Lexical Gestures and Lexical Access: A Process Model

Robert M. Krauss, Yihsiu Chen and Rebecca F. Gottesman

Columbia University

Lexical Gestures and Lexical Access: A Process Model

Robert M. Krauss, Yihsiu Chen and Rebecca F. Gottesman

Columbia University

Observers of human behavior have long been fascinated by the gestures that accompany speech, and by the contributions to communication they purportedly make.[1] Yet, despite this long-standing fascination, remarkably little about such gestures is well understood. The list of things we don't understand is long, but some of the most important open questions concern their form and function: *why do different gestures take the particular form they do,* and *what it is that these ubiquitous behaviors accomplish?* Traditionally, answers to the function question have focused on the communicative value of gesture, and that view is at least implicit in most contemporary thinking on the topic. Although we do not question the idea that gestures *can* play a role in communication, we believe that their contribution to communication has been overstated, and that the assumption that communication is gesture's only (or even primary) function has impeded progress in understanding the process by which they are generated.

Our goal in this paper is to provide a partial answer to the question of the origin and function of gesture. It is a partial answer because we suspect that different kinds of gestures have different origins and serve different functions, and the model we propose attempts neither to account for all gestural behavior nor for all of the functions gestures serve. Our account also must be regarded as tentative, because formulating a model requires us to make a number of assumptions for which we lack substantial empirical warrant.

---

[1]See Kendon (1982) for an informative review of the history of the study of gesture.

Our concern is with what we will call *lexical gestures,* [2] and we will begin by discussing the kinds of distinctions among gestures we find useful and the functions different types of gestures are hypothesized to serve. Following this, we will describe a model of the process by which lexical gestures are produced, and through which they accomplish what we believe to be their primary function. Then, we will examine some of the empirical evidence that bears on our model, and briefly consider alternative approaches.

<div align="center">

GESTURE TYPOLOGIES AND FUNCTIONS

</div>

The lexical gestures that are the main focus of our model are only one of the kinds of gestures speakers make, and it probably would be a mistake to assume that all gestures are produced by the same process or that they serve the same functions. In the next two sections, we will describe what we believe to be the major kinds of gestures, and the functions the different types of gestures might serve.

<div align="center">

Gesture Types

</div>

Gestural typologies abound in the literature, virtually all of them deriving from the category system initially proposed in Efron's (1941/1972) seminal monograph. We espouse a minimalist approach to gesture categories, based on our belief that some of the proposed typologies make distinctions of questionable utility and reliability. However, we do believe that some distinctions are needed.

Symbolic Gestures

Virtually all typologies distinguish a category of gestural signs—hand configurations and movements with widely-recognized conventionalized meanings—that we will call *symbolic gestures* (Ricci Bitti & Poggi, 1991). Other terms that have been used are *emblems* (Efron, 1941/1972; Johnson, Ekman, & Friesen, 1975), *autonomous gestures* (Kendon, 1983), and *semiotic gestures* (Barakat, 1973). Symbolic

---

2In previous publications (Chawla & Krauss, 1994; Krauss, 1995; Krauss, Chen, & Chawla, 1996) we have used the term lexical *movements* for what we are now calling lexical gestures.

gestures frequently occur in the absence of speech—indeed, they often are used to communicate when distance or noise renders vocal communication impossible—but they also are commonly found as accompaniments to speech, expressing concepts that also are expressed verbally.[3]

Deictic Gestures

A second class of gestures called *deictic gestures* consists of indicative or pointing movements, typically formed with the index finger extended and the remaining fingers closed. Deictic gestures usually are used to indicate persons, objects, directions, or locations, but they may also be used to "point to" unseen, abstract, or imaginary things. Unlike symbolic gestures, which generally have fixed meanings, the "meaning" of a deictic gesture is the act of indicating the things pointed to. They function in a way that is similar to demonstrative pronouns like *this* and *that*. [4] Deictic gestures often accompany speech, but they may also be used to substitute for it. Such autonomous uses may be especially common when the gesture constitutes the response to a question about a location or direction. Often it's easier to reply to a question like "Where's the Psych Department office?" by pointing to the appropriate door than it is to describe its location. Sometimes the speaker will add "It's right over there" if the nonverbal response alone seems too brusque or impolite, but in such cases the verbal message really has little meaning in the absence of the accompany deictic gesture, and, questions of politeness aside, the gesture adequately answers the question.

Motor Gestures

A third gestural type consists of simple, repetitive, rhythmic movements that bear no obvious relation to the semantic content of the accompanying speech

---

3For example, in a videotape of a birthday party we have studied, a man can be seen extending his arm in a "Stop" sign toward another man, who is offering a box of chocolates. At the same time, the gesturer is saying "No, thanks, not right now."

4In the birthday party videotape, a woman points to a child whose face is smeared with cake, exclaiming "Will you look at her!" to her conversational partner, who then looks in the direction she has indicated.

(Feyereisen, Van de Wiele, & Dubois, 1988) . Typically the gesturer's hand shape remains fixed, and the movement may be repeated several times. We will refer to such gestures as *motor gestures;*[5] they also have been called "batons" (Efron, 1941/1972; Ekman & Friesen, 1972) and "beats" (Kendon, 1983; McNeill, 1987) . According to Bull and Connelly (1985) motor gestures are coordinated with the speech prosody and tend to fall on stressed syllables (but see McClave, 1994), although the synchrony is far from perfect.

Lexical Gestures

The three gesture categories we have discussed are relatively uncontroversial, and most gesture researchers acknowledge them as distinct gesture types. The fourth major category is less easy to define and there is considerable disagreement among researcher as to their origins and functions. What we will call *lexical gestures* are similar to what have been called "representational gestures" (McNeill, Cassell, & McCollough, 1994), "gesticulations" (Kendon, 1980; Kendon, 1983), "ideational gestures" (Hadar, Burstein, Krauss, & Soroker, 1998; Hadar & Butterworth, 1997) and "illustrators" (Ekman & Friesen, 1972). Like motor gestures, lexical gestures occur only as accompaniments to speech, but unlike motor gestures they vary considerably in length, are nonrepetitive, complex and changing in form, and many appear to bear a meaningful relation to the semantic content of the speech they accompany.   Providing a formal definition for this class of gestures has proved difficult—Hadar (1989) has defined ideational gestures as hand-arm movements that consist of more than two independent vectorial components—and it is tempting to define them as a residual category: co-speech gestures that are not deictic, symbolic, or motor gestures.

This is not to suggest that lexical gestures are nothing more than what remains in a speaker's gestural output when symbolic, deictic and motor gestures (along with

---

5We have previously called these motor *movements* (Chawla & Krauss, 1994; Krauss, 1995; Krauss et al., 1996).

other, nongestural hand movements) are eliminated. As the model sketched below makes clear, we believe lexical gestures to be a coherent category of movements, generated by a uniform process, that play a role in speech production. It is relevant that although we do not have a mechanical technique that can distinguish lexical gestures from other types of gestures solely on the basis of their form, naive judges can identify them with impressive reliability (Feyereisen, et al., 1988).

Functions of Gestures

Communication

The traditional view that gestures function as communicative devices is so widespread and well-accepted that comparatively little research has been directed toward assessing the magnitude of gestures' contribution to communication, or ascertaining the kinds of information different types of gestures convey. Reviewing such evidence as exists, Kendon has concluded that:

> The gestures that people produce when they talk do play a part in communication and they do provide information to co-participants about the semantic content of the utterances, although there clearly is variation about when and how they do so (Kendon, 1994, p. 192).

After considering the same studies, we have concluded that the evidence is inconclusive at best, and is equally consistent with the view that the gestural contribution is, on average, negligible. (See Krauss, Dushay, Chen & Rauscher, 1995 for a discussion of the evidence.) We will return to this issue below, in our discussion of the origins of lexical gestures.

Tension Reduction

Noting that people often gesture when they are having difficulty retrieving an elusive word from memory, Dittmann and Llewelyn (1969) have suggested that at least some gestures may be more-or-less random movements whose function is to dissipate tension during lexical search. The idea is that failures of word retrieval are frustrating, and that unless dealt with, this frustration-generated tension could interfere

with the speaker's ability to produce coherent speech. Hand movements provide a

means for doing this. Other investigators (Butterworth, 1975; Freedman & Hoffman,

1967) have noted the cooccurrence of gestures and retrieval problems, although they

have not attributed this to tension management. We are unaware of any evidence

supporting the hypothesis that people gesture in order to reduce tension, and find the

idea itself somewhat implausible. However, we do think that the association of

gesturing and word retrieval failures is noteworthy.

Lexical Retrieval

Gesturing may be common when speakers are trying to access words from their

lexicons because it plays a direct role in the process of lexical retrieval. This is not a new

idea and has been suggested by a remarkably diverse group of scholars over the last 75

years (DeLaguna, 1927; Freedman, 1972; Mead, 1934; Moscovici, 1967; Werner &

Kaplan, 1963). Empirical support for the notion that gesturing affects lexical access is

mixed. In the earliest published study, Dobrogaev (1929) reported that speakers

instructed to curb facial expressions, head movements, and gestural movements of the

extremities found it difficult to produce articulate speech, but the experiment apparently

lacked the necessary controls.[6] More recently, Graham and Hayward (1975) analyzed

the speech of five speakers who were prevented from gesturing as they described

abstract line drawings, and concluded that "… elimination of gesture has no particularly

marked effects on speech performance" (p. 194).   On the other hand, Rimé (1982) and

Rauscher, Krauss and Chen (1996) have found that restricting gesturing adversely

affects speech. See Krauss, et al. (1996) for a review of the relevant studies.

Although the idea that gesturing facilitates lexical retrieval is not a new one, none

of the writers who have suggested this possibility has described the mechanism by

which gestures affect lexical access. The model presented below describes an

---

6As was commonly done in that era, Dobrogaev fails to report procedural details, and
describes his results in impressionistic, qualitative terms.

architecture that accounts for both the production of lexical gestures and their
facilitative effects on lexical retrieval.

THE PRODUCTION OF SPEECH AND GESTURE

Speech Production

We assume that lexical gestures and speech involve two production systems that
operate in concert. It may be helpful to begin by reviewing briefly our understanding
of the process by which speech is generated. Of course, the nature of the speech
production is not uncontroversial, and several production models have been proposed.
These models differ in significant ways, but for our purposes their differences are less
important than their similarities. Although we will follow the account of Levelt (1989) ,
all of the models with which we are familiar distinguish three stages of the process.
Levelt (1989) refers to the three stages as *conceptualizing, formulating,* and *articulating.*

Conceptualizing involves, among other things, drawing upon declarative and
procedural knowledge to construct a communicative intention. The output of the
conceptualizing stage—what Levelt refers to as a *preverbal message*—is a conceptual
structure containing a set of semantic specifications. At the formulating stage, the
preverbal message is transformed in two ways. First, a grammatical encoder maps the
to-be-lexicalized concept onto a lemma (i.e., an abstract symbol representing the
selected word as a semantic-syntactic entity) in the mental lexicon whose meaning
matches the content of the preverbal message. Using syntactic information contained
in the lemma, the conceptual structure is transformed into a *surface structure* (see also
Bierwisch & Schrueder, 1992) . Then, by accessing word forms stored in lexical memory
and constructing an appropriate plan for the utterance's prosody, a phonological
encoder transforms this surface structure into a *phonetic plan* (essentially a set of
instructions to the articulatory system). The output of the articulatory stage is overt
speech, which the speaker monitors and uses as a source of corrective feedback. The

process is illustrated schematically in the right-hand (shaded) portion of Figure 1 below,
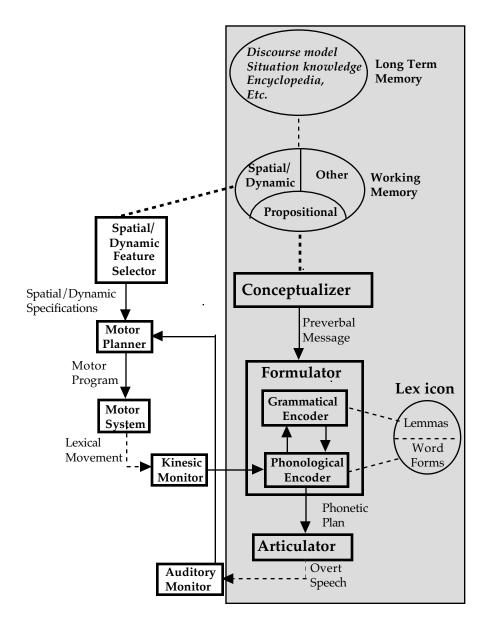
which is based on Levelt (1989).

Figure 1. A cognitive architecture for the speech-gesture production process (speech processor redrawn from Levelt, 1989).

Gesture Production

Our account of the origins of gesture begins with the representations in working

memory that come to be expressed in speech. An example may help explicate our

view. Kendon (1980) describes a speaker saying "…with a big cake on it…" while

making a series of circular motion of the forearm with the index finger extended

pointing downward.   In terms of the speech production model, the word "cake" derives

from a memorial representation we will call the *source concept*. In this example, the

source concept is the representation of a particular cake in the speaker's memory. The

conceptual representation that is outputted by the conceptualizer and that the

grammatical encoder transforms into a linguistic representation typically incorporates

only a subset of the source concept's features.[7] Presumably the particular cake in the

example was large and round, but it also had other properties—color, flavor, texture,

and so on—that might have been mentioned but weren't presumably because, unlike

the cake's size, they were not relevant to the speaker's goals in the discourse. A central

theoretical questions is whether or not the information that the cake was round—i.e,.

the information contained in the gesture—was part of the speaker's communicative

intention. We assume that it was not. Below we will consider some of the implications

of assuming that such gestures are communicatively intended.

In developing our model we have made the following assumptions about

memory and mental representation:

(1) Memory employs a number of different formats to represent knowledge, and much of the contents of memory is multiply encoded in more than one representational format.

(2) Activation of a concept in one representational format tends to activate related concepts in other formats.

(3) Concepts differ in how adequately (i.e., efficiently, completely, accessibly, etc.) they can be represented in one or another format. The complete mental representation of some concepts may require inputs from more than one representational format.

7The difference between the source concept and the linguistic representation can most clearly be seen in reference, where the linguistic representation is formulated specifically to direct a listener's attention to some thing, and typically will incorporate only as much information as is necessary to accomplish this. Hence one may refer to a person as "the tall guy with red hair," an expression that incorporates only a few features of the far more complex and differentiated conceptual representation.

      (4) Some representations in one format can be translated into the
          representational form of another format (e.g., a verbal description can give
          rise to a visual image, and vice-versa).

None of these assumptions is particularly controversial, at least at this level of

generality.

      We follow Levelt in assuming that inputs from working memory to the

conceptualizing stage of the speech processor must be in propositional form. However,

the knowledge that constitutes a source concept may be multiply encoded in both

propositional and nonpropositional representational formats, or encoded exclusively in

nonpropositional formats. In order to be reflected in speech, nonpropositionally

encoded information must be "translated" into propositional form.

      We have tried to illustrate this schematically in Figure 2. In that figure, the

hypothetical source concept A is made up of a set of features (Features A_1 - A_{10}) that

are encoded in propositional and/or in spatial format. Some features (e.g., A_1 and A_4

are multiply encoded in both formats. Others are represented exclusively in

propositional (e.g., A_3 and A_6) or in spatial form (A_2 and A_5) form. Our central

hypothesis is that *lexical gestures derive from nonpropositional representations of the source

concept.* Just as a linguistic representation may not incorporate all of the features of the

source concept's mental representation, lexical gestures reflect these features even more

narrowly. The features they incorporate are primarily spatio-dynamic.   As Figure 2

illustrates, in the hypothetical example, the lexical item the speaker selects to represent

concept A incorporates six of its propositionally-encoded features. The gesture that

accompanies it incorporates three features, two of which are also part of the lexical
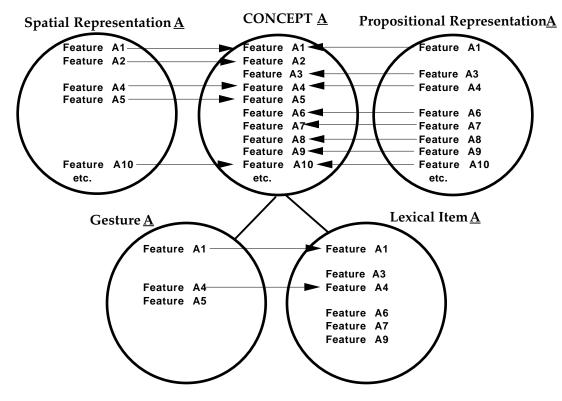
representation.

Figure 2.   Mental representation of a hypothetical source concept and its reflection in speech and gesture

How do these nonpropositionally represented features come to be reflected gesturally? Our model assumes that a spatial/dynamic feature selector transforms information stored in spatial or dynamic formats into a set of *spatial/dynamic specifications*—essentially abstract properties of movements.[8] These abstract specifications are, in turn, translated by a motor planner into a *motor program* that provides the motor system with a set of instructions for executing the lexical gesture. The output of the motor system is a gestural movement, which is monitored kinesthetically. The model is shown as the left-hand (unshaded) portion of Figure 1.

_____

[8]We can only speculate as to what such abstract specifications might consist of. One might expect articulation of lexical items representing concepts that incorporate the features STRAIGHT and CURVED to be accompanied by rectilinear and curvilinear movements, respectively, that lexical items representing concepts that incorporate the feature FAST would be represented by rapid movements, that lexical items representing concepts that incorporate the feature LARGE would be represented by movements with large linear displacements, etc. However, we are unaware of any successful attempts to establish systematic relations between abstract dimensions of movement and dimensions of meaning. For a less-than-successful attempt, see Morrel-Samuels (1989.

Gestural Facilitation of Speech

We believe that an important function of lexical gestures is to facilitate lexical retrieval. How does the gesture production system accomplish this? The process is illustrated in Figure 1, which shows the output of the gesture production system, mediated by the kinesic monitor, feeding into the phonological encoder where it facilitates retrieval of the word form.[9]

In our model, the lexical gesture provides input to the phonological encoder via the kinesic monitor. The input consists of features of the source concept represented in motoric or kinesic form. As discussed above, the features contained in the lexical gesture may or may not also be features of the sought-for lexical item, which is to say that they may or may not be elements of the speaker's communicative intention. These features, represented in motoric form, facilitate retrieval of the word form by a process of cross modal priming. This is represented in Figure 1 by the path from the kinesic monitor to the phonological encoder. The figure also shows a path from the auditory monitor to the motor planner. This path (or its equivalent) is necessary in order to account for gesture termination. Since a gesture's duration is closely timed to articulation of its lexical affiliate (Morrel-Samuels & Krauss, 1992), a mechanism that informs the motor system when to terminate the gesture is required. Essentially, we are proposing that hearing the lexical affiliate being articulated serves as the signal to terminate the gesture.[10]

9This represents a minor change from an earlier version of the model (Krauss et al., 1996), in which the output of the gesture production system went to the grammatical encoder.

10 Hadar and (in press) propose a somewhat different mechanism for gesture initiation and termination: lexical movements are initiated by failures in lexical retrieval and terminated when the sought-for word is accessed. In our architecture, the two systems would appear quite similar structurally, but their operation would be somewhat different. For the Hadar and Butterworth, the path from the phonological encoder to the motor planner would carry a signal that lexical access had not occured and would initiate the gesture; the termination of the signal would terminate the gesture.

SOME THEORETICAL ALTERNATIVES

Formulating a theoretical model requires choosing among alternative conceptualizations, often in the absence of compelling evidence on which to base those choices. In this section we will examine a number of the theoretical choices we made and indicate our reasons for choosing them.

Autonomous vs. Interactive Processes

A question that any model attempting to account for the production of gesture and speech must address is how the two systems function relative to each other. At the most fundamental level, the two systems can be either *autonomous* or *interactive.* Autonomous processes operate independently once they have been initiated; interactive systems can affect each other during the production process. Clearly we have opted for an interactive model. In our view, an autonomous model is inconsistent with two kinds of data.

(1) *Evidence from studies of the temporal relations of gesture and speech:* It is reasonably well established that lexical gestures precede their "lexical affiliates"—the word or phrase in the accompanying speech to which they are related. (Butterworth & Beattie, 1978; Morrel-Samuels & Krauss, 1992; Schegloff, 1984) . Morrel-Samuels and Krauss (1992) examined 60 carefully selected lexical gestures and found their asynchrony (the time interval between the onset of the lexical gesture and the onset of the lexical affiliate) to range from 0 - 3.75 s, with a mean of 0.99 s and a median of 0.75 s; none of the 60 gestures was initiated after articulation of the lexical affiliate had begun. The gestures' durations ranged from 0.54 s to 7.71 s. (mean = 2.49 s), and only three of the 60 terminated before articulation of the lexical affiliate had begun. The product-moment correlation between gestural duration and asynchrony is +0.71. We can think of only two ways that such a temporal relation could exist without interaction between the gesture and speech production systems: (a) gestures of long duration are associated with unfamiliar words; (b) speakers somehow can predict *a priori* how long it will take

them to retrieve a particular word form, and enter this variable into the formula that determines the gesture's duration. We know of no data that supports either proposition, and neither strikes us as plausible.

(2) *Evidence of the effects of preventing gesturing on speech:* Rauscher, Krauss and Chen (1996) found that preventing speakers from gesturing reduced the fluency of their speech with spatial content, compared to their speech when they could gesture. Non-spatial speech was unaffected. Rauscher et al. interpret this finding as support for the proposition that gesturing facilitates lexical access, something that is incompatible with an autonomous model of speech and gesture production. The finding by Frick-Horbury and Guttentag (in press) that preventing gesturing increases the rate of retrieval failures in a Tip-of-the-Tongue (TOT) situation also is relevant.

To explain the association of gesturing with lexical retrieval failures, De Ruiter (this volume) hypothesizes that speakers anticipating problems in accessing a lexical item will try "…to compensate for the speech failure by transmitting a larger part of the communicative intention in the gesture modality" (p. $$). However, Frick-Horbury and Guttentag (in press) found that subjects were likely to gesture in a TOT state—i.e., unable to retrieve a sought-for word. It would be odd if these gestures were intended for the benefit of the addressee, since the addressee (the experimenter) already knew what the target word was.

If one takes De Ruiter's proposal seriously, it would follow that the gestures speakers make when they can't be seen by their conversational partner are different from those that are visually accessible. since the former couldn't possibly transmit information. Not much relevant data exists, but the little we are aware of is not supportive. For example, we have coded the grammatical types of 12,425 gestural lexical affiliates from a previously reported experiment (Krauss, Dushay, Chen, & Bilous, 1995). In that experiment, subjects described abstract graphic designs and novel synthesized sounds to a partner who was either seated face-to-face with the speaker or

in another room, using an intercom. Not surprisingly, the grammatical categories of gestural lexical affiliates differed reliably depending on whether the stimulus described was an abstract design or a sound. However, they did not differ as a function of whether or not the speaker could be seen by their partners (Krauss, Gottesman, Chen & Zhang, in preparation).

It also is relevant that people speaking a language they have not mastered do not attempt to compensate for their deficiency by gesturing. Dushay (1991) found that native English speakers taking second year Spanish gestured significantly *less* when speaking Spanish than when they spoke English in a referential communication task, and allowing listeners to see speakers' gestures did not enhance their communicative effectiveness.

Certainly it is possible that speakers rely on gestures to convey information they cannot convey verbally, as de Ruiter contends, but we know of no attempt to ascertain whether (or how frequently) they do. Our guess is that it is not a common occurrence, and another explanation for the frequency of lexical gestures is needed.

Conceptualizer vs. Working Memory Origins

Our model specifies the speaker's working memory as the source of the representations that come to be reflected in lexical gestures. An alternative source is the speech processor itself. Logically, there are two stages at which this could occur: the conceptualizing stage, where the preverbal message is constructed, and the formulating stage (more specifically during grammatical encoding) where the surface structure of the message is constructed. We regard both as dubious candidates. To explain why, we need to consider the nature of the information in the two systems.

The speech processor is assumed to operate on information that is part of the speaker's communicative intention. If that is so, and if gestures originate in the speech processor, gestural information would consist *exclusively* of information that was part of the communicative intention. However, we believe that gestures can contain

information that is not part of the communicative intention expressed in speech. Recall Kendon's previously described example of the speaker saying "…with a big cake on it…" accompanied by a circular motion of the forearm.   Although it may well have been the case that the particular cake the speaker was talking about was round, ROUND is not a semantic feature of the word *cake* (cakes come in a variety of shapes), and for that reason ROUND was not be part of the speaker's communicative intention *as it was reflected in the spoken message.*

The reader may wonder (as did a reader of an earlier draft of this chapter) how a feature that was not part of the word's feature set could aid in lexical retrieval. For example, if ROUND is not a feature of *cake,* how could gesture incorporating that feature help the speaker access the lexical entry? We agree that it could not. and we do not believe that all of the lexical gestures speakers make facilitate retrieval, or that speakers make them with that purpose in mind. It is only when the gesture derives from features that are part of the lexical item's semantic that gesturing will have this facilitative effect. In our view, gestures are a product of the same conceptual processes that ultimately result in speech, but the two production systems diverge very early on. Because gestures reflect representations in memory, it would not be surprising if some of the time the most accessible features of those representations (i.e., the ones that are manifested in gestures) were products of the speaker's unique experience and not part of the lexical entry's semantic. Depending on the speaker's goals in the situation these features may or may not be relevant. When they are, we would expect the speaker to incorporate them into the spoken message; when they are not, we would expect the spoken message and the gesture to contain different information. Note that in Kendon's example, the speaker elected to incorporate the feature LARGE (which could have been represented gesturally) but not ROUND in the verbal message. In some cases, an attentive addressee will be able to interpret the gestural information (e.g., to discern that the case was both large and round), although the bulk of the gestures

speakers make are difficult to interpret staightforwardly (Feyereisen et al., 1988; Krauss, Morrel-Samuels, & Colasante, 1991). Evidence that viewers are able to do this (e.g., McNeill et al., 1994) really does not address the question of whether such gestures are communicatively intended.

Kendon and others e.g., (Clark, 1996; deRuiter, this volume; Schegloff, 1984) implicitly reject this view of gestures' function , and assume that speakers partition the information that constitutes their communicative intentions, conveying some information verbally in the spoken message, some visibly, via gesture, facial expression, etc., and some in both modalities. In the *cake* example, Kendon presumably would have the speaker *intending* to convey the idea that the cake was both big *and* round, and choosing to convey ROUND gesturally. No doubt that speakers can do this, and occasionally they probably do.[11] However, we see no reason to assume that *most* of the speech-accompanying gestures speakers make are communicatively intended. Apart from the intuitive impressions of observers, we know of no evidence that directly supports this notion, and much of the indirect evidence is unsupportive. For example, speakers gesture when their listener's cannot see them, albeit somewhat less than they do when they are visually accessible (Cohen, 1977; Cohen & Harrison, 1972; Krauss et al., 1995). Gestures that are unseen (e.g., made while speaking on the telephone) cannot convey information, and yet speakers make them.[12] In addition, experimental findings suggest that on average lexical gestures convey relatively little information (Feyereisen et al., 1988; Krauss, et al., 1991, experiments 1 and 2), that their contribution to the semantic interpretation of the utterance is negligible (Krauss et al., 1991, experiment 5), and that having access to them does not enhance the effectiveness of communication, as that is indexed by a referential communication task (Krauss et al.,

---

11This is often the case with deictic gestures. For example, a speaker may say "You want to go through that door over there," and point at the particular door.

12As we have noted above, the distribution of grammatical types of the lexical affiliates of gestures that cannot be seen do not appear to differ from those of gestures that can be seen.

1995) . So if such gestures are communicatively intended, the evidence we have

suggests that they are not notably effective.

De Ruiter argues that the occurrence of gesturing by speakers who cannot be

seen and findings that gestures have relatively little communicative value do not reduce

the plausibility of the idea that they are communicatively intended.

> Gesture may well be intended by the speaker to communicate, and fail to do so
> in some or even most cases. The fact that people gesture on the telephone is also
> not necessarily in conflict with the view that gestures are generally intended to
> be communicative. It is conceivable that people gesture on the telephone
> because they always gesture when they speak spontaneously—they simply
> cannot suppress it (DeRuiter in press, p. $$).

The idea that such gestures reflect overlearned habits is not intrinsically

implausible. However, the notion that gestures are both communicatively intended *and*

largely ineffective is incompatible with a modern understanding of how language (and

other behaviors) are used communicatively. The argument is rather an involved one,

and space constraints prevent us from reviewing it here.[13] Suffice it to say that De

Ruiter is implicitly espousing a view of communication that conceptualizes participants

as autonomous information processors. Such a view stands in sharp contrast with what

Clark (1996) has termed a collaborative view of language use, in which communicative

exchange is a *joint* accomplishment of the participants, who work together to achieve

some set of communicative goals. From the collaborative perspective, speakers and

hearers endeavor to ensure that they have similar conceptions of the meaning of each

message before they proceed to the next one. The idea that some element of a message

is communicatively intended, but consistently goes uncomprehended, makes little sense

from such a perspective.

We do not believe that De Ruiter (and others who share the view that gestures

are communicatively intended) can have it both ways. If such gestures are part of the

---

13Some of the relevant issues are discussed by Clark (1996) and Krauss and Fussell
(1996).

communicative intention and convey information that is important for constructing

that intention, speaker/gesturers should make an effort to insure that gestural

meanings have been correctly apprehended, and it should be possible to demonstrate

straightforwardly that the gestures make a contribution to communication. Again, at

risk of being redundant, let us reiterate our belief that there certainly are some gestures

(e.g., symbolic gestures, some iconic gestures, some pantomimic-enactive gestures) that

are both communicatively intended *and* communicatively effective; that much is not

controversial. The question we raise is whether there is adequate justification for

assuming that all or most co-speech gestures are so intended. In our view, such a

justification is lacking.

The issue of whether we should assume that gestures are part of the speaker's

communicative intention is not simply an esoteric theoretical detail. The assumption is

important precisely because of the implications it has for an explanation of gestural

origins. If gestures are communicatively intended, they must originate within the

speech processor—in the conceptualizer or formulator. If they are not, they can

originate elsewhere. Note that an origin outside the speech processor would not

preclude gestures from containing information that is part of the communicative

intention, or from conveying such information to a perceptive observer.[14] However. it

would constrain a rather different sort of cognitive architecture.

Does Gestural Facilitation Affect Retrieval of the Lemma or Lexeme?

Lexical retrieval is a two-step process, and difficulties might be encountered at

either stage—during grammatical encoding (when the lemma is retrieved) or during

phonological encoding (when the word form or lexeme is retrieved). Do lexical

gestures affect retrieval of the lemma, the word form or both? At this point there is

---

14Of course, it is possible that gestures convey information (e.g., about the speaker's internal state) that is *not* part of the communicative intention but may be of value to the addressee, in much the same way that an elevated pitch level conveys information about a speaker's emotional state. We have explored this idea elsewhere (Chawla & Krauss, 1994; Krauss et al., 1996).

some evidence that gestures affect the phonological encoder, although the evidence
does not preclude the possibility that they also can affect retrieval of the lemma. Our
choice of the phonological encoder is based on findings from research on lexical
retrieval, especially studies using the "tip of the tongue" (TOT) paradigm, that retrieval
failures in normal subjects tend to be phonological rather than semantic (Brown, 1991;
Brown & McNeill, 1966; Jones, 1989; Jones & Langford, 1987; Kohn, Wingfield, Menn,
Goodglass, & al., 1987; Meyer & Bock, 1992). It is especially significant that preventing
gesturing increases retrieval failures in the TOT situation (Frick-Horbury & Guttentag,
in press). Since the definitions by means of which the TOT state is induced are roughly
equivalent to the information in the lemma, the finding suggests that preventing
subjects from gesturing interferes with access at the level of the word form.[15]

Lexical, Iconic and Metaphoric Gestures

The movements we are calling lexical gestures often are partitioned into
subcategories, although there is relatively little consensus as to what those
subcategories should include. Probably the most widely accepted is the subcategory of
"iconic gestures"—gestures that represent their meanings pictographically, in the sense
that the gesture's form is conceptually related to the semantic content of the speech it
accompanies. For example, in a videotape of a tv program about regional dialects we
have studied, the sociolinguist Roger Shuey, explaining how topographic features come
to mark boundaries between dialects, is shown saying: "… if the settlers were stopped
by a natural barrier or boundary, *such as a mountain range* or a river…"   As he
articulates the Italicized words, his right hand traces a shape resembling the outline of a
mountain peak.

However, not all lexical gestures are iconic. Many speech-accompanying
movements clearly are not deictic, symbolic or motor gestures, but seem to have no

---

15 However, since the grammatical and phonological encoders interact in the
Formulator module (see Figure 1), information inputted to the grammatical encoder could affect
retrieval of the lexeme.

obvious formal relationship to the conceptual content of the accompanying speech.[16]

What can be said of these gestures? McNeill (1985; 1992) deals with this problem by

drawing a distinction between iconic and *metaphoric* gestures.

> Metaphoric gestures exhibit images of abstract concepts. In form and manner of
> execution, metaphoric gestures depict the vehicles of metaphors...The metaphors
> are independently motivated on the basis of cultural and linguistic knowledge
> (McNeill, 1985, p. 356).

Despite its widespread acceptance, we have reservations about the utility of such

distinctions. Our own observations have lead us to conclude that iconicity (or *apparent*

iconicity) is a matter of degree rather than kind. While the form of some lexical

gestures do seem to have a direct and transparent relationship to the content of the

accompanying speech, for others the relationship is more tenuous, and for still others

the observer is hard-pressed to find any relationship at all. So, in our view, it makes

more sense to think of gestures as being more or less iconic rather than either iconic or

metaphoric (or noniconic). Moreover, the iconicity of many gestures seems to exist

primarily in the eyes of their beholders. Viewers will disagree about the "meanings"

they impute to gestures, even when viewing them along with the accompanying

speech.[17] In the absence of speech, their imputed meanings evidence little iconicity

(Feyereisen et al., 1988; Krauss et al., 1991).

We find the iconic/metaphoric distinction even more problematic.   The

argument that metaphoric gestures are produced in the same way linguistic metaphors

---

[16]The proportion of lexical movements that are iconic is difficult to determine and
probably depends greatly on the conceptual content of the speech.

[17]We have found considerable disagreement among naive viewers on what apparently
iconic gestures convey. As part of a pretest for a recent study, we selected 140 gestures that seem
to us iconic from narratives on a variety of topics (e.g., directions to campus destinations,
descriptions of the layouts of apartments, instructions on how to make a sandwich, etc.).
Following the method used by Morrel-Samuels and Krauss (1992), subjects saw the video clips
containing the gestures and heard the accompanying speech, and underlined the gestures'
lexical affiliates on a transcript. Despite the fact that these gestures had been selected
because we judged them to be iconic, subjects' agreement on their lexical affiliates averaged
43.86% (SD=23.34%). On only 12% of the gestures was there agreement among 80% or more of
our subjects. These results suggest that even iconic gestures convey somewhat different things to
different people, making them an unreliable vehicle for communication.

are generated does not lead to a tractable explanation of the process by which these gestures are produced. Since our understanding of the processes by which linguistic metaphors are produced and comprehended is incomplete at best (cf. (Glucksberg, 1991; Glucksberg, in press), to say that such gestures are visual metaphors may be little more than a way of saying that their iconicity is not obvious.

In what ways might such classifications be useful? The idea that a gesture is iconic provides a principled basis for explaining why it takes the form it does. Presumably the form of Shuey's gesture was determined by the prototypical form of the concept with which it was associated—the mountain range. Calling a gesture metaphoric can be seen as an attempt to accomplish the same thing for gestures that lack a formal relationship to the accompanying speech. We believe that a featural model provides a more satisfactory way of accounting for a gesture's form.

Featural vs. Imagistic Models

Our model assumes the memorial representations that come to be reflected in gesture are made up of sets of elementary features. This "compositional" view of concepts is an old one in psychology, and it is not without its problems (see Smith and Medin, 1981 for a discussion of theories of the structure of concepts). An alternative to featural models of gestural origins is a model that views memorial representations as integrated and nondecomposible units that are retrieved holistically. We will refer to these as *imagistic* models. This approach has been employed by De Ruiter (this volume) , Hadar and Butterworth (in press) and McNeill (1992; this volume), among others.

We find a number of problems with imagistic models of gesture production. In the first place, specifying an imagistic origin of a gesture does not eliminate the need for a principled account of how the image comes to be represented by a set of hand movements. That is to say, some mechanism must abstract relevant aspects of the image and "translate" it into a set of instructions to the motor system. This is the function of the "sketch generation" module in DeRuiter's Sketch Model. Secondly,

images are concrete and specific—one cannot have an image of a generic cake. Finally, although imagistic models of gesture production may offer a plausible account of the production of so-called iconic gestures, many gestures lack apparent *physical isomorphy* with the conceptual content of the speech they accompany, among them the class of gestures McNeill has termed *metaphoric.* It is not obvious how an imagistic production model can account for the production of such gestures.

CONCLUDING COMMENT

It probably is unnecessary for us to point out that the cognitive architecture we describe is only one of a number of possible arrangements that would be consistent with what we know about gestures. As things currently stand, there is such a dearth of firm data on gesture production to constrain theory that any processing model must be both tentative and highly speculative. Nevertheless, we believe that model building in such circumstances is not an empty activity. In the first place, models provide a convenient way of systematizing available data. Secondly, they force theorists to make explicit the assumptions that underlie their formulations, making it easier to assess in what ways, and to what extent, apparently different formulations actually differ. Finally, and arguably most importantly, models lead investigators to collect data that will confirm or disconfirm one or another model.

We have few illusions that our account is correct in every detail, or that all of the assumptions we have made will ultimately prove to have been justified. Indeed, our own view of how lexical gestures are generated has changed considerably over the last several years as data has accumulated. Nor do we believe that our research strategy, which rests heavily on controlled experimentation, is the only one capable of providing useful information. Experimentation is a powerful method for generating certain kinds of data, but it also has serious limitations, and an investigator who does not recognize these limitations may be committing the same error as the *savants* in the parable of the blind men and the elephant. Observational studies have enhanced our understanding

of what gestures accomplish, and the conclusions of careful and seasoned observers deserve to be taken seriously.

At the same time, we are uncomfortable with the idea that experimental evidence is irrelevant to some propositions about gesture, whose validity can only be established by interpretations of observed natural behavior. Ultimately we look for the emergence of an account of the process by which gestures are generated, and the functions they serve, that is capable of accommodating the results of both experimental and systematic-observational studies.

Acknowledgments

References

Barakat, R. (1973). Arabic gestures. *Journal of Popi;ar Culture*, **6**, 749-792.

Bierwisch, M., & Schrueder, R. (1992). From concepts to lexical items. *Cognition* , **42**, 23-60.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin* , **109**, 204-223.

Brown, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, **4**, 325-337.

Bull, P., & Connelly, G. (1985). Body movement and emphasis in speech. *Journal of Nonverbal Behavior*, **9**, 169-187.

Butterworth, B. (1975). Hesitation and semantic planning in speech. *Journal of Psycholinguistic Research*, **4**, 75-87.

Butterworth, B., & Beattie, G. (1978). Gesture and silence as indicators of planning in speech. In R. N. Campbell & P. T. Smith (Ed.), *Recent Advances in the Psychology of Language: Formal and experimental approaches* . New York: Plenum.

Chawla, P., & Krauss, R. M. (1994). Gesture and speech in spontaneous and rehearsed narratives. *Journal of Experimental Social Psychology* , **30**, 580-601.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Cohen, A. A. (1977). The communicative functions of hand illustrators. *Journal of Communication*, **27**, 54-63.

Cohen, A. A., & Harrison, R. P. (1972). Intentionality in the use of hand illustrators in face-to-face communication situations. *Journal of Personality and Social Psychology*, **28**, 276-279.

DeLaguna, G. (1927). *Speech: Its function and development*. New Haven: Yale University Press.

deRuiter, J. P. (this volume). The production of gesture and speech. In A. Kendon, D. McNeill & S. Wilcox (Ed.), *Gesture: An emerging field* . New York: Cambridge University Press.

Dittmann, A. T., & Llewelyn, L. G. (1969). Body movement and speech rhythm in social conversation. *Journal of Personality and Social Psychology* , **23**, 283-292.

Dobrogaev, S. M. (1929). Ucnenie o reflekse v problemakh iazykovedeniia [Observations on reflexes and issues in language study]. *Iazykovedenie i Materializm*, , 105-173.

Dushay, R. D. (1991). *The association of gestures with speech: A reassessment.* Unpublished doctoral dissertation, Columbia University.

Efron, D. (1941/1972). *Gesture, race and culture.* The Hague: Mouton (first edition 1941).

Ekman, P., & Friesen, W. V. (1972). Hand movements. *Journal of Communication*, **22**, 353-374.

Feyereisen, P., Van de Wiele, M. ., & Dubois, F. (1988). The meaning of gestures: What can be understood without speech? *Cahiers de Psychologie Cognitive*, **8**, 3-25.

Freedman, N. (1972). The analysis of movement behavior during the clinical interview. In A. W. Siegman & B. Pope (Ed.), *Studies in dyadic communication* (pp. 153-175). New York: Pergamon.

Freedman, N., & Hoffman, S. (1967). Kinetic behavior in altered clinical states: Approach to objective analysis of motor behavior during clinical interviews. *Perceptual and Motor Skills*, **24**, 527-539.

Frick-Horbury, D., & Guttentag, R. E. (in press). The effects of restricting hand gesture production on lexical retrieval and free recall. *American Journal of Psychology.*

Glucksberg, S. (1991). Beyond literal meanings: The psychology of allusion. *Psychological Science*, **2**, 146-152.

Glucksberg, S. (in press). How metaphors work. In A. Ortony (Ed.), *Metaphor and thought (2nd edition)* . Cambridge, Eng.: Cambridge University Press.

Graham, J. A., & Heywood, S. (1975). The effects of elimination of hand gestures and of verbal codability on speech performance. *European Journal of Social Psychology* , **5**, 185-195.

Hadar, U. (1989). Two types of gesture and their role in speech production. *Journal of Language and Social Psychology*, **8**, 221-228.

Hadar, U., Burstein, A., Krauss, R. M., & Soroker, N. (. (1998). Ideational gestures and speech: A neurolinguistic investigation. *Language and Cognitive Processes*, **13**, 56-76.

Hadar, U., & Butterworth. (1997). Iconic gestures, imagery and word retrieval in speech. *Semiotica,* **115,** 147-172.

Johnson, H., Ekman, P., & Friesen, W. V. (1975). Communicative body movements: American emblems. *Semiotica*, **15**, 335-353.

Jones, G. V. (1989). Back to Woodworth: Role of interlopers in the tip-of-the-tongue phenomenon. *Memory and Cognition* , **17**, 69-76.

Jones, G. V., & Langford, S. (1987). Phonological blocking in the tip of the tongue state. *Cognition* , **26**, 115-122.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *Relationship of verbal and nonverbal communication* . The Hague: Mouton.

Kendon, A. (1983). Gesture and speech: How they interact. In J. M. Weimann & R. P. Harrison (Ed.), *Nonverbal interaction* . Beverly Hills, CA: Sage.

Kendon, A. (1994). Do gestures communicate?: A review. *Research on Language and Social Interaction* , **27**, 175-200.

Kohn, S. E., Wingfield, A., Menn, L., Goodglass, H., & al., e. (1987). Lexical retrieval: The tip-of-the-tongue phenomenon. *Applied Psycholinguistics* , **8**, 245-266.

Krauss, R. M. (1995). *Gesture, speech and lexical access*. Talk given at the conference "Gestures Compared Cross-Linguistically," under the auspices of the Linguistics Institute, July 7-10 in Albuquerque, NM:

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 389-450). San Diego, CA: Academic Press.

Krauss, R. M., Dushay, R. A., Chen, Y., & Rausher, F. (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology* , **31**, 533-552.

Krauss, R. M., & Fussell, S. R. (1996). Social psychological models of interpersonal communication. In E. T. Higgins & A. Kruglanski (Ed.), *Social psychology: A handbook of basic principles* (pp. 655-701). New York: Guilford.

Krauss, R.M., & Hadar, U. (in press). The role of speech-related arm/hand gestures in word retrieval. R. Campbell and L. Messing (Eds.), *Gesture, speech and sign*. Oxford: Oxford University Press.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology*, **61**, 743-754.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.

McClave, E. (1994). Gestural beats: The rhythm hypothesis. *Journal of Psycholinguistic Research* , **23**, 45-66.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, **92**, 350-371.

McNeill, D. (1987). *Psycholinguistics: A new approach*. New York: Harper & Row.

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

McNeill, D., Cassell, J., & McCollough, K.-E. (1994). Communicative effects of speech-mismatched gestures. *Language and Social Interaction* , **27**, 223-237.

Mead, G. H. (1934). *Mind, self and society*. Chicago: University of Chicago Press.

Meyer, A. S., & Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation?.   *Memory and Cognition* , **20**, 715-726.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **18**, 615-623.

Moscovici, S. (1967). Communication processes and the properties of language. In L. Berkowitz (Ed.), *Advances in experimental social psychology* . New York: Academic Press.

Rauscher, F. B., Krauss, R. M., & Chen, Y. (1996). Gesture, speech and lexical access: The role of lexical movements in speech production. *Psychological Science* , **7**, 226-231.

Ricci Bitti, P. E., & Poggi, I. A. (1991). Symbolic nonverbal behavior: Talking through gestures. In R. S. Feldman & B. Rimé (Ed.), *Fundamentals of nonverbal behavior* (pp. 433-457). New York: Cambridge University Press.

Rimé, B. (1982). The elimination of visible behaviour from social interactions: Effects on verbal, nonverbal and interpersonal behaviour. *European Journal of Social Psychology*, **12**, 113-129.

Schegloff, E. (1984). On some gestures' relation to speech. In J. M. Atkinson & J. Heritage (Ed.), *Structures of social action* . Cambridge: Cambridge University Press.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Werner, H., & Kaplan, B. (1963). *Symbol Formation*. New York: Wiley.