

So You Think Gestures Are Nonverbal?

David McNeill
University of Chicago

In this article I argue that gestures and speech are parts of the same psychological structure and share a computational stage. The argument is based on the very close temporal, semantic, pragmatic, pathological, and developmental parallels between speech and referential and discourse-oriented gestures. Most of the article consists of a description of these parallels. A concept that unites outer speech and gesture is the hypothesis of inner speech.

Many cognitive psychologists hold that overt acts of linguistic production are the result of internal "computations." My aim in this article is to make the following point concerning gestures: Gestures share with speech a computational stage; they are, accordingly, parts of the same psychological structure. The metaphor of a shared computational stage captures the processing aspects of speech: that sentences and gestures develop internally together as psychological performances. The metaphor of a common psychological structure captures the idea that speech and gesture respond to the same forces at the same times.

Taking into account concurrent gestures suggests that in the dynamic situation underlying sentence generation two opposite kinds of thinking, imagistic and syntactic, are coordinated. The types of gestures that provide this insight are the *referential* and *discourse-oriented* gestures that spontaneously accom-

pany acts of speaking. In such gestures the hands function as symbols that are closely connected to the speech channel in terms of both time and semantic and pragmatic function. In the idiom of my title, such gestures are verbal. They are the overt products of the same internal processes that produce the other overt product, speech.

Because there are such close connections between gesture and overt speech, gestures offer themselves as a second channel of observation of the psychological activities that take place during speech production—the first channel being overt speech itself. The channels of gesture and speech are close, yet different. Combining a spoken sentence and its concurrent gesture into a single observation gives two simultaneous views of the same process, an effect comparable to triangulation in vision. In this article I consider some of the implications that can be drawn from the systematic comparison of speech and gesture.

The statement that gestures and speech are parts of the same psychological structure is contrary to the idea of body language, that is, a separate system of body movement and postural signals that is thought to obey its own laws and convey its own typically affective and unconscious meanings. It is also contrary to the assumptions of many linguistic analyses that hold that language structures should be analyzed only in terms of speech sounds plus grammar. We tend to consider *linguistic* what we can write down, and *nonlinguistic*, everything else; but this division is a cultural artifact, an arbitrary limitation derived from a particular historical evolution. Both body language and customary linguistic analyses follow a narrow approach, in which connected

This article was written while I was a fellow at the Netherlands Institute for Advanced Study, Wassenaar, The Netherlands. Financial support for my research was provided by the National Science Foundation (linguistics section), the Spencer Foundation, and the University of Chicago. The gesture research was carried out with the help of the sharp eyes of Catherine Greeno, Mitsuko Iriye, Elena Levy, Laura Pedelty, and Debra Stephens.

I have greatly benefited from comments at various times and in various forms from friends and colleagues, among them E. Bates, J. Bybee, J. Cosnier, G. B. Flores d'Arcais, S. Frey, M. Johnson, D. Kelter, A. Kendon, J. Lawler, W. J. M. Levelt, L. Mackenzie, N. B. McNeill, N. Quinn, M. Silverstein, and W. Washabaugh. The title owes something to the late Glenn Gould.

Requests for reprints should be sent to David McNeill, Committee on Cognition and Communication, Department of Behavioral Sciences, University of Chicago, 5848 South University Avenue, Chicago, Illinois 60637.

parts of a single psychological structure are studied separately. Both begin by radically separating gestures from speech, and from this starting point they go in opposite directions. That gesture and speech are parts of a single psychological structure has been a view held, however, by a number of psychologists and psychiatrists (Argyle, 1975; Barroso, Freedman, Grand, & van Meel (1978); Condon & Ogston, 1971; Cosnier, 1982; Ekman & Friesen, 1969; Kendon, 1972, 1980, 1983—to mention a few), but not by linguists (for a brief but quite definite opinion, see Chomsky, in Rieber, 1983). In this article I present new evidence that supports the broader view, that the whole of gesture and speech can be encompassed in a unified conception, with gesture as part of the psychology of speaking, along with, and not fundamentally different from, speech itself.

Acts of speaking are often accompanied in our culture by movements of the arms and hands that are termed *gestures* (Kendon, 1972, uses the term *gesticulation*). These are movements that (with a class of exceptions to be described) occur only during speech, are synchronized with linguistic units, are parallel in semantic and pragmatic function to the synchronized linguistic units, perform text functions like speech, dissolve like speech in aphasia, and develop together with speech in children. Because of these similarities, a strong case can be made for regarding gestures and speech as parts of a common psychological structure.

The similarities hold for the kinds of spontaneous and semiconscious gesture that can be seen accompanying much conversational and narrative speech. Such gestures play pictorial and discourse-related roles (Efron, 1941) and have been variously called “pictograms,” “ideograms,” “illustrators,” and “batons” (Ekman & Friesen, 1969).

There is another better known, more stereotypic type of gesture that should not be confused with these. The French and Italians are acknowledged as master gesticulators, but the Italianate gesture (a type called the *emblem* by Ekman & Friesen, 1969, and documented throughout Europe by Morris, Collett, Marsh, & O'Shaughnessy, 1979—see Kendon's, 1981, review of the latter) is not the referential and discourse-oriented gesture that

is described here. Emblems are gestures that have a specific social code of their own, have conventional paraphrases or names (Kendon, 1983), are learned as separate symbols, and can be used if they were spoken words; in many uses they are, in fact, unspoken words. An example is the *okay* sign or the *psychologically peculiar* sign (the first finger pointing at the temple while the hand moves in a small circle). These kinds of gesture are interpretable in the absence of speech, and this is one of their chief functions. The types of gesture described in this article, in contrast, are *not* interpretable in the absence of speech, are individual and spontaneous, are outside of any special social code that regulates them, and show interesting cross-cultural similarities (again, in contrast to emblems, which, as Morris et al. show, are highly specific to particular linguistic and national groups and even subgroups).

In this article I first explain and present arguments for the statement that gestures are manual symbols. Second, I exemplify the similarities that point to gesture and speech sharing a computational stage. Third, I consider and reply to counterarguments; and last, I present conclusions regarding the psychological process of speech production based on considerations of gesture and speech data.

Gestures Are Manual Symbols

To demonstrate the symbolic character of gestures, I will present examples of gestures produced by 5 adult female subjects who were narrating the same event from a cartoon story (they had been shown the cartoon just previously and were now telling the story to a listener—each subject separately). The point I wish to make is that, like conventional linguistic symbols, there is great commonality among subjects. Despite individual variation in the detailed manner of performance, every gesture includes upward movement. This commonality is produced, not by a shared code or language that exists specifically for gestures, but by each speaker separately creating her own manual symbol of upward movement. The concurrent sentences also conveyed the idea of upward movement, and this fact shows that semantic parallels existed between gestures and concurrent speech. As

they performed their gestures the subjects were saying¹

he tries going up inside the drainpipe

[hand rises and
points upward]

he tries climbing up the drainspout of the building

[hand rises
and starts to point upward]

and he goes up through the pipe this time

[hand rises quickly
and fingers open to
form a basket]

this time he tries to go up inside the raingutter

[hand steady and pointing] [hand rises quickly (still pointing)]

he tries climbing up the rain barrel

[hand flexes back and up]

Although a quantitative comparison is not realistic, there seem to be at least as much variation among the sentences of these examples as among the gestures; for example, among the words "pipe," "spout," "gutter," and "barrel," the variation is not less than in the manner of moving the hands upward, pointing or not pointing, in a basket shape or with a backward flexion.

Saussure (1916/1959) defined a symbol as a semiotic entity in which there are two distinguishable but inseparable sides—inseparable in the sense that considering one side in isolation from the other destroys the symbol (Saussure used the term *sign*, but in this article I use *symbol*, with the same meaning). The two sides of the symbol, Saussure called the *signifier* and the *signified*. The signifier is the *form* of the symbol; in the case of a word, for example, the form is the sequence of sounds or letters, /k æ t/ or c-a-t. The other side is the concept that the symbol signifies, the concept of a cat or cathood. Clearly, much can be said about verbal concepts. In this discussion the term *concept* is used to refer to subjects' mental representations, in-

cluding their intended meanings as well as their memories. Whether these concepts are categories, propositions, prototypical examples, contextualized individual ideas, or something else, is left unresolved, and undiscussed.

The gestures above are symbols in Saussure's (1916/1959) sense. Even though they are not conventional or arbitrary, these gestures are analyzable as paired signifiers and signifieds. The physical upward movement of the gesture itself is the signifier, and the associated memory of the character's upward movement, the signified concept. Considering just the gesture's upward movement in isolation from the speaker's memory of the character's upward movement, would destroy the symbol. Taking into account both sides of the gesture-symbol allows us to see that the gestural and linguistic channels exhibited the same meanings.

Further evidence that gestures are symbols in their own right comes from a phenomenon that can be seen when the same hand configuration and/or movement occurs twice in immediate succession. It can happen under this circumstance that there are two gestures that utilize the same configuration and motion of the hand, but convey different meanings. In such cases there is added a small extra and otherwise superfluous movement that separates the two meanings and indicates the moment that the meaning changes. The following example illustrates this movement:

and as he's coming up and the bowling ball's coming down

[left hand descends (= the bowling ball)
while right hand rises (= he)]

[both hands loop out and back (= the extra movement)]

¹ The reader is urged to act out the gesture examples. The experience of motor-speech synchrony is part of what the example is meant to demonstrate. Although two streams of behavior are produced, the experience should be of a single coordinated action. The left-most bracket that sets off the gesture description indicates approximately when to start the gesture. Thus, in the first example, the hand should begin to rise as you are saying the word "inside." Both the spoken and gestured parts of the example should be performed in a continuous and natural manner, without pauses.

he swallows it

[left hand descends again (= still the bowling ball)
while right hand rises and engulfs left hand (= now a mouth)]

Even though the posture and motion of the hands is the same in the second gesture, the meaning of the right hand has changed: At first it stood for the whole character, but then became just a part of this character (his mouth). Where the meaning shifted, the hands briefly withdrew and reentered the gesture space. The extra movement implies that the hands and their configuration had symbolic value for the speaker, and this value changed at the moment of the extra movement.

Another indication that gestures are true symbols occurs when the speaker divides a cognitive representation into two parts. One part is carried in the familiar way via the sentence, while the other is conveyed through the gesture. Speech and gesture cooperate to present a single cognitive representation. The gestures in these examples are not repairs. The sentences are well formed and adequately convey meanings. What seems to happen is that there is no grammatically convenient way to combine some aspects of the common cognitive representation with other aspects. In the example below ("she chases him out again") the gesture conveys the idea of the instrument of the act, whereas the act itself is described in the concurrent sentence. The speaker could have conveyed the instrument with a more complex sentence, but by producing a gesture the speaker avoided this extra complexity. The gesture in such a case is necessarily regarded as a symbol on an equal footing with the sentence.

she chases him out again

[hand, gripping an
object, swings
from left to right]

The sentence conveys the concepts of pursuit ("chases") and recurrence ("again"), but not the means of pursuit. The gesture shows this method—swinging an umbrella. The sentence is well formed and the gesture is not a repair or transformation of the sentence. To get the

full cognitive representation that the speaker had in mind, both the sentence and the gesture must be taken in account. (Examples of cooperation between gesture and speech are not difficult to find; several are cited in Kendon, 1984.)

It is important to note in this example the existence of sentence and gesture synchrony. Evidently the basis for synchronization is not that the gesture and sentence are translations of one another, but that they arise from a common cognitive representation, which neither exhibits completely.

For several reasons, therefore, it is apparent that gestures of the kind demonstrated are *referential symbols*. I have shown that (a) gestures are similar between individual speakers when the meanings conveyed are similar; the similarities are no less than among the concurrent sentences. (b) When the hands change meaning, small extra movements mark the boundary between one meaning and the next. (c) Complex meanings are divided between the speech and gesture channels; speech and gesture cooperate in these examples to present a single complex meaning.

Evidence That Gestures and Speech Share a Computational Stage

The principal argument of this article is that gestures and speech share a computational stage (are parts of the same psychological structure); they are connected *internally*. This section presents evidence in support of this conclusion: (a) Gestures occur only during speech, (b) they have semantic and pragmatic functions that parallel those of speech, (c) they are synchronized with linguistic units in speech, (d) they dissolve together with speech in aphasia, and (e) they develop together with speech in children.

Gestures Occur Only During Speech

Gestures and speech overwhelmingly occur together. We can understand this on the ground that they are parallel products of a common computational stage. Gestures by listeners also occur, but they are extremely rare. In approximately 100 hours of recording,

during which time there were thousands of gestures by speakers, there was a gesture by a listener only once (Stephens, 1983). Speakers and listeners are experiencing the same linguistic forms and content, but the passive comprehender role does not evoke gestural activity to anything like the same degree as the active producer role. One can speculate about the cause of this huge difference; for example, gesture production is part of a signal system used to extend the speaker's claim over the speaking turn (Duncan & Fiske, 1977), or the motor arousal needed to produce speech generalizes and induces activity in other parts of the body. The point remains, nonetheless, that gesture production and speaking are virtually restricted to the same situations.

More narrowly, the majority of gestures also occur during the speaker's actual speech articulation. About 90% of all gestures before clauses in a sample of six narrations occurred during active speech output. Of the 10% of gestures that took place during silence, all were immediately followed by further speech. In their own way these gestures of silence also indicate that speech and gesture share a computational stage. The silence gestures did not exhibit the content of the word or clause that was uttered subsequently. The break in the internal computations that led to speech interruption changed the character of the gesture. The gestures at these moments were of two kinds. One was a simple and rapid hand movement of a type that usually accompanies words whose importance depends on multisentence text relations (called *beats*); the other a type of gesture that itself symbolizes concepts about language and communication (called *conduit* gestures). These two kinds of gestures are related to the breakdown of speech itself. They are gestures that symbolize the functions of silence in this situation. The beat gesture served as a sort of metalinguistic index of the point of the breakdown, and perhaps also as an attempt to get the speech process going again (adding a new dimension to the term *beat*). The conduit gesture, presenting a conceptualization of language as consisting of objects and containers, appeared where the language process was temporarily unable to provide the next such container, and perhaps was also part of the effort to call

this container forth. These gestures occurred during silence, but this kind of silence is part of speaking. Speaking was temporarily transferred to the person's gestures, which provided metalinguistic commentary on the process of speaking itself.

Gestures and Speech Have Parallel Semantic and Pragmatic Functions

The goal of this section is to demonstrate that gestures are symbols equivalent to various linguistic units in meaning and function. The interpretation of such parallels is obvious, given the argument that gestures and speech share a computational stage. During this stage semantic and pragmatic functions are decided on, and both speech and gesture perform these functions in parallel. I start with referential gestures (iconic and various kinds of metaphorical), and end with gestures that have purely off-propositional, including discourse-oriented, functions (beats).

Iconic gestures. Several examples of semantic parallels between gestures and linguistic units have already been described. McNeill and Levy (1982) called gestures of this type *iconic*. An iconic gesture is one that in form and manner of execution exhibits a meaning relevant to the simultaneously expressed linguistic meaning. Iconic gestures have a formal relation to the semantic content of the linguistic unit. The signifier part of the symbol is formed so as to present an image of the signified part. The five versions of the going-up-the-pipe gesture were all iconic. They were iconic because, for instance, when the person's hand rose upward while she was saying "he crawls up the pipe," the gesture exhibited, via its own upward motion, the signified concept of upward movement.

McNeill and Levy (1982) performed an analysis of gesture form in terms of 44 different *movement features*. These specify details of the gesture movement, such as whether the gesture was made with one hand or two hands; whether the palm was open or closed; whether the fingers were extended, spread, or curled; whether movement was upward, left, or right; and so forth. McNeill and Levy also analyzed the linguistically conveyed meanings in terms of 38 *meaning features*, specifically,

the meanings of the verbs in the clauses that the gestures accompanied. These features specify details of meaning, for example, whether the verb described horizontal motion, closure, or reaching an end state (see Miller & Johnson-Laird, 1976). For example, "catch up to" as used in cartoon narrations, has the meaning features of *pursuit* and *horizontal motion* (at least); it also has the meaning feature of *end state* (because you don't catch up to something unless you achieve the end state of reaching or becoming adjacent to it; see Vendler, 1967). In contrast, "chase" in cartoon narrations has the meaning features of *horizontal motion* and *pursuit*, but not of *end state*.

McNeill and Levy (1982) prepared a 38×44 matrix, in which, along the rows, one saw the frequency (in the gestures produced by six narrators) of each movement feature accompanying the meaning feature at that row. Down the columns one saw the frequency with which each meaning feature co-occurred with the movement feature at that column. Looking at this matrix (only part of which is reproduced in McNeill & Levy) it is obvious that the patterns of peaks and valleys in each row and column are different. Every meaning feature has its own distinctive gesture profile and, similarly, every gesture feature has its own distinctive meaning profile. For example, when a verb lacked the meaning feature of *end state* (as in "chase") the gestures tended to be performed with one hand, whereas when a verb possessed this meaning feature (as in "catch up to") the gestures tended to be two handed. In other words, the *end state* meaning feature, in which there are logically two component ideas—the idea of an end state itself and that of something reaching the end state—was exhibited in gestures that also included two parts. This example illustrates the essence of gesture symbols that refer iconically: There is predictability of signifier form from the signified concept (in part).

Stephens and Tuite (1983) proposed a subdivision of iconix according to the degree of alienation of the hands from their primary function of object manipulation. Under this proposal there would be a progression of gesture types running from iconix₁ to iconix₂ to beats (defined later). In iconix₁, the hands

recreate and manipulate a virtual object, for example,

he takes a crate

[both hands held apart,
as if holding onto an object]

The gesture recreates and manipulates a *discourse crate*, and the hands play the part of hands. In iconix₂, the hands function more abstractly to depict an entity, also mentioned in speech, other than hands, for example,

go into Sylvester

[left hand encircles right-
hand space, then right hand
slides down left]

The gesture shows an object being swallowed but does not depict hands. One might expect that with children, iconix₁ would develop earlier than iconix₂, and I show in the section on gesture development that this proves to be the case.

Not only are iconic gestures capable of referring, but they reveal discourse functions. In the following example from a male speaker the gesture reveals the speaker's point of view toward the event he is narrating:

and he bends it way back

[hand rises up, appears
to grasp something, and
pull it back]

This gesture iconically exhibits the event being described in the clause, "he bends it way back." The object that was being bent back was a tree (as is known from the earlier narrative), and the gesture shows the hand of the character grasping this tree and bending it back.

The gesture also reveals that the speaker was "seeing" this event from the viewpoint of the character who was performing the act. The speaker's hand symbolized the character's hand. If the viewpoint had been that of the tree (the other participant), we would have expected a gesture in which the arm moves backward, but the hand—now representing a tree rather than a hand—is not grasping anything.

The viewpoint of the character performing the action also appears in the concurrent sentence. The agent of the act of bending

back is functioning as the discourse reference point, or thematic topic, of the sentence (Clancy, 1980; also see Kuno, 1976). It was this agent's hand that the gesture depicted. If the viewpoint had been that of the tree instead, one would have expected a passive or other sequence form capable of indicating that the tree is the discourse reference point (e.g., "it got bent back").

Gestures that function specifically to emphasize discourse relationships are described under the heading *Beats*.

Metaphoric gestures. Another type of referential gesture, closely related to iconix, is termed *metaphoric* (McNeill & Levy, 1982). Gestures of this type demonstrate that references with gestures are not limited to concrete objects and events. Metaphorix are semantically parallel to sentences with abstract meanings. The shared computational stage between speech and gesture, therefore, is one that generally can be part of acts of linguistic production regardless of the level of conceptual abstraction.

Metaphoric gestures are like iconic gestures in that they exhibit a meaning relevant to the concurrent linguistic meaning. However, the relation to the linguistic meaning is indirect. Metaphoric gestures exhibit images of abstract concepts. In form and manner of execution, metaphoric gestures depict the vehicles of metaphors (Richards, 1936). The metaphors are independently motivated on the basis of cultural and linguistic knowledge. (Thus, the metaphoric gestures that accompany speech are a source of information about cultural beliefs and attitudes—in general, the world view of the speaker.) For example, a common metaphor in our culture (but not in others) is a bounded physical object for the concept of a mental or cultural product (e.g., human knowledge, works of art, and memory; thus we speak of an artistic *object*), and there are gestures that exhibit these metaphors. The following is an example from a narration:

it was one of the Tweetie Pie and Cat cartoons

[both hands are lifted up
as if supporting an object]

The speaker appeared to be holding onto an object while he referred to a work of art. This gesture-created *object* was the meta-

phoric vehicle for the abstract concept of the cartoon.

In conversations between previously unacquainted adult subjects recorded for a research project on face-to-face interaction (Duncan & Fiske, 1977), metaphoric gestures amounted to as much as half of all the gestures the participants produced. Often there were no iconic gestures at all. In these conversations, that is, *all* of the gestures that are referential could be metaphoric.

The next three sections describe some of the varieties of metaphoric gesture that can be observed.

Mathematics gestures. Metaphoric gestures appear frequently in technical discussions where abstract content plays an important role. The following examples are taken from a videotaped discussion between two mathematicians.

Verbal references to the concept of a mathematical dual were accompanied by gestures where the hands alternated between two positions. In a mathematical dual, one relation is replaced by its converse relation. For example, the dual of "A is above B" is "A is below B." The hand rotation thus exhibits a concrete image in which one position is replaced by another: There are two hand positions and two relations, and there is replacement:

and this gives a complete duality

[hand loops up and
backwards through
a circle, ending
at a higher level]

The motion of the hand was synchronized with the verbal reference to the dual concept, which itself of course also is a metaphor.

In other examples, gestures exhibited concepts of limits. For instance, with the concept of a direct limit there were gestures in which the hand moved along a straight line followed by end marking (a tensed stop). An example is the following:

this gives a direct limit

[left index finger slides
along right index finger
and comes to tensed stop
just beyond the tip]

End marking and the barrier it implies provides a metaphor for the concept of a limit, and the straight line along which the hand moved, a metaphor for directness. This interpretation is confirmed by the gestures for inverse limits. Like direct limit gestures, inverse limit gestures also had end marking, but the movement, rather than along a straight line, followed a downward loop. That is, the gesture showed the interchange of two points, an image of the concept of an inverse (in which the start point and end point of a relation are exchanged). The following is an example:

it's an inverse limit

[hand moves down then
upward in continuous
movement, and stops
at original level]

I infer from metaphoric gestures that both an abstract concept and a concrete image (regarded by the speaker as similar to the abstract concept) are part of the same computation underlying the production of the sentence. Together with the abstract dual concept, in which one relation replaces another, there was also an image of a spatial position being replaced by another. Together with the abstract limit concept there was also an image of a moving object coming up to a barrier. Whatever the internal computations are that underlie the production of sentences with abstract meanings, the evidence of metaphoric gestures suggests that they can be highly similar to the computations that underlie sentences whose meanings are concrete (Gildea & Glucksberg, 1983).

Output errors sometimes appear in one channel but not in the other. This can occur if the error is due to a disturbance that takes place after the shared computational stage. In the mathematics discussion there was a speech error of the semantic type (Garrett, 1976), in which the word *inverse* was used where *direct* was appropriate. The synchronized gesture was the gesture for a direct limit. This suggests that conceptual confusion was lacking, and the error occurred after the shared computational stage of gesture and speech. The following is the example:

an inverse limit . . . of . . .

[hand moves upward
and forward in
straight line]

(listener says: it's a direct limit)

I mean a direct limit

[hand makes
same gesture
as before]

If gestures pass through a less complex transformation after the common computational stage, it would not be by chance that the error here occurred in the speech channel and not in the gesture channel. With aphasics also, referential gestures exhibit contextually appropriate meanings when there are paraphasias (see the *Aphasia* section).

Conduit metaphors. This type of gesture has been in use for at least 2,300 years. Montaigne (1958), explaining Zeno's attitudes toward the divisions of the soul, described a conduit metaphoric gesture—in fact, a graded series of them—in which Zeno conveyed the concept of degrees of knowledge with degrees of closure of the hand:²

Zeno pictured in a gesture his conception of this division of the faculties of the soul: the hand spread and open was appearance; the hand half shut and the fingers a little hooked, consent; the closed fist, comprehension; when with his left hand he closed his fist still tighter, knowledge. (p. 372)

Conduit gestures (including Zeno's) are iconic depictions of abstract concepts of meaning and language. Such metaphors are independently motivated by linguistic evidence (Lakoff & Johnson, 1980; Reddy, 1979). A conduit gesture exhibits a container or substance (e.g., forming a cup shape), whereas simultaneously the linguistic channel refers to an abstract concept (an appearance, content, comprehension, knowledge, language, meaning, art, the soul, etc.). The speech and context of speaking need not contain anything about containers or substances. To describe accurately a speaker's cognitive representation of an abstract concept, it is necessary to

² I am grateful to J. Stern of the University of Chicago, Department of Philosophy, for bringing this wonderful passage to my attention.

include the concrete images implied by conduit gestures. An example is the gesture produced by the speaker cited earlier who appeared to be holding onto an object while saying "it was a Tweetie Pie and Cat cartoon." The person's cognitive representation of the work of art evidently included a container with the cartoon inside, or the cartoon itself as a physical substance, in either case invoking the conduit metaphor. The computational stage shared by speech and gesture contained—together with the abstract concept of a work of art—the image of a concrete object, metaphorically interpreted as this abstract concept.

In another example, a conduit metaphoric gesture is part of a virtuosic image of the concept of the past tense of the subjunctive mood:

even though one might

[both hands form cups]
[and spread wide apart]

have supposed (no pause)

[cups vanish]
[abruptly]

The cup shapes could be images of the concept of meaning; the spreading apart, the idea of possibility (cf. an anything-is-possible gesture), whereas the abrupt closure of the cups of meaning, that what might have been, in fact, did not obtain. Such is the essence of the past tense of the subjunctive mood (Curme, 1931).

The hand shape seen in the above gesture is also typical of conduit gestures for *Wh* questions—gestures that accompany, for example, "what did you mean/say/think?" The palm is forward and held upright, while the fingers extend and spread out, usually also curling: a metaphor for the question as an object or container into which the listener is supposed to place an answer. One of the reviewers for this Journal suggested that there are other interpretations of the *Wh*-question gesture that do not rest on the conduit metaphor. For instance, the opened hand could be interpreted as the questioner's readiness to receive the answer. There are several replies

to this. First, this interpretation still takes the gesture to be a metaphor for a concept that is abstract. Second, this particular interpretation is also an instance of the conduit metaphor. If the hand is ready to receive an answer, then the answer is being regarded as a substance.

Metaphoric gestures in other cultures. Conduit gestures enable the speaker to act as if he could hold onto and manipulate abstract concepts such as "knowledge," "art," "language," and "meaning." The manipulation of abstract concepts is not present in the gestures of some other cultures, even though there are gestures in these cultures that present abstract concepts in the form of concrete objects. Specifically, conduit gestures are absent from a series of filmed conversations in Turkana, a language spoken by nomadic herdsman of the Turkana district of Kenya (see Dimmendaal, 1983, for a description of this language). I examined more than 100 gestures recorded in an anthropological film about the Turkana (MacDougall & MacDougall, 1977) and found no clear instances of gestures accompanying verbal references to abstract concepts in which the hand appeared to more than briefly hold or manipulate a container or substance. The Turkana speakers produced numerous gestures in which abstract concepts were metaphorized as concrete entities or places. However, they did not seem to manipulate these entities in the way that North American speakers manipulate objects evoked by the conduit gesture. Without far more knowledge of Turkana culture one cannot begin to interpret the gestures of these speakers. One would at least have to know independently something about their underlying cultural models. It is an observable difference from the behavior of North American speakers, nonetheless, that metaphoric gestures in which abstract concepts are "handled" are essentially lacking. The following is an example of what a Turkana speaker produces, in a context where from a North American speaker there would very probably be a conduit gesture (cited by Stephens & Tuite, 1983):³

³ Translations and transcriptions were kindly provided by R. Dyson-Hudson of the Cornell University, Department of Anthropology.

these Europeans want to extract all our knowledge—pft!	
toditarite ngitunga	lu na kilna yoka —nith!
(they-extract people)	(this-here knowledge
	our-inclusive)

[hand moves to brow
and appears to
remove something]

[fingers create
small object
rising upward]

[object
seems to
fly away
on its
own]

In this gesture an object was presented as though it moves on its own, not requiring manipulation. "Knowledge" was pulled from the speaker's brow (this much manipulation occurring), but then rose up and flew off on its own. Although there might be a universal tendency of human thinking to embody abstract concepts in images of concrete entities, the metaphor of manipulation is by no means universal.

Beats. In the discussion of speech-gesture similarities so far (with the exception of the "point-of-view" discussion), I have been describing parallels of propositional content. There are also gestures that demonstrate parallels of pragmatic function. In the common computational stage of gestures and speech, decisions are made at several levels. Beats are gestures that emphasize off-propositional relations. These gestures add another dimension to the case I am making for gestures being part of the same psychological structures as speech.

In contrast to referential gestures, beats (as they were called by McNeill & Levy, 1982; batons by Ekman & Friesen, 1969) 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 linguistic items. Background sentences in narrations are important because of their relations to other parts of the narration (the main story line), not for their own propositional content as such. Similarly, lexical items that contrast with other lexical items (e.g., "it's Wednesday—no, I mean Thursday") are important because they contrast with the other items. (Beats during silences implicitly contrast the absence of a

word with its desired presence.) These situations typify the circumstances in which beats are likely to occur. The beat signals that what is important is an interitem relation. Such relations are, collectively, at off-propositional levels (both supra- and subpropositional), in contrast to the propositional level of referential (iconic and metaphoric) gestures.

Lacking form, beats say that the material they accompany is *not* part of the story line. Beats should occur at points of significant discontinuity in the discourse structure. McNeill and Levy (1982) found that sentences classified as extranarrative were accompanied only by beats. Iconic gestures accompanied only sentences classified as narrative. Beats also accompanied narrative sentences when the gestures performed the off-propositional function of emphasizing individual lexical or sublexical items (syllables). Without exception, beats did not emphasize the propositional content of sentences. An objective method for identifying points at which discourse discontinuities occur has been developed by Levy (1984).

The definition of an iconic gesture (as stated previously) is that it exhibits in its form and manner of production an aspect of meaning relevant to the meaning of the concurrent linguistic item. This definition is semiotic in character, and refers to the meaning-function of the gesture type. A beat is correspondingly defined as a gesture that does not exhibit meaning iconically. There is, in addition to the semiotic definition, a formal difference between iconix and beats. Iconic gestures are typically large complex movements that are performed relatively slowly and carefully in the central gesture space. Beats are typically small simple movements that are performed more rapidly at or near the rest position of the hands.

These form differences themselves can be regarded as an iconic depiction of the distinction between the narrative and extranarrative functions. The display of propositional content in an iconic gesture takes place on the central stage, whereas the extranarrative commentary on this performance by a beat emanates from the wings.

Moreover, referential gestures are preferentially made with the dominant hand (Stephens, 1983)—right or left, depending on the handedness of the speaker—but beats are made bimanually or with the subordinate hand. Kimura (1973a, 1973b) observed that the right hands (of dextral subjects) were engaged in gesture production during speech. In Stephen's study this was true of iconic and metaphoric gestures, but not of beats.

The connection of referential gesture production to the hand used for object manipulation is presumably not accidental because these gestures often evoke images of objects and their manipulation. Beats do not evoke this sort of imagery, and also are not preferentially performed with the hand used for manipulating objects. Thus, beats occupy the end of the progression of gesture types that are increasingly alienated from the primary function of the hand for object manipulation (iconix₁, iconix₂, beats).

It is also possible that the choice of the hand in gesture production is influenced by the motor complexity of the gesture; the dominant hand is preferred for more complex performances (Kimura & Archibald, 1974). Motor complexity appears to be a minor factor, however. Stephens (1983) observed that the dominant hand was still preferred for conduit metaphoric gestures in which (as in beats) there are not more than two movement components. That is, the dominant hand was preferred for simple movements when the gesture was referential.

The following example is a typical beat (right-handed speaker) performing a discourse-oriented function:

he keeps trying to catch the bird

[left hand at rest position
rotates quickly outward
and back]

The statement comments on the structure of the story as a whole, namely, that it consisted

of a string of events like the one described. The propositional content of the sentence was that the character tries to catch the bird. However, the emphasized aspect of the sentence was that it was providing background commentary. In keeping with this off-propositional role, the form of the gesture was unrelated to catching the bird or anything else. The linguistic unit with which the gesture was synchronized, "keeps" (on), coded the same discourse function. That the gesture and this verb were synchronized further supports the proposal that the gesture and word shared a computational stage. The following illustrates a beat that accompanied a linguistic repair—highlighting the successful completion of a subpropositional element:

Alice li- Alice lives with her father who's a new uh runs a . . . sort of a newsstand

[fingers rise
and then
come down]

The beat synchronized with "newsstand", indicating its contrast as a repair and the (at last) appropriate lexical choice.

Summary. Gestures are semantically and pragmatically parallel to their concurrent linguistic units in speech output. Metaphoric gestures show that the parallels are not limited to concrete references. Iconic and metaphoric gestures exhibit the propositional content of linguistic units. The form of the gesture is determined by the form of the content being presented. Beats emphasize elements that code off-propositional functions and have an abstract form that is unrelated to propositional content. Together iconix-metaphorix and beats provide gesture coverage that parallels linguistic functions ranging from the referential uses of language to the discourse-indexing uses. These parallels imply a shared computational stage at which the semantic and pragmatic values of sentences and gestures are jointly worked out.

Gestures Synchronize With Parallel Linguistic Units

Speakers tend to perform gestures at the same time they produce semantically and pragmatically parallel linguistic items. Iconic and metaphoric gestures almost never cross

clause boundaries. Such synchronization suggests that the gesture reveals the moment at which the speaker's thinking process formulates the concept that the linguistic item signifies. A particularly interesting form of synchronization is shown in the following example, in which there is artificial prolongation of a gesture:

and she dashes out of the house	
dynamic stroke: hand shows "dashing out"	static stroke: hand keeps position
(no pauses)	
[retraction]	

The gesture illustrated something "dashing out" (this was the dynamic stroke). The resulting posture—the hand extended outward and the fingers splayed (definitely a marked posture)—was held in midair until the clause describing the same event came to an end.

In this example the speech channel controlled the timing of the gesture channel. A recent series of experiments by Levelt, Richardson, and La Heij (in press) found evidence of the reverse direction of influence, deictic (pointing) gestures that control the timing of speech. The speaker evidently can compare the gesture and speech channels and confine symbols felt to be equivalent to the same span of time. The source of influence may shift between the speech and gesture channels depending on the complexity of the speech transformation after the common computational stage (the speech transformation in the experiment by Levelt et al. was relatively uncomplex).

That gestures reveal the moment at which the speaker formulates a concept is suggested in another way, in which—seemingly a paradox—there is not synchronization. There exist anticipations where the concept revealed in the gesture becomes available before the sentence can grammatically make use of the linguistic item that signifies the concept. If we are to link the gesture and linguistic item, we must interpret the moment of the gesture to be the moment of the shared computational stage, but the speech product of this stage was stored and later retrieved from memory.

In the following, a gesture for the location of a character anticipated the verbal reference to the location ("there"):

they keep on flashing back to Alice just

[hand moves out and
points to location]

sitting there

The speaker evidently formulated a mental representation of the location of Alice no later than the moment of saying "flashing back." This incident of thinking, however, anticipated by four words the verbal reference to the location.

In general, two kinds of gestural anticipation must be distinguished. On the one hand there are gestures during silence in which no computation of semantic or pragmatic function takes place. In this case, speech comes to a halt and the gesture is either a beat or a conduit metaphoric, which comments on the speech breakdown. On the other hand there is gestural anticipation that occurs during uninterrupted speech output. In this case there is computation of a semantic and pragmatic function, and the gesture exhibits the product of this computation stage, but for grammatical reasons the unpacking of the cognitive representation in the speech channel is delayed. The linguistic version of the semantic-pragmatic representation necessarily occurs at the end of a grammatical constituent, in this case, enforcing delay.

To summarize, gestures synchronize with linguistic units that have the same semantic and pragmatic functions. Synchronization shows that the linguistic unit and gesture belong to the same psychological moment and could arise from a common psychological structure activated on-line during speech.

Gestures Are Affected Like Speech in Aphasia

Gestures and speech are affected in parallel ways by the neurological damage that produces Broca's and Wernicke's types of aphasia. Broca's aphasia consists of a relatively intact ability to use referring terms, together with a radical disturbance of the ability to combine these referring terms into larger grammatical wholes. The speech of these patients is often characterized as telegraphic.

Broca's aphasics produce numerous and sometimes quite elaborate iconic gestures, but few or no beats. This is our interpretation of the findings of Cicone, Wapner, Foldi, Zurif, and Gardner (1979). (A review of various studies that touch on the gesture production of aphasics can be found in Bates, Bretherton, Shore, & McNew, 1983.) Pedelty (1985), who has collected observations of the gesture production of Broca- and Wernicke-type aphasic patients, confirmed that the Broca type produces iconix. Both iconix₁ and iconix₂ occur. Beats also occur, but only during silences when the speaker is attempting to retrieve words (a circumstance that in normal speakers is associated with subpropositional emphasis). There are no beats that perform extranarrative functions. Thus, in parallel with the ability to use referring linguistic terms, Broca-type patients retain the ability to create referential gestures, but have lost the ability to mark interrelations of items in parallel with the dissolution of the ability to combine linguistic symbols. Beats that contrast the absence of a referring term to its desired presence also occur, but these are the only beat gestures. Because as movements the two kinds of beat are identical and quite simple, their absence for performing one function (interrelating parts of text) and presence for performing the other (word retrieval) is quite striking.

In other words, the pattern of preservations and losses that follow brain injury in these patients is quite similar in the speech and gesture domains. For Broca aphasics, speech and gesture appear to share a computational stage. There is lacking from this stage, however, the ability to interrelate symbols, and this affects speech and gesture in parallel.

A similar but complementary picture appears with Wernicke-type aphasics. In this syndrome there appears to be disruption of the ability to form coherent semantic plans, but preservation of the ability to construct sequences of words. The speech of these patients is often described as vacuous. In the gestures of Wernicke-type patients there are few or no interpretable iconix, but there are movements that seem to be beats. The theoretical possibility of beats is suggested by the Wernicke-type aphasic's ability to produce word sequences, and also by the high propor-

tion of pronoun compared to noun occurrences in their speech (Wepman, 1977). Delis, Foldi, Hamby, Gardner, and Zurif (1979) described for Wernicke-type aphasics hand movements at the start of new subordinate clauses when the clauses were semantically discontinuous. These movements may have been beats. Pedelty (1985) has observed with Wernicke-type aphasics gestures that in terms of form are classifiable as beats. There are also conduit metaphors when the speaker refers to his or her own speech problems. These patients, in other words, produce gestures that do not relate to propositional content, and apparently these are the only gestures they do produce. The dissolution of function in Wernicke-type patients therefore also shows parallels of gesture and speech ability.

If gestures pass through a less complex transformation after the shared computational stage than does speech, it should be possible for the speaker's choice of words to be inappropriate (a paraphasia), whereas a concurrent gesture exhibits the appropriate meaning. Pedelty (1985) has observed several such examples. The reverse combination, a correct word choice together with an inappropriate gesture, should be rare or nonexistent, and she has seen no examples of this type. The occurrence of appropriate gestures implies that the paraphasias are not due to conceptual confusions, but arise from disturbances after the speaker has developed an appropriate cognitive representation.

Thus, gestures and speech show an intimate connection to one another in the special situation where dissolution of linguistic ability has taken place. There is simultaneous, and functionally parallel, dissolution of gesture ability. This linkage implies again that gesture and speech are parts of the same psychological structure. The aphasic data add the further suggestion that speech and gesture are mediated by the same parts of the dominant cerebral hemisphere for speech/gesture production as would be implied by a common computational stage.

Gestures Develop Parallel to Speech in Children

Although interpretations vary, a broad characterization of the linguistic development

of children would show that it passes through three overlapping stages: (a) an initial emphasis on denoting concrete objects and situations. This is possible with very simple expressions, single words and simple combinations of words, provided there is a detailed context of speaking shared by the child and the others with whom the child interacts, as invariably there is. This stage typically occupies the first 2 years and part of the third. (b) Gains in flexibility, particularly in the construction of grammatically structured sentences for expressing relations (spatial, temporal, causal, interpersonal, etc.) between objects and objects and persons, make up the second stage. At this stage there is the possibility of linguistically sorting out the ways in which different semantic relations are coded by the language, but not yet of using devices such as anaphora for coding the relations of sentences to text. This stage occupies the 3rd, 4th, and 5th years, approximately. (c) The emergence of text coding constitutes the third stage. This stage occupies the remainder of primary language acquisition.

The point of the above summary (which,

words are used within the same scripts, that is, interactive routines between the baby and objects and people. So, for instance, most babies have words for greeting and farewell like "hi" and "bye," and at the same time have waving gestures that are used in carrying out the same script. Babies who have verbal telephone routines ("hello," "speaking," etc.) also have telephone gestures (placing the phone to the ear etc.; see Bates et al., Table 3.3 for numerous examples).

Pointing and reaching gestures are probably derived from movements in which real objects are manipulated (Carter, 1975). These and imitative performances seem to be the gestural counterparts of the initial stage of linguistic development, where the emphasis is on denoting concrete objects and situations.

Early in the second stage, by 2½ years, iconic gestures are clearly present (McNeill, in press). In the following example from a 2½-year-old, an iconic gesture accompanies a sentence (part of a cartoon narration) describing a scene in which a character was shown rolling transversely across the visual field with a bowling ball inside him:

and it went away with that . . . in th-e-r-e

[arm moves right
in large sweep-
ing movement
while hand
wobbles up and
down]

[turns completely
around in chair
and points to
room behind]

I hope, is broad and exceptionless enough to escape theoretical and factual disputes) is to compare language development, that is, overt speech development to gesture development. That these pass through the same stages is my final piece of evidence that speech and gesture can be usefully regarded as parts of a single psychological structure.

The earliest referential gestures described in the developmental literature are pointing (deictic) gestures and various imitative performances such as waving hello/goodbye (see Bates et al., 1983, for extensive review and discussion). Bates et al. present a number of convincing arguments that such gestures and the earliest spoken symbols develop together as a single system. In particular, gestures and

Notwithstanding the occurrence of iconic gestures at an early age, the production of these gestures follows a developmental path parallel to that of speech development. Werner and Kaplan (1963) described an increasing symbolicization of linguistic forms during language development that they term *distancing*. This refers to the relation between signifiers and signifieds. Young children act as if properties of the signifier must reproduce properties of the signified. There is accordingly little distance between them. In the speech realm, lack of distance can be shown with examples of onomatopoeism (Werner & Kaplan, 1963). For instance, a child learning French said "boom-er" (with the French verbal suffix -er) for the meaning of something

falling down; a child learning German said "whee-en" (with the German verbal suffix -en) for the meaning of going down a slide; a child learning English said "wau-wau dog" for dogs; and when children are asked to repeat a word such as *chicken* spoken by an adult, they might say, rather, "peep-peep" or some other such onomatopoeicism. A signifier has been formed that has a (phonetic) property that reproduces a property of the signified. The lack of distance between signifiers and signifieds shown in these examples also exists in children's early iconic gestures.

In the iconic gestures of adults the same posture and movement of the hands can be correlated with new references. As described earlier, when changes of meaning occur in otherwise identical gesture movements, the gestures are accompanied by small extra movements that mark the point at which the meaning changes. A closed hand can, accordingly, one moment play the part of one character and the next moment, the part of a different character. The gesture space can one moment be a windowsill, the next moment a street, and so forth. Within limits set by the need for symbols to remain iconic (limits that are admittedly hard to state), the hands and gesture space of adults are free to have the meanings the speaker designates them to have. The iconic gestures of young children are less flexible. The limits on accepting new meanings, however we state them, are less wide. The child's hands tend to play only the part of the character's *hands*. If they have to play the part of something else, the child includes additional movements that exhibit nonhand properties and differentiate the gesture from a hand, for example, the wobbly movements in the example that exhibited the rolling of the bowling ball and differentiated this gesture depicting an entire character from a gesture depicting a hand (adults perform the corresponding gesture with an unperturbed sweep). Children's gestures tend to be *more* detailed and *less* abstract, therefore, than the corresponding gestures of adults for the same references, and this effect follows from the lesser distance children tolerate between signifier and signified.

Children prefer to make gestures showing running, not with their hands, but with their

feet, and iconic gestures incorporating the head and legs are far more common with little children than with adults. Moreover, the gesture space of young children is (in many cases, absolutely) larger than that of adults and, more significantly, is centered on themselves. For adults the gesture space is a more or less flattened disc in front of the body onto which images (gestures) are projected and gestures rarely cross the frontal plane of the body; but for children it is a sphere that includes the child and his or her own movements as the center point. The above example again is an illustration. The child totally turned around in her chair to complete the gesture (the gesture space was wrapped around her).

Children's gestures, in short, appear to be *enactments*. It is for this reason that iconix₁ precedes iconix₂, developmentally. The entire body and all of its relevant parts reproduce the movements of the character whose acts are being described. Body parts of the character tend to be played by the corresponding body parts of the child (and if played by the hands, with movements added to simulate the real thing), the gestures are large, like real actions, and the gesture space, like the space of real actions, centers on the child. From this situation of minimal distance between the gesture and the concept or memory that it represents, the child develops increasingly symbolic gestures in which size, space, and meaning are (comparatively) freely assigned. The development of symbolic gestures out of enactments thus parallels the increasing distance in Werner and Kaplan's sense between signifiers and signifieds that is appearing at the same time in the linguistic domain. This parallel joins the others mentioned earlier to imply that gestures and speech are parts of a single psychological structure. Speech and gesture rise along the same developmental curve.

Also showing the phenomenon of enactment by children is the absence from their gestures of sharp boundaries corresponding to meaning changes. A gesture exhibiting one meaning gradually turns into a gesture exhibiting another meaning, just as one physical motion leads by degrees to another physical motion. Adult gestures, in contrast, have distinct physical boundaries corresponding to

the changes of meaning between the gestures (cf. the extra movements, described earlier, that adult speakers add to mark the division between otherwise identical gesture movements).

The phenomenon of enactment implies a quantitatively, if not qualitatively, different kind of thinking during the common computational stage on the part of little children, in which actions are mentally recreated. The roots of this style of thinking in the sensorimotor intelligence of early childhood can be easily imagined (Piaget, 1954; for theoretical linkage of sensorimotor intelligence to gesture production, see McNeill, 1979).

Enactment appears not only in gesture symbols, but also in the smaller signifier-signified distance of children's linguistic symbols. Onomatopoeicisms are *acoustic* enactments; "whee-en" acoustically enacts sliding down a slide, "wau-wau" enacts (to a child if not to a dog) what it is like to be a dog, "boom-er" enacts an object falling down, and so forth.

The third stage of language development, text coding, is accompanied by the first occurrences of beats. Although beats demand very little in the way of movement control, they are absent from the gesture performances of very young children. Thus children's developmental order is iconix_1 - iconix_2 -beats, corresponding to the progression of increasing alienation of the hands from their primary function of object manipulation (Stephens & Tuite, 1983).

The earliest beat we have observed was produced by a 5-year-old and emphasized a subpropositional contrast in word meaning:

execute is something else

[both hands repeatedly
move up and down on
armrest of chair]

This gesture (assuming it was one of the first) occurred long after the first iconic gestures. Clearly such simple movements are possible as motor performances in children much younger than 5 years. The complexity of the beat gesture is internal. That beats do not appear in speech development until there is the beginning of text coding is another ges-

ture-speech parallel, therefore, which supports the argument that gestures and speech share a computational stage. Not until this stage includes decisions about off-propositional indexing, do beats arise from it.

Metaphoric gestures also appear for the first time during the third ontogenetic stage. The first examples are conduits. These can be observed by 5 or 6 years. More original metaphors do not appear until much later (at the conclusion of the article I cite one of the earliest, from a 9-year-old). Children's ability to interpret metaphoric usages is likewise a late development (Winner, Rosenstiel, & Gardner, 1976). Thus, like other gesture types, the development of metaphoric gestures parallels that of the corresponding linguistic ability in children.

Recapitulation

To recapitulate: Speech and gesture appear together and perform a variety of semantic and pragmatic functions in parallel. Iconix emphasize the propositional content-bearing elements of speech, ranging from words to full clauses. Beats emphasize discourse and other interitem relations in which there is discontinuity or contrast. There are parallels therefore at the propositional and at several off-propositional levels. These parallels support the hypothesis that gestures and linguistic items are parts of the same psychological structure and share a computational stage. The unity of speech and gesture is also demonstrated by the synchronization of gestures with linguistic units, particularly when there is artificial prolongation of the gesture to maintain synchrony, and by the ability of speakers to transfer linguistic functions between gesture and speech (a) during silences caused by the temporary failures in the computational stage that result in hesitations, and (b) during fluent speech, when part of the propositional content of speech is exhibited gesturally and the rest conveyed verbally. Finally, the unity of gestures and speech is seen in aphasia, where speech and gesture dissolve together, and in children's development, where speech and gesture develop together through the same stages of increasing symbolicization.

The next section of the article presents counterarguments to the proposal that speech and gesture share a computational stage. The final section considers the shared computational stage itself. What can be said about it on the basis of observable gestural and linguistic data?

Counterarguments (and Answers)

In this section I discuss three counterarguments that the reader might also have thought of. Replying to these will clarify the concept of a computational stage shared by speech and gesture. The counterarguments are (a) gestures are like photographs held up while speaking, (b) gestures are translations of sentences into the manual-visual medium, and (c) the shared computational stage of sentence generation is a linear-segmented verbal plan.

Gestures are like photographs held up while speaking. Just as holding up a real photograph would have nothing to do with the process of sentence generation, so performing a gesture could have nothing to do with it either. The first point in reply to this counterargument is that, unlike a photograph, a gesture is something the speaker himself is creating on-line while speaking. Second, the gesture is very closely connected to the sentence temporally, semantically, and pragmatically, all of which point to a coordination between the gesture and the sentence that is quite different from holding up a photograph. Finally, when meaning is divided between a gesture and sentence, it is a true division. It is not that the gesture is called out, as a photograph might be held up, to repair an otherwise interrupted communication. We do not observe, for example,

so he ...

[gesture to complete the idea]

but rather,

so he chases her out again

[gesture to show the means]

These points reduce to one reply: Gesture and speech are operations that have been connected *within*. This is the sense in which they share a computational stage.

The gesture is a translation of the sentence into a different medium. This counterargument runs as follows. The visual-actional medium of the gesture has its own qualities, but these do not characterize the psychological structure of the sentence. There are in fact two psychological structures. The sentence medium has one characteristic set of qualities (linear and segmented), and the gesture medium has another set (global and synthetic or imagistic). When a gesture occurs, the linear-segmented qualities of the sentence medium are translated into the imagistic qualities of the gesture medium. This does not mean that there is a shared computational stage in which both sets of qualities exist simultaneously. On the contrary, it implies that there is computation of the sentence, then a separate computation of the gestural translation.

One way of replying to this counterargument is to show that the gesture medium does *not* impose a characteristic structure of its own, that instead it is a pliable medium capable of assuming the same linear and segmented qualities as the sentence medium. If the gesture medium is free to code meanings segmentally, the imagistic global-synthetic property of gestures cannot be explained as a translation of the sentence medium into the gesture medium.

The following are two cases (among many that could be cited) that demonstrate that the gesture medium is not inherently global and synthetic. In both, the medium normally (in the first case) or spontaneously (in the second) provides a linear and segmented representation. The first is American Sign Language (ASL), a language in which segmentation and linearity are fundamental properties. Gestures in ASL divide events into idea segments, and sentences arrange these segments into linear successions (Klima & Bellugi, 1979). The second case is the behavior of hearing subjects (naive to ASL), who are required to narrate stories without use of speech. Bloom (1979) observed that under these conditions subjects spontaneously subdivide event meanings into idea segments by making a different gesture for each segment. To convey the meaning of an event as a whole, the subjects performed gestures in linear, sentencelike successions.

In view of such cases, one concludes that

The shared computational stage of sentence generation is a linear-segmented verbal plan from which the overt sentence is generated. Spontaneous gestures, too, are generated from the verbal plan:

Covert Verbal Plan

Speech

Gesture

Utterances also exhibit a form of thinking, but in contrast to imagistic thinking, the thinking revealed in speech is linear and segmented. Thus the occurrence of gestures, along with speech, implies that during the act of speaking two kinds of thinking, imagistic and syntactic, are being coordinated. The concept of inner speech explains this coordination. We assume that an image and inner speech symbol are generated at the same time. Inner speech symbols have syntactic implications. These implications and the imagistic properties that the gesture exhibits are jointly present from the start, and can develop internally together. In the final outer speech and gesture streams, each form of output should therefore display properties of the other, and this would occur because they have developed simultaneously from a

common base of inner speech. Examples below demonstrate such an interpenetration of syntactic and imagistic forms of representation.

The concept of inner speech refers to thinking that utilizes words and other linguistic symbols as cognitive tools (Vygotsky, 1962, 1978). A cognitive tool crystallizes the mental operations that are connected with the symbol (Leont'ev, 1979), and inner speech is thinking carried out by means of such mental operations. In the intended use of this concept, the symbols of inner speech are purely mental operations that do not necessarily engage the movements of the speech articulators.

Because of imagery, the mental operations crystallized in the word or larger item of inner speech can be carried out as if they were about entities in space (cf. Pinker, 1984). Words of inner speech are impregnated with what Vygotsky called *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.

An example that comes close to the syntax of inner speech, as Vygotsky described it, is the following:

and finds a big knife

[hand in grip shoots out
to rear and grasps "knife"]

The gesture showed (a) the act of picking up the knife, (b) that it was picked up from the rear, and (c) that the cross-sectional shape of the knife was round (the knife's handle). Yet the gesture synchronized with the single word *finds*, a word conventionally carrying none of this information. Such gestures can be called *holophrastic* (McNeill, in press, discusses holophrastic gestures of children), that is, a gesture that shows an association of a full clause meaning with a single word *within* a grammatical clause.

In my study of children's holophrastic gestures, the words with which the gestures synchronized denoted the contextually novel parts of situations as judged from the narration and the child's choices of referring terms.

If such words are interpreted as cognitive tools, they illustrate what Vygotsky meant by the term *psychological predication*: thinking in terms of the contextually novel. To judge from the speaker's choice of referring terms, the gesture of the "finds" example was comparable. Reference to the agent was suppressed and was made by the maximally presupposing referring "term" \emptyset , a form that altogether rids inner speech of the psychological subject (which in this case was also the grammatical subject).

The form of the speaker's thinking in this example can be inferred from the combination of the mental operations attached to the cognitive tool "finds" and the image of the person leaning back and grasping an object with a round handle. Among the concepts associated with "finds" would be the presupposition that the action is accidental. Thus, the thinking expressed incompletely in the clause, "and finds a big knife," included the idea of an accidental discovery taking place while a particular spatial arrangement obtained—the person groping around behind her for a round-handled object (a referentially equivalent verb like *grabs for* would have been a comment of a quite different kind).

Thus "finds," enriched with the image of the complete event, was plausibly a mental predication about the subject and an appropriate cognitive tool for the speaker's remembering and organizing his thoughts of the event that had taken place.

Can the form of thinking that we infer from gestures and synchronized linguistic items be regarded as the sentence in the early phases of its internal development? In this example a considerable part of the information required for the outer sentence was already present in what we reconstruct of inner speech. Inner speech implied an agent and object in the image of a hand reaching out to grasp something; an object, moreover, small and round (a knife handle) in the shape of the hand; and an agent and object related transitively by the logic of the cognitive tool "finds." Thus the main structural and lexical choices of "and \emptyset finds a knife" were already implied by the organization of the speaker's ongoing thinking, when this utilized "finds" as a symbol of inner speech with the sense exhibited in the gesture.

Not only are the main features of outer speech forecast by inner speech, but the surface form of outer speech might exhibit the imagistic global-synthetic properties of inner speech. Global-synthetic properties do not disappear when inner speech is converted to outer, but reappear in a new linear-segmented form. My final example illustrates the preservation of global-synthetic properties in surface sentence structure (from a 9-year-old boy; but the point of the example does not depend on the fact that the speaker is a child):

and they wanted to get where Anansi was

[both hands held up and facing
each other; the left hand is
motionless, while the right
is "fluttering" back and forth]

(This sentence and gesture occurred during a narration of one of the Anansi stories—folktales of the Ashanti people of Ghana—that the narrator had learned from a filmed presentation.)

To reconstruct inner speech from this gesture and its synchronized speech, we would say that the phrase "they wanted" is shown by the gesture to have been the cognitive tool on this occasion (the pronoun "they" refers to Anansi's six sons). The mental operations crystallized in this phrase, combined with the image that the gesture exhibits, provide a form of thinking that could have retrieved and structured the boy's memory of the event. The verb "wants" presupposes that the action has *not* taken place. "Wants" also permits the references of Anansi and the sons to be in separate clauses. Thus as a cognitive tool "they wanted" provided mental operations geared to the crux of the narrative situation, inaccessibility. These mental operations were utilized to think about inaccessibility in the form of spatial separation.

At the point in the narration where the example occurred, the boy had already explained that Anansi had been swallowed by a fish and that the six sons were setting out to rescue him. The gesture exhibited two objects that stood apart, one of which was motionless, whereas the other fluttered back and forth ineffectually. If the nonmoving object on the left is Anansi imprisoned inside

the fish, the moving object on the right is the sons collectively setting out on the rescue mission, and the lack of closure of the objects, the inaccessibility of Anansi inside the fish. With this interpretation, the gesture can be seen as having metaphorized into a compact concrete symbol all of the ideas that we ponderously express with the words *pursuit* and *inaccessibility*.

The following are some of the global-synthetic parallels that can be noted between the surface grammatical form of the sentence and the synchronized gesture:

1. Two hands appear in the gesture, and two participants are mentioned in the sentence ("they" and "Anansi").

2. The right hand (dominant for this speaker) becomes the thematic reference point of the sentence (correspondingly, the sons are referred to with a presupposing pronoun, "they").

3. The hands are held apart in the gesture, and the references in the sentence to the two participants are contained in different clauses (Clause 1: "they wanted to get somewhere"; Clause 2: "Anansi was somewhere"), syntactic distance reproducing the spatial distance of the gesture. (Mention of the participants within a single noun phrase would have been possible; e.g., "Anansi and the sons couldn't get together.")

4. The right hand is in motion, and the corresponding participants in the sentence are the subject of the verb "get," conveying movement. (Alternatively, the fluttering movement could represent the numerosity of the participants, coded with a plural pronoun.)

5. The left hand is not in motion, and the corresponding participant in the sentence is contained in a locative stative phrase, "where Anansi was," emphasizing lack of movement.

6. The two hands do not close on one another, despite the movement of the right hand, and the subject of the movement verb is also the subject of "want", the use of which presupposes that the action denoted by "get" has not taken place. The choice of this verb also permitted the two-clause structure of the sentence.

The surface form of the sentence, in other words, had many of the same properties as the global-synthetic representation of the ges-

ture. In addition to coding the intentions of Anansi's sons, the surface sentence exhibited the meaning of pursuit plus inaccessibility. Rather than replacing the metaphoric image of two nonclosing objects, the constituents of the sentence *unpacked* this image, the same image presented in a new way. (Other examples also show preservation of global-synthetic properties in surface structure; e.g., an image from the point of view of one who performs the action was unpacked with a choice of active voice and pronominal agent in "and he bends it way back.")

Thus an argument can be put forth that gestures reveal a stage deep within the speaking process. At this level two kinds of thinking are being coordinated: imagistic (thinking that is global and synthetic) and syntactic (thinking that is linear and segmented). This dialectic is suggested by the concept of unpacking a global-synthetic representation with a linear-segmented string that exhibits the same global-synthetic sense. An expression of the dialectic is that there is no system break between thinking and speaking. Grammatical features exist in thought (agentivity, transitivity), and thought features exist in grammar (a global-synthetic image in the surface form of the sentence). If this reveals a language of thought, this language is much more imagistic than Fodor (1975), the author of this evocative phrase, believed. Yet it is capable of explaining the impression of ease and speed of action that one has of normal speech generation. For related and supporting arguments, see Kendon (1983; 1984).

I finally can say why gestures bear the same relation to the verbal process of speaking as does the speech output itself. Gestures and speech share a computational stage that includes the operations of inner speech—thinking utilizing verbal symbols as cognitive tools—and this is also the initial phase of sentence generation. Such is one answer to the hypothetical skeptic addressed in my title.

References

- Argyle, M. (1975). *Bodily communication*. New York: International Universities Press.
- Barroso, F., Freedman, N., Grand, S., & van Meel, J. (1978). Evocation of two types of hand movements in information processing. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 321-329.
- Bates, E., Bretherton, I., Shore, C., & McNew, S. (1983). Names, gestures, and objects: Symbolization in infancy and aphasia. In K. E. Nelson (Ed.), *Children's language* (Vol. 4, pp. 59-123). Hillsdale, NJ: Erlbaum.
- Bloom, R. (1979). *Language creation in the manual modality: A preliminary investigation*. Unpublished undergraduate honors paper, University of Chicago, Chicago.
- Carter, A. L. (1975). The transformation of sensorimotor morphemes into words: A case study of the development of "more" and "mine." *Journal of Child Language*, 2, 233-250.
- Cicone, N., Wapner, W., Foldi, N. S., Zurif, E., & Gardner, H. (1979). The relation between gesture and language in aphasic communication. *Brain and Language*, 8, 324-349.
- Clancy, P. M. (1980). Referential choice in English and Japanese narrative discourse. In W. L. Chafe (Ed.), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production* (pp. 127-202). Norwood, NJ: Ablex.
- Condon, W. S., & Ogston, W. E. (1971). Speech and body motion synchrony in the speaker-hearer. In D. L. Horton & J. J. Jenkins (Eds.), *Perception and language* (pp. 150-173). Columbus, OH: Merrill.
- Cosnier, J. (1982). Communications et langages gestuels [Communication and gesture language]. In J. Cosnier, A. Berrendonner, J. Coulon, & C. Orecchioni (Eds.), *Les voies du langage. Communications verbales gestuelles et animales* (pp. 255-304). Paris: Dunod.
- Curme, G. (1931). *Syntax*. New York: Heath.
- Delis, D., Foldi, N. S., Hamby, S., Gardner, H., & Zurif, E. (1979). A note on temporal relations between language and gestures. *Brain and Language*, 8, 350-354.
- Dimmendaal, G. J. (1983). *The Turkana language*. Dordrecht, Holland: Foris.
- Duncan, S. D., & Fiske, D. W. (1977). *Face-to-face interaction: Research, methods, and theory*. Hillsdale, NJ: Erlbaum.
- Efron, D. (1941). *Gesture and environment: A tentative study of some of the spatio-temporal and "linguistic" aspects of the gestural behavior of Eastern Jews and Southern Italians in New York City, living under similar as well as different environmental conditions*. Morningside Heights, NY: King's Crown Press.
- Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavioral categories—origins, usage, and coding. *Semiotica*, 1, 49-98.
- Fodor, J. A. (1975). *The language of thought*. Cambridge, MA: Harvard University Press.
- Garrett, M. F. (1976). Syntactic processes in sentence production. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 231-256). Amsterdam: North-Holland.
- Gildea, P., & Glucksberg, S. (1983). On understanding metaphor: The role of context. *Journal of Verbal Learning and Verbal Behavior*, 22, 577-590.
- Kendon, A. (1972). Some relationships between body motion and speech. In A. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 177-210). New York: Pergamon Press.
- Kendon, A. (1980). Gesticulation and speech: Two aspects

- of the process of utterance. In M. R. Key (Ed.), *Relationship of verbal and nonverbal communication* (pp. 207-227). The Hague: Mouton.
- Kendon, A. (1981). Geography of gesture. *Semiotica*, 37, 129-163.
- Kendon, A. (1983). Gesture and speech: How they interact. In J. M. Wiemann & R. P. Harrison (Eds.), *Nonverbal interaction* (pp. 13-45). Beverly Hills, CA: Sage Publications.
- Kendon, A. (1984). Some uses of gesture. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 215-234). Norwood, NJ: Ablex.
- Kimura, D. (1973a). Manual activity during speaking: I. Right handers. *Neuropsychologia*, 11, 45-50.
- Kimura, D. (1973b). Manual activity during speaking: II. Left handers. *Neuropsychologia*, 11, 51-55.
- Kimura, D., & Archibald, Y. (1974). Motor functions of the left hemisphere. *Brain*, 97, 337-350.
- Klima, E., & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.
- Kosslyn, S. M. (1975). Information representation in visual images. *Cognitive Psychology*, 7, 341-370.
- Kuno, S. (1976). Subject, theme, and the speaker's empathy—A re-examination of relativization phenomena. In C. N. Li (Ed.), *Subject and topic* (pp. 417-444). New York: Academic Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Leont'ev, A. N. (1979). The problem of activity in psychology. In J. V. Wertsch (Ed. and Trans.), *The concept of activity in Soviet psychology* (pp. 37-71). Armonk, NY: Sharpe.
- Levelt, W. J. M., Richardson, G., & La Heij, W. (in press). Pointing and voicing in deictic expressions. *Journal of Verbal Learning and Verbal Behavior*.
- Levy, E. T. (1984). *Communicating thematic structure in narrative discourse: The use of referring terms and gestures*. Unpublished doctoral dissertation, University of Chicago, Chicago.
- MacDougall, D., & MacDougall, J. (Producers and Directors). (1977). *Lorang's way* [Film]. Berkeley: University of California Extension Media Center.
- McNeill, D. (1979). *The conceptual basis of language*. Hillsdale, NJ: Erlbaum.
- McNeill, D. (in press). Iconic gestures of children and adults. *Semiotica*.
- McNeill, D., & Levy, E. (1982). Conceptual representations in language activity and gesture. In R. Jarvella & W. Klein (Eds.), *Speech, place, and action: Studies in deixis and related topics* (pp. 271-295). Chichester, England: Wiley.
- Miller, G. A., & Johnson-Laird, P. N. (1976). *Perception and language*. Cambridge, MA: Harvard University Press.
- Montaigne, M. de (1958). *The complete essays of Montaigne* (D. Frame, Trans.). Stanford, CA: Stanford University Press.
- Morris, D., Collett, P., Marsh, P., & O'Shaughnessy, M. (1979). *Gestures, their origins and distribution*. New York: Stein & Day.
- Pedelt, L. (1985). *Gestures in aphasia*. Unpublished doctoral dissertation, University of Chicago, Chicago.
- Piaget, J. (1954). *The construction of reality in the child* (M. Cook, Trans.). New York: Basic Books.
- Pinker, S. (1984). Visual cognition: An introduction. *Cognition*, 18, 1-63.
- Polyshyn, Z. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, 80, 1-24.
- Reddy, M. (1979). The conduit metaphor—a case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought* (pp. 284-324). New York: Cambridge University Press.
- Richards, I. A. (1936). *The philosophy of rhetoric*. New York: Oxford University Press.
- Rieber, R. W. (1983). *Dialogues on the psychology of language and thought*. New York: Plenum Press.
- Saussure, F. de (1959). *Course in general linguistics* (W. Baskin, Trans.). New York: Philosophical Library. (Original work published 1916).
- Stephens, D. (1983). *Hemispheric language dominance and gesture hand preference*. Unpublished doctoral dissertation, University of Chicago, Chicago.
- Stephens, D., & Tuite, K. (1983, October). *The hermeneutics of gesture*. Paper presented at the symposium on Gesture at the meeting of the American Anthropological Association, Chicago.
- Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 118-152). New York: Academic Press.
- Vendler, Z. (1967). *Linguistics in philosophy*. Ithaca, NY: Cornell University Press.
- Vygotsky, L. S. (1962). *Thought and language* (E. Hanfmann & G. Vakar, Trans.). Cambridge, MA: MIT Press.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.) Cambridge, MA: Harvard University Press.
- Wepman, J. M. (1977). Aphasia: Language without thought or thought without language? *ASHA*, 18, 131-138.
- Werner, H., & Kaplan, B. (1963). *Symbol formation: An organismic-developmental approach to language and the expression of thought*. New York: Wiley.
- Winner, E., Rosenstiel, A. K., & Gardner, H. (1976). The development of metaphoric understanding. *Developmental Psychology*, 12, 289-297.

Received May 30, 1984

Revision received December 27, 1984 ■