

The production of gesture¹

KEVIN TUIITE

Speech, the utterance of linguistic forms, is but one component of a multichannel complex of physical activities which, in one way or another, signify. In this paper I will discuss one particular component of this complex which reflects an especially close correspondence to the verbal channel: gesticulation, or gesture, the often depictive hand movements which accompany speech. The article is divided into three parts. The first section will be given over to a discussion of the various classes of nonverbal activity which researchers have recognized; in the second part I will briefly discuss the semiotics of gesture; and in the third part I will propose a model of gesture production.

Classes of nonverbal activity

In an important article published in 1969, Ekman and Friesen discuss five major classes of nonverbal activity, labelled emblems, illustrators, regulators, affect displays, and adaptors. These are ranked by degree of awareness typically accorded them by interlocutors. Emblems are in a sense lexicalized gestures — they have highly specific meanings, can be interpreted in the absence of speech, and vary significantly from culture to culture. Familiar examples from our culture include the 'OK' sign and the shoulder shrug with upturned palms ('beats me'). Illustrators correspond to what I call gesture; we will turn to them shortly. Regulators play an important role in the process of conversational interaction, but we seldom notice them under normal conditions. These also are highly culture-specific: the head nods, gaze-direction changes, shifts in interpersonal distance and body posture, eyebrow activity, and so forth that provide information on the status of the interaction and regulate turn-taking (Duncan and Fiske 1977). Affect displays primarily involve the face. According to Ekman and Friesen (1969: 70–81), there is evidence that certain 'primary affects' (happiness, fear, surprise, anger, disgust) are universally associated with particular

facial expressions. Cultures differ in the degree to which affect displays are suppressed or exaggerated in given contexts. The last category, adaptors, includes a fairly diverse assortment of activities, the origin and function of which have been a topic of investigation at least since Freud: grooming, scratching, licking one's lips, fiddling with small objects, rubbing one's face and so forth. Though the behavioral repertoires subsumed under each of Ekman and Friesen's five categories probably all have a semantic component — Birdwhistell (1970: 183) elevates this hypothesis to the status of an axiom — there are qualitative as well as quantitative distinctions to be drawn concerning the correlation between kinesics and speech. Kendon (1980) reports a relevant experiment: twenty subjects were shown a film of a New Guinea native addressing a crowd, with the sound turned off. The subjects were in agreement that 'arm extensions and elaborate movements in the space in front of the body' belonged to the speech performance, while 'self-touching, postural movements and movements involved in the manipulation of an axe were regarded as being quite separate'.

Gesture

The nonverbal activity interpreted as being part of the speech performance corresponds to what Ekman and Friesen termed 'illustrators', and will be referred to here as 'gesture'. Several investigators, among them Efron (1941), Kendon (1980), and McNeill (1985), have further categorized gesture in terms of its relation to the content of the cooccurring speech. The categories devised by Efron for describing gesture have formed the basis for those employed by more recent investigators:

A gestural movement may be 'meaningful' by [a] the emphasis it lends to the content of the verbal and vocal behavior it accompanies, [b] the connotation (whether deictic, pictorial or symbolic) it possesses independently from the speech of which it may, or may not, be an adjunct. In the first case its 'meaning' is of a logical or a discursive character, the movement being, as it were, a kind of gestural portrayal, not of the object of reference, or 'thought', but of the course of the ideational process itself (i.e. a bodily reenactment of the logical pauses, intensities, inflections, etc. of the corresponding speech sequence). This type of gesture may in turn be [a] simply *baton-like*, representing a sort of 'timing out' with the hand of the successive stages of the referential activity, [b] *ideographic*, in the sense that it traces or sketches out in the air the 'paths' and 'directions' of the thought-pattern. The latter variety might also be called *logico-topographic* or *logico-pictorial*. In the second case the 'meaning' of the gesture is 'objective', and the movement may be [a] *deictic*, referring by means of a sign to a visually present object (actual pointing), [b] *physiographic*, depicting either the form of a visual object or a spatial

relationship (iconographic gesture), or that of a bodily action (kinetographic gesture), [c] *symbolic* or emblematic, representing either a visual or a logical object by means of a pictorial or a non-pictorial form which has no morphological relation to the thing represented. (1941: 69–70)

In this section, each of Efron's categories of gesture will be discussed in turn.

Beats

The simplest class of gesture corresponds to Efron's baton gestures (termed 'beats' by McNeill). Unlike pictorial gestures, which tend to be motorically elaborate and performed in central gesturing space by the speaker's dominant hand, beats 'are typically small simple movements that are performed more rapidly [than representational gestures — KT] at or near the rest positions of the hands' (McNeill 1985: 359), often by the speaker's non-dominant hand (Stephens 1983). According to McNeill, beats 'have no propositional content of their own. The gesture is an abstract visual indicator. Hence beats are particularly appropriate for emphasizing discourse-oriented functions where the importance of a linguistic item arises, not from its own propositional content, but from its relation to other items' (1985: 359). Beats can accompany emphasized or contrasted lexical items and shifts in discourse topic, fulfilling a function similar in many ways to that of intonation. (I will have more to say about this later.)

Ideographic gestures (spatial beats)

The textual fragment² given below comes from an extranarrative digression in the description of a film the speaker has just seen. Most of the gesticulation is in the form of beats: motorically-simple, low-energy gestures. The speaker is evaluating the relative moral statuses of an ex-convict trying to perpetrate blackmail on two 'respectable' people who, in turn, are attempting to cover up a killing by blaming it on the blackmailer. This sense of contrast is conveyed through gesticulation. Through the positioning of some of the beat gestures — the ones marked as 'spatial' — the speaker diagrams the important actors in his narrative. In the passage quoted earlier, Efron referred to a class of gestures (ideographics) which 'trace or sketch out in the air the paths or directions of the thought-pattern'. The spatial beats in our example fulfill this kind of function, in that the same fundamental metaphor underlies them. The speaker has in effect represented the plot of the discourse as a spatially-extended object,

with distinct loci corresponding to distinct plot components — characters, episodes, time segments, props, etc. Two such diagrams occur in Example {1}, both indicating characters. The first [lines 2–7] is a binary one with Tracy to the right, the killer and her boyfriend to the left; the second diagram [lines 10–14] represents three parties (the above two plus the victim). As is often the case, relative location is more important than absolute location — notice that the ‘good guys’ are associated with the left-side locus in the first diagram and the right in the second.

These gestures reflect only a single semantic feature — spatial locus — which enables them to contrast meaningfully with one or more gestures of the same kind. Since the assignment of spatial features is driven by purely structural considerations (that is to say, the primary consideration is that the gestures be executed in distinct spatial loci, to reflect the underlying opposition of narrative elements, linguistic items, etc.), spatial beats are fundamentally different from superficially-identical deictic gestures, where the choice of locus is dictated by existing objects or people.³

Example {2} contains spatial gestures connected with two very distinct levels of the discourse. The leftward movements in gestures (c), (d), (f), (g), and (k) are associated with reference to a shop owned by one of the characters in the story being narrated. As such they derive their significance from their opposition to the rightward gesture in (1), connected with mention of the apartment where another character lives. Imbedded within this sequence of narrative-related spatial gestures are two metalinguistic ones, (h) and (i), connected with a correction of lexical choice made by the speaker: the locus on the right associated with the wrong word is opposed to that on the left associated with the correct one.

Anaphoric (inertial) beats

Not infrequently the hand will retain its configuration after an iconic gesture has been completed, and one or more beat-like gestures will be performed without a change in hand-shape. These so-called ‘inertial beats’ usually serve a cohesive function, linking the gestures accompanying a particular thought. The repetitions of the hand with the wiggling fingers in Example {2} are a case in point. In gesture (c) this combination of handshape and motion appears to be associated with nearness, almostness.⁴ None of its recurrences are motivated in this fashion. Rather, they serve to connect the successive references to Alice’s father’s shop within one sentence-like unit. The last iterations, gestures (j) and (k), accomplish the additional function of reestablishing the train of discourse after a metalinguistic interruption.

- {1} 1. / [and i* it flips]_a [back]_b and forth because
 a. *metaphoric*: BH, palms flat, flip back and forth in alternation
 b. BH relax
2. /]_c [then]_d Tracy [has got a]_e criminal record
 c. RH up
 d. *spatial*: RH points R
 [R locus = Tracy]
 e. RH loops up and down
3. / [the police]_f are looking for him
 f. RH up and down
4. [he was seen]_g at the* the crime [that night]_h
 g. RH moves up and down
 h. RH up and down
5. / [they can pin it]_i on him
 i. *spatial*: RH loops up and L; then down
 and R [L locus = 'good guys'; R locus = Tracy]
6. [but he's uh]_j he's [got*]_k
 j. *spatial*: RH loops up and L, pointing down
 and L [L = 'good guys']
 k. *spatial*: RH moves R, palm up
 [R = Tracy]
7. [he's blackmai-]_l [ling them]_m
 l. *spatial*: RH up and L, pointing L
 [L = 'good guys']
 m. RH down, relaxed
8. / [so i* it's a* th* the uh]_n
 n. *metaphoric*: BH, slightly cupped, palms facing, move up and
 down in alternation, synchronized with speech
9. [everyone's morals]_o [are]_p [very]_q ambiguous /]_r
 o. continues as in (n)
 p. BH hold
 q. LH a bit forward
 r. BH relax

Diagram I:
 Right = Tracy
 Left = 'good guys'

10. ['cause]_s [they're sup-]_tposed to be the good
[guys]_u
s. BH up and down
t. *spatial*: BH cupped, move R
[R locus = 'good guys']
u. BH relax, palms down
11. [but she]_v really did kill [him/]_w
v. *spatial*: RH moves L and points L
[L locus = victim]
w. RH retracts, closes to fist
12. and [he's a]_x bad [guy]_y
x. *spatial*: BH extend forward, palms up
[forward locus = Tracy]
y. BH drop slightly
13. [but he really]_z didn't kill him
z. *spatial*: BH move L, RH points L
[L = victim]
14. [/and shouldn't]_{aa} be [blamed]_{ab}
aa. *spatial*: BH move R, pointing forward
[forward = Tracy]
ab. BH drop to lap and relax

Diagram II:
Right = 'good guys'
Forward = Tracy
Left = victim

Deictic gestures

Efron's second major class of gestures are those he calls 'objective', divided into three classes corresponding to Peirce's three types of semiotic relation: icons, indices (deictics), and symbols. Deictics, or pointing gestures, are usually foregrounded by framing devices (e.g., 'she went over there [point]'), but need not be. Deictics not explicitly intended to draw the attention of the listener often occur when reference is made to an entity or event whose associated locus is nearby.⁵ In Example {3}, excerpted from a Georgian-language narration recorded in our laboratory, the listener asks for clarification concerning the reference of a noun phrase (*es k'aci* 'this man') used by the narrator. In doing so he spontaneously points in the direction of a spot in empty space which the speaker had previously associated, through the gesture in line (2), with the noun phrase in question.

- {2} 1. and she says uh [Alice says]_a something to the effect
 a. LH (on leg): slight beat
2. [I live]_b [just around the]_c corner
 b. LH on leg uncurls
 c. *metaphoric + spatial (narrative-related)*: LH rises, fingers spread and wiggle; moves slightly L metaphor ('just') + spatial (L locus = Alice's father's shop)
3. [I* y'know]_d you know that news stand
 d. *inertial beat*: LH rises, swings further L, fingers wiggle (cohesive with gesture c. above)
4. [that's where my father]_e/
 e. *inertial beat*: LH uncurls, rises slightly, fingers wiggle (cohesive with gesture c. above)
5. that's [my father's news]_f[stand]_g/
 f. *spatial*: LH swings up and L; fingers spread
 g. *spatial + inertial beat*: LH up and L; fingers wiggle (cohesive with gesture c.)
6. not [news stand]_h [news shop]_i/
 h. *spatial (metalinguistic)*: LH points forward and R
 i. *spatial (metalinguistic)*: LH points L and forward
 (correction of lexical error:
 R locus = word 'news stand'
 L locus = word 'news shop')
7. [that's my]_j father's news shop — [that's/_k where I live/
 j. *inertial beat*: LH opens; rises slightly, fingers wiggle
 k. *inertial beat*: LH swings up and L, wiggles slightly, moves R (cohesive with gesture c.)
8. and uh//he says well really I [live right here]
 1. *spatial (narrative-related)*: RH forward and R; rotates palm up (R locus = artist's apartment)

referential chain (maintained by inertial beat gestures)

spatial contrast (metalinguistic)

spatial contrast (narrative-related)

- {3} 1.DK: *magram es k'aci sinamdvile-shi k'ino-shi*
 but this man: nominative reality-in movies-to
ar c'avida
 not he: went
 'But in reality this man didn't go to the movies'.
 2.DK: [*garet*]_a *aris kucha-shi/ rest'oranis c'in*
 outside he: is street-on restaurant: genitive before
 'He's outside on the street, in front of the restaurant.'
 a. *spatial*: head nods R
 (anaphoric; indicates Frank at R locus)
 3. KT: [*es p'ir-*]_b *veli [k'aci?]*_c
 this first man: nominative
 'This first man?'
 b. *spatial*: LH forward, points forwards, then L
 (point to locus set by DK to her R [= Frank])
 c. *spatial*: LH points L
 (similar to above)

Physiographic (iconic) gestures

The class of gestures termed 'iconic' by McNeill correspond to Efron's physiographics. These have been further divided by Stephens and Tuite (1983) into two types, based on the nature of the mapping of referent onto gesturer's body. In type I iconics the speaker's hands (and other body parts, if used) represent the corresponding body parts of some character in the narrative. Consider Example {4}:

- {4} 1. and uh he takes her dress
 [which has been]_a[lying]_b across the top of the screen/
 a. *spatial*: RH points R
 b. *iconic*: RH arcs up and L; retracts
 (depiction of dress lying over screen)
 2. and [throws it]_c into another part of the apartment/
 c. *iconic*: LH (fist) up, swings inward toward self; then swings
 forward and L, opens, fingers spread, palm down
 (artist throwing dress; also, trajectory of dress to L is opposed to
 the location of the screen [and Alice] on the R)
 3. and refuses [to get it for her]_d when she asks for it/
 d. *iconic*: LH closes to fist; pulls straight back toward face
 (artist keeping dress from Alice [metaphoric use of iconic gesture])
 4. instead [he sits]_e [at the piano]_f

- e. RH swings forward and R; uncurls
- f. *iconic*: RH fingers ripple (as though playing a piano)
(artist playing the piano)

The speaker's hands can be said to be throwing a 'discourse dress' or playing a 'discourse piano' much as they would a real one. In each case the hands are performing a stylized manipulation repertoire associated with an object mentioned in the accompanying speech. To use a theatrical metaphor, the speaker's hands are playing the role of the hands belonging to the person denoted by the grammatical subject of the clause. Now look at Example {5}, from the McNeill and Levy corpus:

- {5} [and as he's coming up and the bowling ball's]_a coming [down]_b
 a. LH (= ball) descends while RH (= 'he') rises
 b. BH loop out and back
 [he swallows it]_c
 c. LH (= ball) descends again, while RH (now = 'his' mouth) rises
 and engulfs LH

This sequence of gestures (cited in McNeill 1985: 352) occurred while the subject was retelling the plot of a Tweety Bird and Sylvester cartoon. At one point Tweety drops a bowling ball down a drainpipe Sylvester is crawling up inside of; the latter swallows the ball and rolls back down the pipe. One could say that the narrator's right hand is playing Sylvester and the left hand the bowling ball. Note that the semiotic mapping is less direct than in {4} — the speaker's hands represent entire characters, rather than characters' hands. These are instances of type II iconics.

Symbolic (metaphoric) gestures

Two gestures labelled 'metaphoric' were described in Example {1}, and are repeated here:

- (a) /[and i* it flips]_a back and forth
 a. *metaphoric*: BH, palms flat, flip back and forth in alternation
- (b) [/so i* it's a* th* the uh everyone's morals]_b are very ambiguous
 b. *metaphoric*: BH, slightly cupped, palms facing, move up and down in alternation, synchronized with speech

They differ from ordinary iconics in that their form 'depicts the vehicle of a metaphor. The gesture is iconically related to this vehicle, not to the meaning, or tenor, of the metaphor' (McNeill and Levy 1982: 274). Metaphoric gestures involve the representation of abstract notions through concrete forms. Gesture (a) corresponds closely to the lexical metaphor

occurring simultaneously in the speech channel ('it flips back and forth'), while gesture (b) reflects a metaphoric image associated with the word 'ambiguous': the comparison of two interpretations, reified as concrete objects whose weight one compares by hand. Metaphoric gestures are frequently used by adult speakers, and are by no means restricted to our culture. Numerous instances of this type of gesticulation have been observed in the speech of the Turkana of eastern Kenya — for instance, Example {6} (Stephens and Tuite 1983):

- {6} *todarite ngitunga lu na [kilna yoka]_a — [nith!]_b*
 extract: 3PL people this here knowledge our_{incl}
 'These people <Europeans> will extract all of our knowledge: *nith!*'
 a. LH, cupped, fingertips touching, near L side of head
 b. LH moves rapidly away from head, fingers spread apart;
 LH moves up and L

The image portrayed here is one of knowledge — transformed into a concrete object — being pulled from the speaker's head and flying away.⁶

The semiotic nature of gesture

The fundamental assumption underlying the gestural taxonomies of Efron and McNeill is that, in some sense or other, the gesturer is transforming stored information (internal representations) into patterned movement. Such an assumption arises through a variant of the hermeneutical spiral: (a) certain instances of functionally non-transparent nonverbal behavior bear a striking resemblance to actual objects or activities; (b) stored information derived from manipulations or actions — performed or witnessed — is presumed to underlie the behavior in (a); (c) less obviously iconic gestures are interpreted on the assumption that internal representations underlie them as well. Let us examine each stage of the spiral more closely. The first step presupposes that observers associate typical functions with various manual activities. Gestures, then, are distinguished by their lack of motivation in this sense. In the gesture cited here, for example, the gesturer performs some of the movements associated with lifting a crate without a crate actually being present.

[he takes a crate]
 both hands rise and stop, apart, palms facing each other

The impression of resemblance between gesture and 'functionally transparent' activity is often augmented by the accompanying speech, as

in the instance just cited. Ricoeur once wrote that a text can be understood 'to the extent that it is not closed on itself, but open to the world which it redescribes and refashions' (1978: 155). The second step in our spiral is taken when the observer performs the operation described by Ricoeur in reverse — that is, interprets kinesics of the sort described in (a) as refashionings of the world as perceived by the gesturer. (Put another way, one sees gesture as a kind of text.)⁷ Iconic gesture originates in the abstraction of sensorimotor-based features (types) from tokens of activity observed or engaged in by the gesturer. On this basis, representations can be formed. By allowing for greater degrees of selectivity in the abstraction and reflection process, one can hypothesize that less readily interpretable gestures can arise from internal representations of essentially the same nature as those underlying blatantly iconic gesticulation.

Gesture as representation

Merleau-Ponty observed that 'I become involved in things with my body, they coexist with me as an incarnate subject, and this life among things has nothing in common with the elaboration of scientifically conceived objects' (1962: 185). As Rudolf Arnheim argued, the process of representation building through which we become 'incarnate subjects' involves continual, active give-and-take with the world. Arnheim took a particular interest in children's drawings, which for him reflect

the kind of problem solving of which I spoke earlier as the intelligence of perception. ... Thinking requires more than the formation and assignment of concepts. It calls for the unravelling of relations, for the disclosure of elusive structure. Image-making serves to make sense of the world. (1969: 257)

Now let us consider gesture from this perspective. Arnheim spoke of descriptive gestures as 'the forerunners of line drawings' in that

the portrayal of an object rarely involves more than some one isolated quality or dimension, the large or small size of the thing, the sharpness or indefiniteness of an outline. ... By the very nature of the medium of gesture, the representation is highly abstract. What matters for our purpose is how common, how satisfying and useful this sort of visual description is nevertheless. ... The gesture limits itself intelligently to emphasizing what matters. (1969: 117)

So, iconic gesture, like children's drawings, does not represent 'scientifically conceived objects' as such, but rather arises from the process of perceptual thinking, the abstraction of structures, relations, and features. The interpretive method consonant with Arnheim's perspective would be

evaluating gestures as reflections of features selected from some richer underlying image. Semantically non-transparent gestures would be explained away as highly abstract representations, abstract beyond the point of recoverability.

Gesture and communication

This question of recoverability merits an extended comment. It is clear from our corpus of recorded discourses that speakers make a distinction between gestures they want their interlocutors to attend to and 'ordinary' gesticulation. Gestures of the first type are specifically marked: by gaze and posture shift, and by framing devices in the accompanying speech ('then the guy went like this: [gesture]', 'her cheeks puffed out like this [gesture]', 'un poisson grand comme ça: [gesture]'). As Kendon (1983) noted, this kind of gesticulation serves to complete an utterance, for which reason the speech will often be interrupted while the gesture is enacted. The interlocutor responds differently to framed gestures as well, shifting gaze direction from the speaker's face to his or her hands.⁸ Such gestures stand in distinct contrast to those we produce quite unconsciously while speaking on the telephone to an interlocutor who clearly cannot see them. Cosnier (1982: 278) reports that subjects seated back-to-back continue to gesticulate while talking, though the frequency decreases somewhat.

This evident lack of responsiveness to feedback from the discourse situation would indicate that as a semiotic activity, gesture is *production-centered* rather than reception-oriented, in which respect it differs markedly from most other systems that signify. In other words, *the activity of gesture primarily occurs for the 'benefit' of the speaker/gesturer, and not for the listener*. It follows that, unlike most semiotic systems, gesture is not subject to the pragmatic constraint of surface recoverability — that is, the requirement that the information underlying the sign not be degraded, or pared down through abstraction, beyond the point that certain features are no longer retrievable by the recipient. The lack of this constraint is easily observed in the case of spontaneous gesticulation. Under controlled conditions, when subjects are describing cartoons shown to them by experimenters, a sizeable proportion of their gestures are uninterpretable, even when viewed in slow motion. In a system constrained by surface recoverability, one can speak of successful (or unsuccessful) instances of sign presentation. Except for the case of framed gestures mentioned above, the notion of 'success' in enabling the interlocutor to reconstruct the gesturer's representation cannot be applied coherently to the description of gesticulation.

Gesture production

For the reasons cited above, the analysis of gesture presented here will be primarily in terms of production rather than communication in the more typical sense. The model of gesture and speech production proposed in this paper is summarized in Figure 1. On the basis of observations of speech and gestural behavior and their correlation, two internal components have been postulated: an internal representation, and a 'visceromotor' component, which determines the pace and timing of muscular activity. We will discuss each of these components in turn.

Internal representation

According to McNeill, the representation underlying speech production consists of 'an image and an inner speech symbol, [which] are generated at the same time. Inner speech symbols have syntactic implications.⁹ These implications and the imagistic properties that the gesture exhibits are generated at the same time, and can develop internally together' (1985: 367). Inner speech symbols are based on words and other linguistic symbols 'impregnated with what Vygotsky calls sense, a synthesis of several meanings into a single symbol and an enrichment of the word's meaning. The sense of inner speech — this enrichment of the mental operations attached to the word — according to the evidence of spontaneous gestures consists, in part, of images' (1985: 368). McNeill backs up his leading idea — that gesture and speech are two separate surface manifestations stemming from a common underlying representation — with detailed analyses of normal adult gesture-and-speech fragments as well as evidence from children and aphasics (summarized in McNeill 1985). Especially significant are instances of gesture/speech complementarity in representing information. In a recent article, he reports an experiment in which subjects were shown a videotape of an experimenter narrating the stimulus of choice in our laboratory (a Tweety Bird cartoon). The subjects were not shown the cartoon itself. Afterward they retold the story as they recalled it from this videotaped presentation. Example {7} is an excerpt from one such secondhand narration (McNeill 1986: 303):

{7} *Experimenter's narration:*

1. Sylvester ingests the bowling ball,
2. comes back out the drainpipe,
3. and goes [rolling down the street]
 - a. LH moves up and down while swinging to the left

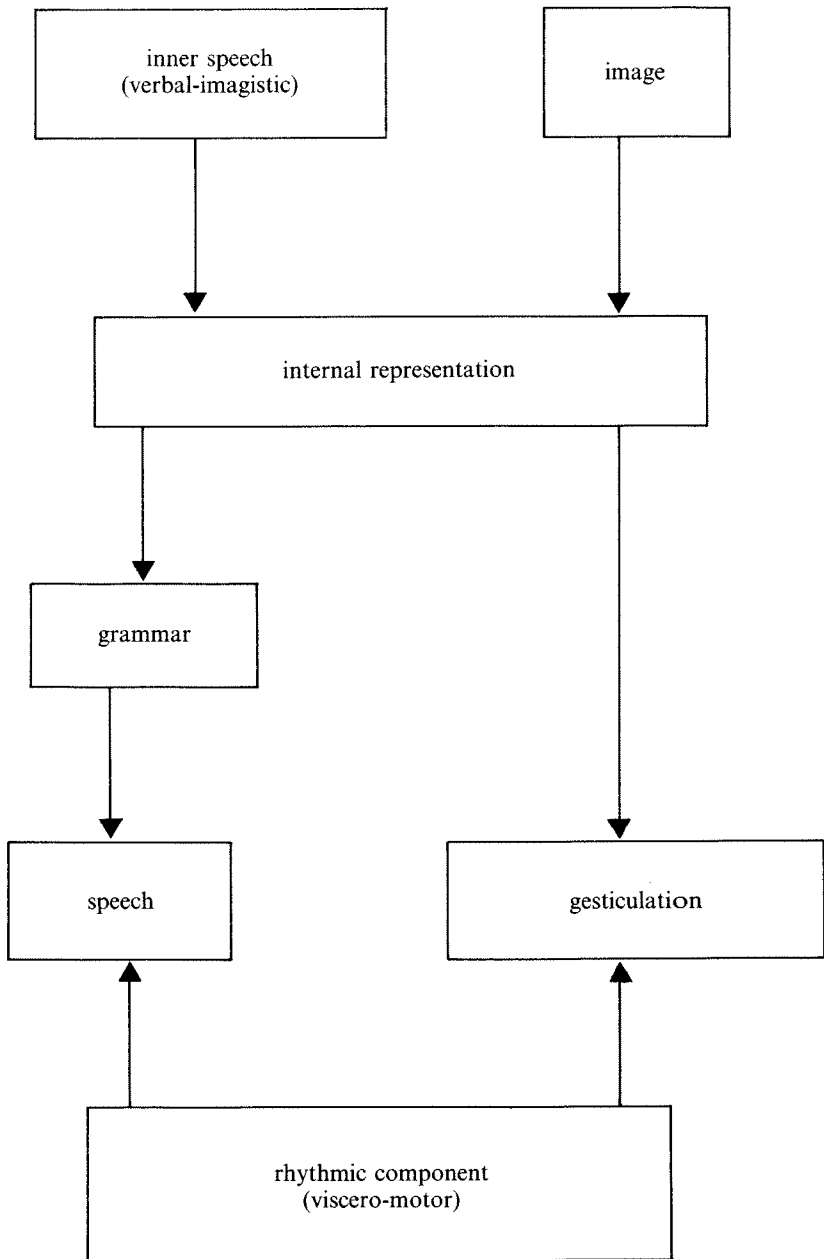


Figure 1. *A model of the production of speech and gesticulation.*

Subject's narration:

1. Sylvester ate — ingested the bowling ball,
2. falls back down to the bottom of the drainpipe,
3. and [bounces down the street]
 - a. RH sweeps smoothly to right

The information picked up from the experimenter's gesture concerning the manner of movement was expressed by the subject through lexical choice ('bounces'). Conversely, the subject produced a smooth gestural trajectory reflecting only the direction of Sylvester's movement; the bouncing motion was left out.

Discourse, for McNeill, is a process of unpacking global-synthetic representations — associations of images and inner-speech symbols — into external speech and gesture.¹⁰ Though by its nature a linear, symbolic, digitized medium, speech can reflect some of the imagistic flavor of its underlying representation through lexical choice and the iconicity implicit in syntax (McNeill 1985: 369). But it is gesture, the synthetic output medium *par excellence*, that illustrates 'thought as something whole that includes all of its parts simultaneously. Gestures combined with the concurrent linguistic evidence thus show as it is happening (in "real time") the full development of speech from a unitary undivided thought to an extremely articulated utterance' (1987: 93).

Rhythmic component

Recent research on gestural activity, especially that of Kendon and our group in Chicago, indicates that speech production consists in the presentation of signs — linguistic and gestural — accompanied by a more or less regular rhythmic pulse which is, like the semiotic content, expressed through both verbal and nonverbal channels. We will begin with some observations concerning the correlation between peaks of energy output in speech and in gesture. In a recent paper on the uses of intonation in speech, Bolinger concluded that 'the gestural domain is where intonation truly belongs' (1982: 18). Careful measurements of the temporal contours of nonverbal activity cooccurring with speech lend support to this claim. Particularly noteworthy in this respect are two studies by Kendon concerning parallelisms in the hierarchical structure of spoken discourse, as marked by intonation patterns and pauses, and the structure of the accompanying nonverbal activity. Kendon partitioned the recorded speech into 'tone groups', defined as 'the smallest grouping of syllables over which a completed intonation tune occurs' (Kendon 1972: 186). The minimal gestural

unit is the so-called 'G-phrase', comprising a gestural stroke (the moment of peak effort) which may be accompanied by a preparation 'in which the limb moves away from its rest position to a position at which the stroke begins', and/or a recovery 'in which the limb is either moved back to its rest position or in which it is readied for another stroke' (Kendon 1980: 9). Kendon noted that 'an examination of the relationship in time between the nucleus of a tone unit and the stroke of its associated G-phrase shows that the stroke of the G-phrase is completed either before the tone unit nucleus, or just at its onset' (1980: 15). Most gestural strokes that precede the tonic nucleus do so by at most one syllable's length. Consider the discourse segment given as Example {8}, transcribed by L. Pedelty and J. Cassell. The text is segmented roughly into tone groups, with the nucleus (stressed syllable) indicated by italics. Most of the gestures in Example {8} are motorically-simple beats. Note how frequently the gestural strokes coincide or slightly precede these syllables. Further, the strokes occur at more or less regular intervals in this speech segment, as if their generation were governed by a kinesic pulse.

- {8} 1. [so he]_a [*dresses*]_b [up]_c
 a. <preparation> RH drops back and down and outward on wrist
 b. <stroke> RH rises and points up and L
 c. <stroke> RH makes emphatic movement down and back at wrist
2. [like a]_d [*doorman*]_e
 d. <hold> RH holds previous configuration
 e. <stroke> RH fingers extend, RH tilts back and rotates inward
3. [*goes up there*]_f
 f. <stroke> RH rises slightly and rotates up and out
4. // and um // and the [*grandma shouts*]_g
 g. <stroke> RH opens and rotates outward, fingers extend and curl
5. [out *oh*]_h
 h. <stroke> RH flips over palm down
6. [the the]_i [*bags are*]_j
 i. <preparation> RH moves up with fingers slightly curled
 j. <stroke> RH flips palm down, moves downward
7. [*right*]_k
 k. <stroke> RH moves up and down from wrist
8. [*beneath*]_l
 l. <stroke> slight quick movement of LH thumb
9. [*behind*]_m the door
 m. <stroke> RH turns sideways, flicks outward

Even when speakers are not gesticulating, in most cases the pulse continues to find a kinesic outlet: head, leg, or foot movements, twitching of the hand or fingers, raising of the eyebrows, etc.¹¹

The generation of different types of gesture is represented schematically in Figure 2. The kinesic base which underlies both gesture and the gesture-like component of speech (i.e., intonation) is represented as a rhythmic pulse. If expressed gesturally, the pulse peak corresponds to the stroke portion of the gesture. If expressed vocally, the pulse peak is represented by an intonational peak, which — presumably because of the more elaborate processing required by speech production — may be somewhat delayed relative to the gestural stroke.

The simplest form of gesticulation is the beat gesture. It is a kinetic realization of the underlying pulse, and almost entirely lacking in semantic content derived from the speaker's internal representation. Gestures may have a spatial component, which can be generated in either of two ways. In the case of a deictic gesture, the

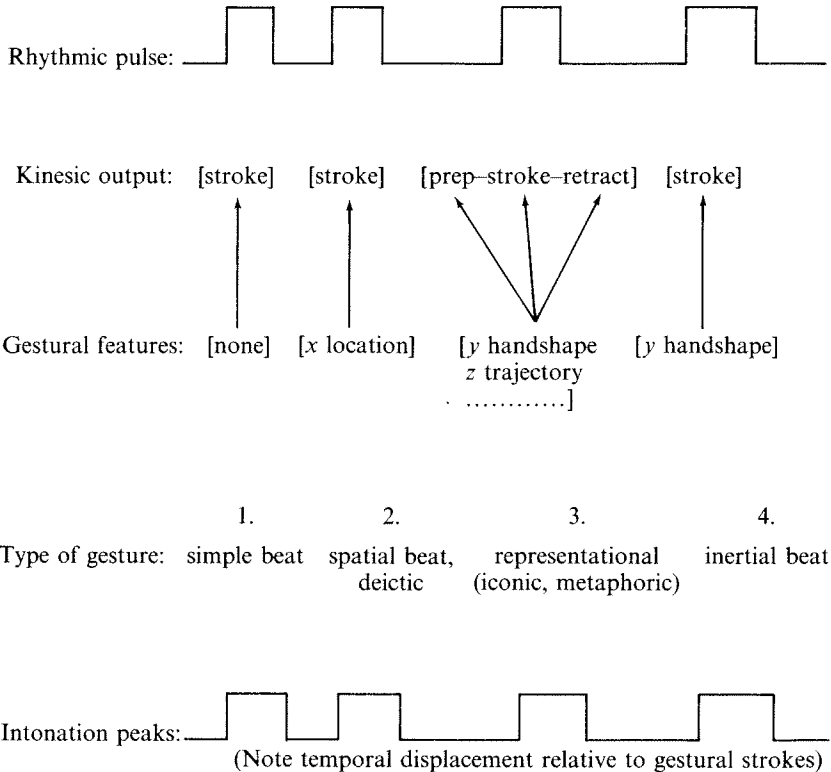


Figure 2. Superimposition of features upon rhythmic base.

movement is directed toward the actual location of a referent, or to a location which has been associated with the referent in some way (see note 5). The locus indicated by a so-called spatial beat gesture is purely abstract, deriving its significance only by its distinctness from the locus of one or more other gestures of the same type. In this way spatial beats reflect some sort of opposition between elements of the narrative or even of the speech itself (as was illustrated in Example {2}).

Representational gestures (iconics and metaphorics) differ from beats in terms of both their semiotics and their motoric complexity. The components of the gesture are abstracted from an underlying image, which bears an iconic or symbolic (metaphoric) relation to the content of the speech. As was mentioned earlier, there is no 'recoverability constraint' which ensures that the gestural representation is interpretable by the interlocutor. As with beats, the stroke portion of iconic and metaphoric gestures tends to coincide with the nuclear syllable of the accompanying tone group. This leads me to believe that iconic and metaphoric gesticulation is built upon an underlying pulse as well, but with different types of features extracted from the speaker's internal representation.

After the enactment of a representational gesture, reference to some aspect of the situation represented can be maintained by beat gestures with retention of the handshape (inertial beats). Although the handshape may be complex, inertial beats are structurally simple: they are functionally comparable to anaphors in speech.

Loose ends

In an article on the structure of communicative behavior, Cosnier declared that 'toute parole est le fruit d'une énonciation corporelle, et tout langage est langage du corps' (1982: 302). This paper represents an attempt to formulate a model of how gesture, the component of the 'énonciation corporelle' most closely coordinated with speech, is produced. According to the model proposed here, the speech process begins with an internal representation formed from images and verbal-imagistic inner speech. The content of this representation is distributed to the verbal and gestural channels, and is operated upon in correspondingly different ways. One significant difference between the verbal and gestural channels is the lack of any sort of surface-recoverability constraint upon the latter. The abstraction of features from the internal representation is not restricted by considerations of comprehensibility to the interlocutor. The realization of the output, verbal as well as gestural, is synchronized with a pattern of rising

and falling energy levels, a sort of rhythmic pulse. This pulse is reflected in the intonational component of speech, and the timing of the gestures.

The gesture-production model presented here is provisional; it is intended as a starting point for further research. I will mention here two important areas in which more work needs to be done.

The function of gesture

If it is indeed the case that the spontaneous production of gesture is not oriented toward communication in the usual sense, how has it come about that such an activity is so ubiquitous in human speech behavior? At least two proposals have been offered to explain the occurrence of gesture during speech. Hewes (1973) and Kendon (1975) have hypothesized that spoken language developed relatively late in the course of human evolution, and that the first well-elaborated systems of communication were gestural. Gesture would represent the vestiges of this earlier communicative system. According to Kendon,

The fact that gesticulation tends to anticipate speech, that speech may be disrupted though concurrent gesticulation continues to go forward smoothly, and the fact that on occasion a gestural response may be given first before any speech whatever is begun, does perhaps provide a hint that the gestural channel is easier and more readily called upon, that the process by which an idea is transformed into public behavior of a reportive character is more swiftly accomplished gesturally, that there are fewer steps to the process than there are when formulation of the 'idea' into speech is to occur. The least we can say about this is that we would not, perhaps, expect a more elaborate and time-consuming method of utterance to be the one that was first developed in language evolution. (1975: 367)

If gestural communication preceded speech, it may well have been more linguistic in nature than spontaneous gesticulation of the sort discussed here. The 'home sign' communicative systems developed by several deaf children born to hearing parents — who were exposed to no language model, either spoken or signed — have structural features which are characteristic of human languages, and which are conspicuously absent in gesture (Goldin-Meadow 1982; Goldin-Meadow and Mylander 1983). If the language capacity evolved before the changes necessary for articulate speech occurred in the hominid vocal tract, then early human gestural communication would probably have been more like home sign than gesture.

A much different explanation has been proposed by Cosnier, who sur-

mises that gestural activity in some way assists the speaker in the process of encoding an underlying representation into the verbal modality.

La valeur énergétique (ou 'économique') de cette activité motrice est encore loin d'être éclaircie mais plusieurs constatations amènent à penser qu'elle est fondamentale pour tous les locuteurs, tout en s'organisant pour chacun d'une façon idiosyncrasique dans ses rapports avec l'activité verbale et l'activité végétative. ... La gestualité servirait au locuteur à ordonner son message en fournissant des repères temporo-spatiaux à sa traduction parolière. (1982: 281)

Gesticulation would then be part of the process of modality transfer: from verbal-imagistic internal representation to the coordinated muscular activity of speaking. The underlying images, first of all, are derived from the body's encounter with the world, and are not only visual — they may also include auditory, tactile, and kinesthetic elements. The body is likewise involved in the formation of concepts. As studies of metaphor have shown, we draw upon concrete visual and kinesthetic experiences in the creation of metaphors to represent abstract notions (Lakoff and Johnson 1980). Even after the verbal metaphor has been 'frozen' into a rote-memorized lexical item, its imagistic counterpart often remains alive for speakers, as shown by gesture. Just as the body is intimately involved in the activity of perceiving and representing the world, so it is involved in the process of acting upon these representations in order to communicate. If gesture appears to be an epiphenomenon from a strictly communication-theoretic point of view, it is nonetheless a necessary component of the process.

Information pick-up from gestures

While the semantic content of individual gestures is often not retrievable, there are indications that interlocutors do pick up at least some information from the (spontaneous, non-foregrounded) gestures of their speech partners. Examples {3} and {7} illustrate how a conversational interactant may reflect by speech or gesture information picked up from the interlocutor's gesticulation.¹² Another indication that gestural behavior is attended to is the fact that cultural differences in regard to gesture do exist. Our preliminary observations of the spontaneous gestures produced by native speakers of such languages as Chinese, French, Georgian, and Turkana indicate that all four types of gesture described in this paper are used. The differences appear, as Efron (1941) noted in his comparative study of Italian and East European Jewish nonverbal behavior, in what might be termed gestural style: the sweep and energy of gesticulation, the relative frequency of different types of gesture (which is probably related to cultural differences

in the nature of preferred conversational topics), and the use of gestural metaphors (Stephens and Tuite 1983).

It is our hope that researchers in kinesics, especially those from foreign countries, will assist in the tasks of expanding the data base to include a wider range of cultures, and developing methods for more closely studying the transfer of information through the gestural channel.

Notes

1. The following individuals have contributed significantly to the preparation of this paper with helpful comments, kinesic cues, and affect displays: Justine Cassell, Brecky Church, Nancy Dray, Dodona Kiziria, Karl-Erik McCullough, David McNeill, and Laura Pedelty. Much of the research reported on here was supported by NSF research grant BNS 85-18324 to the University of Chicago.
2. The following abbreviations are employed in the gestural transcriptions: L(left), R(right); BH (both hands), LH (left hand), RH (right hand). / = pause; * = interrupted word (false start). Brackets [] enclose the speech cooccurring with the gesture. All of the examples used in this paper have been recorded, transcribed, and analyzed by researchers in the gesture and language workshop of the Committee on Cognition and Communication at the University of Chicago, directed by David McNeill.
3. These two types of gesticulation have very different developmental histories as well. According to McNeill (forthcoming: §5), children first point toward objects around the end of their first year. Beat-type gestures do not appear until age five or six.
4. The spread hand with wiggling fingers is probably a metaphor conveying smallness, slightness, corresponding to the word 'just' in the verbal channel [tenor = smallness of distance, vehicle = smallness of muscle movement].
5. An associated locus need not be an actual locus. One of our subjects gestured in the direction of the video recording equipment on a nearby table while mentioning the experimenters (all of whom were in another room). This person had drawn an association between us and the equipment most prominently employed in the experimental situation into which we had brought her, a nice example of deixis mediated through metonymy.
6. Somewhat different from this is a very common gestural metaphor, at least in our culture: the so-called conduit metaphor, in which information or knowledge is reified as a concrete object, passed from speaker to hearer. This gesture frequently accompanies the beginning of a new episode or the introduction of important new information into the discourse: the speaker will gesturally present the listener with an 'object' resting on the palm of his or her hand, or held with both hands (McNeill 1985: 357–358).
7. In the essay from which the above quote was cited, Ricoeur argues that the process of textual interpretation must involve 'understanding' as well as 'explanation', the latter being scientific analysis in the usual sense, the former — Dilthey's *Verstehen* — being grounded in 'our ability to transfer ourselves into another's psychic life on the basis of signs the other gives us' (1978: 150). In the case of gesticulation we gain entry into this private domain through a process of triangulation, working from the accompanying speech and our own experience as participants in conversational dyads, as gesturers and responders to gesture.
8. Particularly enlightening are rare instances of misinterpretation by the listener of the speaker's intentions in regard to a gesture. In one discourse videotaped in our laboratory,

two subjects are discussing a story from a comic book one of them has read. At one point in the narration, the speaker is describing the layout of an underground bunker to which the protagonist of the story has been taken. Just before beginning a sentence she points upward and rightward with her fully extended right arm. The interlocutor instantly turns her head, as though seeking the object in the room that the speaker has just pointed to. At the same time, the speaker begins her new sentence, in which flashing lights on the ceiling of the underground bunker are described. The interlocutor, evidently realizing that the pointing gesture, despite its wide sweep, was purely narrative-internal, and not foregrounded, quickly shifted her gaze back to the speaker.

9. McNeill states elsewhere that one of the properties of inner speech symbols is a 'sensitivity to linguistic potential and ability to select other words', comparable to that of 'verbs, prepositions and other linguistic symbols that convey an implicit syntax' (1987: 101).
10. Compare the remark by Vygotsky that 'thought is always something whole, something with significantly greater extent and volume than the individual word That which is contained simultaneously in thought unfolds in sequence in speech' (1984: 290, cited in McNeill 1987: 93).
11. One of our subjects, a native speaker of Hebrew, claimed that he hardly gestured at all. In the course of the hour or so of speech we recorded he kept his hands in his lap, folded together, effectively impeding gesticulation of the usual sort. Nonetheless, his hands were in constant motion: his thumbs moved up and down, and his wrists rotated, in synch with the intonational peaks of his speech.
12. Also to be noted is evidence that gesticulation plays a role in the regulation of conversational turn-taking. Duncan and Fiske's (1977) analysis of videotaped conversations indicates that listeners interpret the occurrence of gesticulation by the speaker as a signal to refrain from taking over the speech turn. The role played by gesticulation in the regulation of turn-taking does not, however, depend upon any semantic interpretation of the content of gestures, nor is gesticulation the sole or primary means of signalling the speaker's intent to continue his or her conversational turn.

References

- Arnheim, Rudolf (1969). *Visual Thinking*. Berkeley: University of California Press.
- Birdwhistell, Ray (1970). *Kinesics and Context*. Philadelphia: University of Pennsylvania Press.
- Bolinger, Dwight (1982). Intonation in nondeclaratives. In *Chicago Linguistics Society Parasession on Nondeclaratives*, 1–15. Chicago: Chicago Linguistics Society.
- Cosnier, Jacques (1982). Communications et langages gestuels. In *Les voies du langage*, J. Cosnier, A. Berrendoner, J. Coulon, and C. Orecchioni (eds.), 255–304. Paris: Dunod.
- Duncan, S. and Fiske, D. (1977). *Face-to-Face Interaction: Research Methods and Theory*. Hillsdale, NJ: Erlbaum.
- Efron, David (1941). *Gesture and Environment*. New York: King's Crown Press.
- Ekman, P. and Friesen, W. V. (1969). The repertoire of nonverbal behavior: Origins, usage and coding. *Semiotica* 1, 49–98.
- Goldin-Meadow, Susan (1982). The resilience of recursion. In *Language Acquisition: The State of the Art*, E. Wanner and L. Gleitman (eds.). New York: Academic Press.
- Goldin-Meadow, S. and Mylander, C. (1983). Gestural communication in deaf children: The non-effects of parental input. *Science* 197, 401–403.
- Hewes, Gordon (1973). Primate communication and the gestural origin of language. *Current Anthropology* 14, 5–24.

- Kendon, Adam (1972). Some relationships between body motion and speech. In *Studies in Dyadic Communication*, A. Siegman and B. Pope (eds.), 177–210. New York: Pergamon.
- (1975). Gesticulation, speech and the gestural theory of language origins. *Sign Language Studies* 9, 349–373.
- (1980). Gesticulation and speech: Two aspects of the process of utterance. In *The Relationship of Verbal and Nonverbal Communication*, M. R. Key (ed.), 207–227. The Hague: Mouton.
- (1983). Gesture and speech: How they interact. In *Nonverbal Interaction*, J. M. Wiemann and R. Harrison (eds.), 13–45. Beverly Hills, CA: Sage Publications.
- Lakoff, G. and Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.
- McNeill, David (1985). So you think gestures are nonverbal? *Psychological Review* 92 (3), 350–371.
- (1986). Speech-gesture synthesis, or natural acts. *Chicago Linguistics Society* 22 (1), 286–307.
- (1987). *Psycholinguistics: A New Approach*. New York: Harper and Row.
- (forthcoming). *Hand and Mind*. Chicago: University of Chicago Press.
- McNeill, D. and Levy, E. (1982). Conceptual representation in language activity and gesture. In *Speech, Place and Action*, R. Jarvella and W. Klein (eds.), 271–295. Chichester: Wiley.
- Merleau-Ponty, Maurice (1962). *The Phenomenology of Perception*, trans. by C. Smith. London: Routledge and Kegan Paul.
- Ricoeur, Paul (1978). Explanation and understanding. In *The Philosophy of Paul Ricoeur*, C. Reagan and D. Stewart (eds.). Boston: Beacon.
- Stephens, Debra (1983). Hemispheric language dominance and gesture hand preference. Unpublished Ph.D. dissertation, University of Chicago.
- Stephens, D. and Tuite, K. (1983). The hermeneutics of gesture. Paper presented at the American Anthropological Association annual meeting (Chicago).
- Vygotsky, Lev (1984). *Collected Works, Volume 2: Problems of General Psychology*, trans. by N. Minick. New York: Plenum Press.

Kevin Tuite (b. 1954) is affiliated with the Department of Anthropology at l'Université de Montréal, in Montreal, Quebec, Canada. His principal research interests include historical linguistics and language typology, the relation between gesture and speech, language acquisition, and South Caucasian (Georgian) folklore. Among his publications are 'Indirect transitives in Georgian' (1987), 'The geography of Georgian *q'e*' (1989), 'The acquisition of Georgian' (forthcoming, with N. Imedadze), and 'The Svans' (forthcoming).