsolicited and unrehearsed, recorded during storytelling. A participant watches an approximately eight-minute-long animated Tweety and Sylvester cartoon or a full-length Hitchcock film, *Blackmail* (1929), then retells the story, without notes, to a listener who has not seen it. For readers unfamiliar with the genre, Tweety is a feisty, large-headed canary belonging to a feisty elderly grandmother. Sylvester, an enterprising cat with culinary goals, endlessly pursues Tweety; his unchanging lot is frustration and disaster. The genre was familiar to our narrators but not the specific cartoon (“Canary Row,” Warner Brothers, 1950). The cartoon consists of eight episodes, all with the same pursuit-catastrophe theme in amusing variations. McNeill and Levy (1982) chose it originally to show to children, but adults, university-educated and professional, also find it engaging.

One participant (the “speaker”), chosen at random at the start of the session, is shown the cartoon/film in its entirety. The speaker is told in advance that immediately after viewing the cartoon/film he or she will tell the story to the second participant “as accurately and completely as possible, as your listener will have to retell the story based on your narration,” or words to this effect. The second participant was a genuine listener, not one of the experimenters (usually a friend or spouse, never a stranger). The performance was recorded on video with the seated speaker in full camera view and at least the front half of the listener as well. The instructions emphasized that the experiment was about storytelling. Gesture was not mentioned.

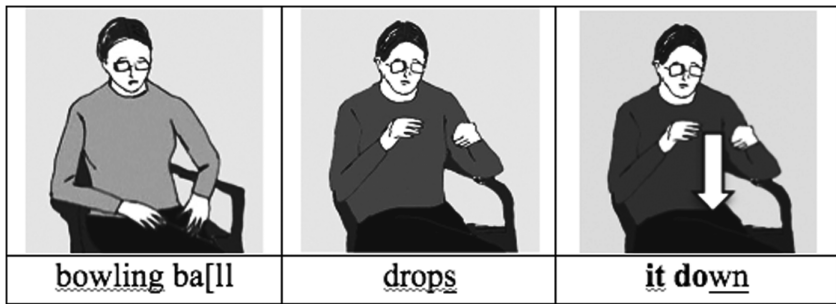
### 1.3.2 Coding validity

Gesture coding requires close attention to kinesic details and an accurate and detailed transcription of speech. Unlike some researchers who, fearing “contamination,” code gesture and speech separately, our method requires them to be coded jointly. Ultimately we are coding *gesture–speech units*, not “gestures” alone (see also Chapter 2). Separating the streams removes the very phenomenon we want to code. Each narration is coded by at least two coders working independently, who then agree on the final coding (a finished transcript includes notes of any disagreements). The development of this method and notation is the achievement of two researchers, Susan Duncan and Karl-Erik McCullough.

How are gestures interpreted? The answer is not obvious. I am of the view that “true” interpretation of a gesture is a hypothesis to be judged for its correspondence to observable facts. These facts include the form and timing of the gesture, and (equally important) how it fits into the immediate context of

speaking. The exact timing of the gesture with speech is important. An example is in the “it down” gesture in the following graphic (the two hands, facing down and curved as if around an object—the bowling ball, and thrusting down), timed as shown:

(1.2) “and Tweety Bird runs and gets a bowling ba[ll and  $\emptyset_{tw}$  drops it down the drainpipe]”



The square brackets mark the beginning and end of motion; the boldface the image-bearing stroke, when the arched hands moved down; the underlining two holds, a prestroke on the final sibilant of “drops” and a poststroke at the end of “down.” This example illustrates a number of points and will be analyzed in full in Chapter 4. However, we can make the point now that the syntax of co-occurring speech—“it” and “down” in different constituents—is dominated by the gesture, as gesture-orchestrated speech implies, and is not a guide for rendering it.

## 1.4 The beat

One type of gesture, the “**beat**,” may seem an exception. A beat appears to be a gesture without an image-bearing stroke. Rather than embodying meaning, beats appear to synchronize with speech rhythm. Bresse (2010) has tracked different hand shapes and orientations of beats with shadings of this function. If anything, beats appear to show the reverse of gesture-orchestrated speech—speech-orchestrated gesture. Indeed, Kevin Tuite (1993) argued that iconic gestures contain rhythmic pulses, in effect inner beats, deriving from speech rhythmicity.

Rhythmicity is unquestionably a factor in all forms of gesture but it does not operate autonomously. Both beat and rhythm arise from meanings on the discourse level. They share a source in contextual highlighting. For example,



the beats and rhythmic emphasis in the following shift successively to follow what is new in each speech unit: “his girlfriend– **Á**lice—Alice **W**hite”—first her role in the story, then her first name, then her last (beats in boldface and rhythmic pulses in enlarged font with accents).

The principle seems to be that highlighting draws effort in both speech and gesture. We can go further and say that an observed beat is a reduced version of a full gesture. It contains its meaning on this level. Some beats stand alone, but many ride on an imagery-bearing gesture stroke—the hands depicting something and, while depicting it, moving up and down or in and out, in beats. The added effort takes this form. We will see an example of such beats later in Figure 1.5. These beats seem to be replications of the gesture on which they ride. They are the gesture made twice or thrice, all at the same time, and have the function of emphasizing the gesture for the significance it has in the discourse beyond its speech unit (as with “Alice White”).

### 1.5 Emblems and pointing

Two other slots on the gesture continuum, the **emblem** and the **point**, have yet other features. Although usually thought of as different gestures, they are surprisingly similar on these features and can be described together.

First, standardization of form. The upward movement of Figure 1.2 indicates the location of the pipe, its position relative to the character and the location inside. This deixis was accomplished not with a dedicated point but was built into the gesticulation itself. A dedicated, stand-alone point, however, has properties that make it more like an emblem than a gesticulation, and the gesture continuum combines it with the emblems. Like an emblem, a point meets form standards—while all kinds of gestures can be used to indicate a locus, the extended index finger is standard in North American and Northern Europe; a flat hand is standard in some British Isle uses (Kendon 2004); and lip points are standard in Laos (Enfield 2001). All have in common depicting an iconic vector from a zero point or “origo” (Bühler 1982) to some target. The vector is the image, and cultures standardize different forms of it. This is one similarity of points to emblems.

The other similarity, less obvious but more profound, is how points and emblems relate to speech. While points and demonstrative pronouns (“this,” “that,” etc.) can synchronize (Levelt et al. 1985) and thus appear to be like gesticulations, in fact the timing is different and more like that of an emblem. The similarity appears when gesture and speech are asynchronous. For both points and emblems, the asynchronies are meaningful, and are so in both directions. Say “that” and then point; point and then say “that”; or say and point simultaneously—each combination is meaningful and different (the meanings seem metapragmatic, indicating how speech and gesture are being

used pragmatically). The same asynchronies and differences appear with emblems, for example, do the same experiment, with “OK” and the OK sign. Asynchronous gesticulations, on the other hand, are merely slovenly and not meaningful. Duplicate the experiment with the “rising hollowness” gesture of Figure 1.2, either synchronized with speech (“up through”) or not, and the meaning is the same until the asynchrony grows so great that the gesture and speech lose unity and seem to be repetitions. So pointing joins the emblem slot far from gesticulation as one of the most language-like of the non-sign language gestures. This has implications for gesture-orchestrated speech. Points and emblems play a metapragmatic part. They can act on their own or can join speech being organized by something else. The arrow in Figure 1.1, from the “Emblem/Point” slot to the “Gesticulation” slot, shows the possibility that the latter can absorb the former, and in this way endow the emblem or a point with the power of orchestration. At the same time, points bring a metapragmatic indication to “Gesticulation”; the arrow accordingly is two-headed. As a quick example, imagine a gesture with “he barrels up it,” in two versions that we compare; first, the fingers resting together in a relaxed pose as the hand rotates left and right and rises up: the meaning is co-expressive with “barrels up,” its deixis inherent to the gesticulation; next, the forefinger extended, the others folded into the palm (the classic point) while the hand rotates right and left and rises up, like the first: again co-expressive with “barrels up” with inherent deixis but with the addition of a specific point, the utterance not just about the cat’s motion in a certain direction but also calling attention to ascent itself as a discourse relevant dimension.

## 1.6 The conception of language and gesture in this book

The motto is, “abandon all presuppositions, ye who enter here.” Language and gesture as described here need to be grasped as a totality. The conception is non-reductive. It is important to think of words, grammar, sentences, etc. not separately, and gesture–speech unity not as something built out of “parts,” but to see gesture as the orchestrating force of the whole. If you set aside reductive expectations, you can easily follow the logic. You see language in a new (or perhaps an old but forgotten) way. Wilhelm von Humboldt, two centuries ago, saw language in the right way to follow the arguments here and avoid a hell of concepts that slide past each other without fixing into a form. I can do no better than start with Humboldt’s distinction between *Ergon* and *Energeia*:

“An important distinction ... kept reemerging”; this was Humboldt’s distinction between language as *Ergon*—language viewed as structure—and as *Energeia*—language as an “embodied moment of meaning located both in the organism and in the medium that the