Gesture and Thought

DAVID MCNEILL

S. 37 + 38

University of Chicago Press CHICAGO AND LONDON

2005

Table 2.1. Gesture rates (onsets per 1000 secs.) in fluent speech (from Nobe 1996).

	Butterworth & Beattie (Phases Unknown)	Nobe (Onsets of Preparation Phases)
During silent pause	280.1	321.7
During phonation	106.9	372.3

be made that a gesture (G) was initiated in 100 percent of all pause units, but in only 50 percent of all phonated units. However, it is also the case that almost twice as many gestures occur during phonation as pauses. Thus, because pause units of time are less frequent than phonation units of time, examining the initiation of gesture by comparing units of time can lead to a deceptive result.

Furthermore, the probability per unit of time of a gesture launch during a pause may not, in fact, be higher than during phonation. Using the same definitions of fluent speech and pause, Shuichi Nobe compared onsets in pauses and onsets in phonations in the narrations of six subjects delivering a narration of our standard cartoon stimulus, with the results shown in Table 2.1.

Contrary to the 3:1 predominance of initiations during pauses in the Butterworth & Beattie column, Nobe finds a slight predominance during phonation (nonsignificant). In Nobe's case, we know these are the onsets of preparations. The strokes to which the preparations led occurred during phonations to an even greater degree—90 percent of strokes did so.

To explain this difference between the studies, Nobe speculates that the original Butterworth & Beattie result could have been related to the specific speech situation from which their data come (tutorials). The hypothesis suggested by Nobe is that the dialogic situation resulted in a higher occurrence of "turn attempt–suppressing signals" (Nobe 1996, p. 45, quoting Duncan & Fiske 1977). A gesture with a turn-attempt suppression function is obviously more likely during a pause. Given the ambiguity of what the pause signifies, a listener may incorrectly think a turn has ended, and the speaker then uses a gesture to retain the floor. Such a gesture would automatically precede the linked speech—and hence both aspects of the Butterworth & Beattie result are explained by Nobe's hypothesis. This gesture strategy would not have been called for in the cartoon narrations that Nobe examined.

To sum up, there is no basis for claiming that gestures occur during pauses in speech to any particular degree, and the strokes, the meaning-bearing phases of the gesture, only rarely anticipate speech. Preparations, the first sign that the meaning to be presented in the stroke is becoming organized in the speaker's thought, do (necessarily) anticipate the speech segments with which the stroke will ultimately synchronize, and some of these anticipation onsets (less than half) occur during pauses. However, the extent to which gesture onsets occur in pauses can be grossly overestimated if such gestures in pauses are confused with other functions of gestures, specifically the use of gesture as a turn-attempt suppression signal.

THE LEXICAL AFFILIATE

A *lexical affiliate*, a term introduced by Schegloff (1984), is the word or words deemed to correspond most closely to a gesture in meaning. Schegloff introduced the concept to localize the moment at which, during a conversation, a lexical meaning was first put 'in play'. He observed that gestures tended to precede the words that lexically corresponded to them and that the gesture could therefore signal the introduction of the new meaning into the conversational stream before it surfaced in speech. Apart from the problem of whether it is the gesture preparation or stroke that anticipates the lexical affiliate, it is important, lest there be confusion, to distinguish this concept from speech-gesture co-expressivity, as this term is used here.

A lexical affiliate does not automatically correspond to the co-expressive speech segment. A gesture, including the stroke, may anticipate its lexical affiliate but, at the same time, be synchronized with its co-expressive speech segment. Whereas a lexical affiliate can be recognized by comparing gesture and word, co-expressive speech can be defined only in relation to the context of speaking (since such speech, along with the gesture stroke, is what the speaker is differentiating from the context as significant). It is possible that a co-expressive linguistic segment might be a lexical affiliate, but there is no necessity for it to be.

A clear illustration of the lexical affiliate/co-expressive speech distinction appears in Engle (2000). A gesture anticipated a lexical affiliate, consistent with Schegloff's original observation, but the immediate context of speaking suggests that the gesture and the *co-expressive speech* were actually synchronous. This is the example: Attempting to explain how a lock-and-key mechanism works, the subject said, "lift them [tumblers] to a // height, to the perfect height, where it [**enables**] the key to move," and appeared to turn a key as he said "enables." The lexical affiliate is "key" or "key to move" and the key-turning

gesture clearly occurred in advance of it. But from the vantage point of what would be newsworthy content in context, the synchrony of "enables" with the key-turning gesture embodies an idea this speaker might be expected to highlight—that by lifting the tumblers up, you are *able to turn the key*; and this thought is what the combination of a turning gesture plus "enables" captured.

TYPES OF GESTURES

Gesture classification plays little role in the dialectic analysis to follow; gesture *content*, regardless of type, is far more significant. At the end of this section I will explain what replaces gesture categorization in the upcoming dialectic analysis. Nonetheless, it is useful to describe the diversity of gestures. Systems for classifying gestures have been proposed for decades, and many gesture workers have provided ideas for this (e.g., Efron 1941, Ekman & Friesen 1969). The systems differ mainly in the number of categories they use. I myself participated in creating one such system (McNeill & Levy 1982, McNeill 1992). The search for categories, however, now seems misdirected. Most gestures are multifaceted—iconicity is combined with deixis, deixis is combined with metaphoricity, and so forth. Rather than categories we should think in terms of dimensions. One practical effect of shifting to a dimensional framework is that gesture coding is simplified. Coders are not forced to cram a gesture into a single box—say, the iconic—when it also seems deictic, and so forth. Difficulties with intercoder reliability are often traceable to such forced choices. A new dimensional approach was first described in Duncan et al. (1995).

An advantage of the dimensional approach is that we can combine dimensions without being forced to posit a hierarchy of them. A given gesture can have its own loadings across dimensions, and in this way can contain space for all the meanings—semantic, pragmatic, and poetic—that it may embody. I will explain this idea by reviewing our own efforts, and then describe the problems that arose from them.

The Iconic-Metaphoric-Deictic-Beat Quartet

The dimensions I will describe below are based on the original categories we described in McNeill & Levy 1982 and McNeill 1992. Inspired by the semiotic categories of C. S. Peirce (1960), Elena Levy and I proposed a classification scheme with four categories: *iconic, metaphoric, deictic,* and *beat.*

Iconic: Such gestures present images of concrete entities and/or actions. They are gestures in which the form of the gesture and/or its manner of execution embodies picturable aspects of semantic content (aspects of which are also present in speech). An example is the gesticulation example described above (Figure 2.2 and Chapter 1). Various aspects of the gesture—form, hand (standing for the character's hand, and opening it as if grasping an object with some thickness), trajectory (a curved path), direction (backward), etc.—correspond to aspects of the event, a character bending back a tree. The gesture as a referential symbol functioned via its resemblance to this event, iconically. Speech and gesture were not identical, however. As noted before, the gesture was made with a single hand, reflecting the fact that the tree was fastened down at one end. Speech did not mention this fact; being fastened at one end was implied by the use of "it" (a tree), but this was not built into the verbal description itself. Speech and gesture were *co-expressive* in the sense that they portrayed the same event but were not redundant, since each articulated its own aspects of it.

Metaphoric: Gestures can also present *images of the abstract.* Some gestures involve a metaphoric use of form—the speaker appears to be holding an object, as if presenting it, yet the meaning is not presenting an object but rather that she is holding an 'idea' or 'memory' or some other abstract 'object' in her hand. There is an iconic component (the form of the gesture resembles holding an object) and a metaphoric component (holding or presenting something is a metaphor for presenting a meaning). Other gestures involve a metaphoric use of space. A speaker, for example, divides the gesture space before him according to an appearance-reality dimension of morality being attributed to story character (McNeill 1992, p. 155; cf. Cienki 1998). The division is not necessarily along the lines of such a cultural stereotype as good to the right (cf. 'dexterous') and bad to the left ('sinister'). The metaphor may well assign good and bad in the opposite direction. In a metaphoric gesture, an abstract meaning is presented as form and/or space, but not necessarily in terms of stereotypic linkages. The topic of metaphoricity will be taken up in greater detail below.

Deictic: Although the prototypical deictic gesture is the [illegible] hand with an extended index finger, almost any extensible body part or held object can be used for pointing. If the hands are employed otherwise we can improvise pointing with our heads, noses, elbows, feet, etc., as well as with abstractions like the 'eidola' or gaze of ancient ophthalmology. Indeed, some cultures prescribe deixis with the lips (Enfield 2001). Deixis entails locating entities and actions