The Role of Speech-Related Arm/Hand Gestures in Word Retrieval

Robert M. Krauss

Columbia University

and

Uri Hadar

Tel Aviv University

INTRODUCTION

The traditional view that gestures play an important role in communication is so widespread and well-entrenched that comparatively little research has been done to assess the magnitude of their contribution, or to determine the kinds of information different types of gestures convey. Reviewing such evidence as exists, Kendon has concluded:

> 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).

However, other researchers, considering the same studies, have concluded that the available evidence is both inconclusive and equally consistent with the view that the gestural contribution to communication is, on the whole, negligible (Feyereisen and deLannoy 1991; Rimé and Schiaratura, 1991; Krauss et al 1995). Below we will examine in some detail the question of whether gestures communicate.

One reason that gestures are so often ascribed a communicative function may be that it is not obvious what other functions they might serve. Actually, over the past half-century several have been suggested. For example, noting that people often gesture when they are having difficulty retrieving elusive words from memory, Dittmann and Llewelyn (1969) have suggested that at least some gestures may be functional in dissipating the tension that accumulates during lexical search. Dittmann and Llewelyn assume that the failure to retrieve a sought-after word is frustrating, and that the tensions generated by such

frustration could interfere with the speaker's ability to produce coherent speech; hand movements provide a means for dissipating excess energy and frustration. Other investigators have remarked on the cooccurrence of gestures and hesitation pauses (Freedman and Hoffman 1967; Butterworth and Beattie 1978; Christenfeld et al. 1991), although they have not attributed the gestures to tension management. The tension reduction hypothesis has never been tested experimentally, but there is little doubt that gesturing and word retrieval failures co-occur (Ragsdale and Silva 1982; Hadar and Butterworth 1997).

The possibility that gesturing occurs during hesitation because it plays a direct role in the process of lexical retrieval has been suggested by a strikingly diverse group of scholars over the last 75 years (DeLaguna 1927; Dobrogaev 1929; Mead 1934; Werner and Kaplan 1963; Freedman and Hoffman 1967; Moscovici, 1967). Although the idea is not a new one, the details of the process by which gestures might affect lexical access are both grossly underspecified and underconstrained by the available data.

In this chapter, we will first examine, from a conceptual and empirical perspective, the assumption that the primary role of gestures is communicative; we will conclude that their contribution to communication is relatively small. We will then consider evidence bearing on the possibility that they play a facilitative role in lexical retrieval, and conclude that there is some support for this idea. Next, we will propose a cognitive architecture that is consistent with the available evidence and accounts both for gesture production and for the facilitative effects of gestures on lexical retrieval. Finally, we will discuss specifications of the general model and, with them, possible linkages between the speech and the gesture systems.

Gesture and Communication

Do gestures serve a communicative function? Any answer to this question must implicitly or explicitly assume a definition of communication, but, generally speaking, discussions of gestural communicativeness have avoided addressing some of the thorny conceptual issues that arise in formulating such a definition. Space considerations preclude us from considering them in detail (see Krauss and Fussell 1996 for a more extended discussion), but a few points will serve to outline our argument: People use both symbols and signs to convey information, but the two kinds of signals convey information in importantly different ways. The difference corresponds to the distinction Grice (1957) draws between 'natural' and 'non-natural' meanings. Natural meanings are comprehended by virtue of a causal connection between the sign and what it is understood to mean, while non-natural meanings are comprehended by virtue of an understanding of the conventions that govern symbol use.[1] Although there is hardly anything resembling a consensus on the details of the process, a fair amount is understood about the way linguistic communication is accomplished, and there seems to be general agreement on two points:

First, as regards language use, it is assumed that communication involves exchanges of *intended meanings*. In order for a linguistic message to be communicative, (a) the speaker must intend the message to create some particular effect (i.e., a belief) in the addressee; and (b) the speaker must intend that effect to result from the addressee's recognition of the intention. Levinson

---

[1]We are using the term 'sign' in the traditional semiotic sense—i.e., a display that is causally related to its significance or meaning. Another quite different sense of 'sign' (as in 'sign language') is reflected in the way the term is used in this book's title and elsewhere.

puts it succinctly: *'…communication is a complex kind of intention that is achieved or satisfied just by being recognized'* (Levinson 1983, p. 18).

Second, it is assumed that language use is a *joint activity* in which the parties collaborate to produce *shared* meanings. From this perspective, a conversation can be viewed as a series of discursively-related contributions in which speakers and hearers take pains to ensure that they have similar conceptions of the meaning of each message before they proceed to the next one.

Unlike symbols, signs do not need to be performed with the intention of creating a particular effect in order to have that effect, and they do not require the addressee to recognize an intention in order to convey information. Sperber and Wilson (1986) nicely illustrate the distinction with the example of a woman who has a sore throat and wishes to inform someone of this fact. She could accomplish this by saying 'I have a sore throat,' but the information could equally well be conveyed simply by saying anything, and allowing the listener to infer her condition from her hoarse voice. In the former case, in which the information is conveyed linguistically, the speaker's intention is critical, while in the second, her intention to convey the information is irrelevant.

One way of defining communication is as information that has been conveyed in accordance with the intentionality and joint action criteria, Thus defined, communication would include most instances of language use, the use of such symbolic gestures or emblems as the thumbs-up sign, and certain deictic gestures. It would exclude sign behaviors such as a hoarse voice or blushing, which, although they unquestionably convey information, do so in a different fashion. The question we address is: Would it include what Kendon (1994) calls

'the gestures that people produce when they talk,'—what McNeill (1992) calls 'representational gestures' and we have called 'lexical gestures'?

Kendon and others (see Schegloff 1984; de Ruiter in press) contend that it would. They argue that speakers partition the information that constitutes their communicative intentions, choosing to convey some of it verbally in the spoken message, some of it visibly via gesture, facial expression, etc., and some of it in both modalities. For example, Kendon (1980) describes a speaker saying '…with a big cake on it…' while making a series of circular motions of the forearm with the index finger extended pointing downward. Kendon would have the speaker *intending* to convey the idea that the cake was both large *and* round, and choosing to convey ROUND gesturally rather than verbally.

Although there is nothing implausible about Kendon's interpretation, it is unclear from his description of the episode that the speaker's behavior satisfies the intentionality and joint activity criteria. Since it was explicitly included in the utterance, it's reasonable to assume that the idea of the cake being large was part of the speaker's communicative intention, presumably because it was discursively relevant. Can we assume this is also the case about its being round? Such an assumption would need to be justified with other evidence (perhaps from elsewhere in the narrative), but the mere fact that the gesture occurred and is interpretable in context is insufficient to demonstrate that it was communicatively intended.

Something we saw in one of our own experiments suggests how misleading such observations can be. In that experiment subjects learned the definitions of arcane words, and were later videotaped as they tried to recall the definitions. One of the words was 'deasil', which means 'to move in a clockwise direction'. All 14 of the subjects who remembered the word's definition made a

rotary movement of the index finger as they defined it. It seems reasonable to regard the gesture as intended to aid the listener. However, for all but one speaker, the rotation was *clockwise relative to the speaker;* that is to say, from the addressee's perspective, the movement was *counter*clockwise, hence misleading.[2] It is instructive to compare the subjects' behavior in this situation to what speakers do when they formulate spatial descriptions verbally. According to Schober (1993, 1995) speakers formulating referring expressions about locations overwhelmingly formulate them from the perspective of the addressee. The fact that most of the speech-accompanying gestures in our experiment were formulated from the speaker's perspective raises questions about the glib assumption that they were intended in the same way the elements of an utterance are intended. Clearly, if they were so intended, they were defective from the addressee's perspective.

Overall, empirical evidence for the communicativeness of gestures is at best equivocal. Experimental findings indicate that, on average, lexical gestures convey relatively little information (Feyereisen et al. 1988; Krauss et al. 1991, experiments 1 and 2), and that the extent to which they influence the semantic interpretation of utterances is negligible (Krauss et al. 1991, experiment 5). Moreover, adding a speaker's gestures to his/her voice does not enhance listeners' performance on a referential communication task, in which it is possible to measure how well a message accomplishes its intended purpose (Krauss et al. 1995).

Three studies (Graham and Argyle 1975; Rogers 1978; Riseborough 1981) are often cited as providing empirical support for the hypothesis that hand

---

[2]We are grateful to Stephen Krieger and Lisa Son for sharing this observation with us, and to Lauren Michelle Walsh, who coded the gestures.

gestures serve a communicative function, at least in highly specific circumstances. All three purport to find small, but statistically reliable, performance increments on tests of information (e.g., reproduction of a figure from a description; answering questions about an object on the basis of a description) for listeners who could see a speaker's gestures, compared to those who could not, suggesting that the gestures enhanced the effectiveness of the communication. However, all three studies suffer from serious methodological problems, and we believe that little can be concluded from them.

Graham and Argyle (1975) had 6 speakers describe abstract line drawings to a small audience of listeners who tried to reproduce them. On half of the trials, speakers were prevented from gesturing. Listeners reproduced the figures more accurately from descriptions given when speakers were allowed to gesture. However, the design of the experiment does not control for the possibility that speakers who were allowed to gesture produced better *verbal* descriptions of the stimuli, which, in turn, enabled their audiences to reproduce the figures more accurately, quite apart from any information conveyed by the gestures. The Riseborough study found an effect of gesturing for only one of three stimuli, and it is not clear that the relevant contrast was statistically reliable. Rogers found that subjects who viewed-and-heard videotaped descriptions of novel actions scored better on multiple choice tests of information than did subjects who only heard the audio track. However, the result was found only when a relatively high level of noise (signal-to-noise ratios of -3 and -8 dB) was added to the audio track; with more favorable signal-to-noise ratios (+2 and +7 dB), seeing the speaker had no demonstrable effect on communication. Rogers also did not adequately control for the 'speech-reading effect'—the well

established finding that seeing a speaker's lips contributes to the intelligibility of speech (Sumby and Pollack 1954).

More recently, McNeill et al. (1994) have found that gestural information that differed from the information conveyed by the accompanying speech tended to be reflected in listeners' later retellings of the narrative, suggesting that speech and gesture information were combined in memory. This study seems to provide a clear demonstration that at least some gestures contribute to listeners' comprehension, but it is not without problems. In our experience, speech-gesture mismatches of the kind used by McNeill et al. are relatively rare in adult speech; in the study the mismatches were enacted simulations rather than naturally-occurring events. But even if we accept these results at face value, the communicative contribution of the overwhelming majority of the non-mismatched gestures that accompany spontaneous speech still remains to be established

The fact that speakers gesture more often when they can see their addressees than when they cannot (Cohen and Harrison 1972; Cohen, 1977; Rimé, 1982; Bavelas et al. 1992; Krauss et al. 1995) is sometimes cited as evidence that the gestures are communicatively intended (Cohen and Harrison 1972; Kendon, 1994). Certainly it is weak evidence at best, since speakers do gesture when their listeners cannot see them. Are the gestures speakers make when they are visually inaccessible different from the ones they make when they can be seen? Not much research has been directed to this question; the little we know of suggests that they are not. Krauss et al. (in preparation) coded the grammatical types of 12,425 gestural lexical affiliates (i.e., the word or words in the accompanying speech associated with the gesture) from an experiment in which subjects described stimuli to a partner who was either seated face-to-face

or in another room. The grammatical categories of gestural lexical affiliates for the two conditions did not differ, suggesting that speakers gestured at the same points in their narratives regardless of whether or not they could be seen by their partners.

De Ruiter (in press) 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 such gestures 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 (p. $$).

Although the idea that such gestures reflect overlearned habits is not implausible, the contention that they are both communicatively intended *and* largely ineffective runs counter to a modern understanding of how language (and other behaviors) function in communication. De Ruiter's view implicitly conceptualizes participants as 'autonomous information processors' (Brennan 1991). Such a view stands in sharp contrast with what Clark (1996) and his colleagues have called a 'collaborative' model of language use. In this view, communicative exchange is a joint accomplishment of the participants who work together to achieve some set of communicative goals. From the collaborative perspective, communication requires that speakers and hearers endeavor to ensure 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, violates what Clark and Wilkes-Gibbs (1986) have termed the 'Principle of Mutual Responsibility'.

We believe that symbolic (emblematic) gestures, deictic gestures, and the kind of gestural activity that Clark has termed 'demonstrations' are, as a rule, both communicatively intended *and* communicatively effective. The question we raise is whether there is adequate justification for assuming that all or most co-speech gestures are so intended. We believe there is not. No doubt lexical gestures occasionally are intentionally performed, and it seems likely that some do convey information regardless of whether or not they were so intended by the speaker. However, considering, on the one hand, the amount of gesturing that often accompanies speech, and, on the other hand, the paucity of the information gestures seem to convey, it seems to us reasonable to ask whether gestures might be serving some other function.

The question of the functions lexical gestures serve has important implications for models of gesture production. If gestures are communicatively intended in the way utterances are, it seems reasonable to suppose that they share some stages of the production process with speech. If they are not, gesture and speech could have different sources. If gestures facilitate lexical retrieval, the speech production system and the gesture production system must interact. If they do not, the two systems could function autonomously.

Gesture and Word Retrieval

If hand gestures do not serve a communicative function, why would speakers bother to make them? One possibility is that they play a role in speech production. Empirical support for this notion 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 appears to have lacked the necessary controls. As was common in that era, Dobrogaev's report fails to describe procedural details, and presents his results in impressionistic, qualitative terms. More recently, Graham and Heywood (1975) analyzed the speech of five speakers who were prevented from gesturing as they described abstract line drawings, and concluded that '… elimination of gesture had no particularly marked effects on speech performance' (p. 194). On the other hand, Rimé (1982) and Rauscher et al. (1996) found that restricting gesturing adversely affects speech fluency. The Rauscher et al. study is especially relevant to our hypothesis that gestures facilitate access to lexical memory, because the effects of preventing gesturing on speech were found to be similar to the effects of making word retrieval difficult by other means (e.g., requiring subjects to use rare or unusual words).

More evidence supporting the association of gesture with word retrieval difficulties comes from studies with brain-damaged subjects. It has been known for some time that adult with Broca's aphasia produce more gestures per unit of speech than normal controls (Goldblum 1978), but there have been claims that these gestures were often disrupted in a way that paralleled the speaker's language disorder (McNeill 1992). In a single-case study, Butterworth et al. (1981) showed that their aphasic patient tended to gesture prior to a word retrieval failure (either a hesitation or an erroneous production). More recently, Hadar et al. (1998b) found that aphasics whose speech problems primarily concerned word retrieval tended to gesture more than both normal controls and other aphasics whose problems were primarily conceptual. About 70% of the gestures of the aphasics with word retrieval difficulties appeared adjacent to a hesitation

or erroneous production. At the same time, and contrary to the hypothesis of a gestural deficit, the composition and form of their gestures were normal.

In a very different kind of study, Hanlon et al. (1990) showed that aphasic patients' word retrieval in a picture naming task could be improved by training them to perform gestures just prior to their naming attempt.

LEXICAL, ICONIC AND METAPHORIC GESTURES

In the two sections that follow, we present a model of speech and gesture production in which the latter supports the former. Our model does not attempt to account for all speech related gestures, only the subset that we previously have called 'lexical' (Hadar 1989; Krauss et al. 1995). Lexical gestures are relatively long, broad and complex arm-hand movements that often have shapes or dynamics related to the content of the accompanying speech (Hadar et al. 1998a; 1998b). Other co-speech gestures that typically are short, simple and repetitive, called 'beats' by McNeill (1992) and 'motor gestures' by Krauss et al. (in press) do not seem to be involved in lexical search, and probably link to speech through other systems (Hadar 1989; Butterworth and Hadar 1989). Discussions of terminological and descriptive aspects of gesture can be found in Rimé and Schiaratura (1992) and Krauss et al. (in press).   Although the nature of gestural taxonomies is not a settled matter, we do not plan to enter this discussion here, but instead to address some issues that specifically concern lexical gestures.

Investigators often partition the movements we are calling lexical gestures into subcategories, although there is little consensus as to what those subcategories should include and exclude. 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 related conceptually to the

semantic content of the speech it accompanies (Efron 1972). However, not all
lexical gestures are iconic. The proportion of lexical movements that are iconic is
difficult to determine, and probably depends greatly on the conceptual content
of the speech and the flexibility of the standards used to determine iconicity. But
however iconicity is determined, many speech-accompanying movements seem
to have little obvious formal relationship to the conceptual content of the
accompanying speech. What can be said of these gestures? McNeill (1985, 1987,
1992) deals with the problem by drawing a distinction between gestures that are
iconic and gestures that are *metaphoric*.

> 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 the iconic-metaphoric distinction. Our own observations lead us to conclude
that iconicity is more a matter of degree rather than of kind. While the forms 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 finding any relationship at all requires a good deal of
imagination on the part of the observer. In our view, it makes more sense to
think of gestures as being more or less iconic rather than either iconic or
metaphoric. Moreover, although it may be possible to judge the iconicity of
gestures fairly reliably, the 'meanings' of such gestures are highly uncertain,
even when viewing them along with the accompanying speech. In the absence
of speech, the meanings of iconic gestures are indeterminate (Feyereisen et al.
1988; Krauss et al. 1991).

We also find problems with the iconic/metaphoric distinction at the conceptual level. The argument that metaphoric gestures are produced in the same way as linguistic metaphors does not help us understand how they are generated, since our understanding of the processes by which linguistic metaphors are produced and comprehended is incomplete at best (cf. Glucksberg 1991; Glucksberg, in press). So identifying such gestures as visual metaphors may be little more than a way of saying that the nature of 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. 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. However we do not believe that identifying gestures as metaphoric really accomplishes that goal, and instead will opt for a model formulated at the level of features as a more satisfactory way of accounting for gestural form.

SPEECH PRODUCTION

In our model, lexical gestures reflect the use that the speech production system makes of the gesture production system for word retrieval purposes. Before describing the gesture production system, we will review our understanding of the process by which speech is generated. Of course, the nature of speech production is not uncontroversial, and a variety of production models have been proposed. Although these models differ in significant ways, for our purposes their differences are less important than their similarities. Most models distinguish three successive stages of processing, which Levelt (1989) refers to as 'conceptualizing', 'formulating', and 'articulating' (Garrett 1984; Dell 1986;

Butterworth 1989). The process is illustrated schematically in Fig. 1 below, which is based on Levelt (1989). Let us examine the model's stages more closely.

Insert Fig. 1 about here

Conceptualizing involves, among other things, drawing upon declarative and procedural knowledge to construct a communicative intention. We believe that initially many memorial representations, in a variety of representational formats, are activated by contextual triggering and structural (cognitive/affective) biases. Then a more active process of 'focusing' reduces the number of activated representations, and forms connections between those remaining (Sperber and Wilson 1986; Levelt 1989). The output of the conceptualizing stage—what Levelt refers to as a *preverbal message*— can be thought of as a propositional structure containing a set of pragmatic and semantic specifications (Bierwisch and Schrueder 1992).

This preverbal message constitutes the input to the formulating stage, where the message is transformed in two ways. First, a grammatical encoder maps the to-be-lexicalized concept onto an abstract syntactic structure. At the same time the main lexical entries of the message are selected in the form of a lemma, i.e., an abstract symbol representing the selected word as a semantic-syntactic entity. Lemmas are selected whose semantic features match a subset of the semantic features of the preverbal message. A *surface structure* is formed by joining the abstract syntactic structure with the lemmas. Then, by accessing word forms stored in lexical memory and constructing an appropriate plan for the utterance's prosody, the phonological encoder transforms this surface structure into a *phonetic plan* (essentially a set of instructions to the articulatory system) that serves as the input to the articulatory stage. The output of the articulatory

stage is overt speech, which is monitored auditorily and used as a source of corrective feedback.

<div align="center">**GESTURE PRODUCTION**</div>

Our model makes several assumptions about mental representation and memory. They are:

(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.

(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 visual imagery, and vice versa).

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

In our view, gestures originate in the process that precedes conceptualization and construction of the preverbal message. That is to say, we believe their origin *precedes* the formulation of the speaker's communicative intention. Consider Kendon's (1980) aforementioned example of the speaker saying '…with a big cake on it…' while making a series of circular motions of the forearm with index finger pointing downward. We assume that the articulated

word 'cake' derives from a conceptual representation of a *particular* cake in the speaker's long-term memory. The preverbal message outputted by the conceptualizer (and which the grammatical encoder transforms into a linguistic representation) typically incorporates only a subset of the memorial representation's features. From the information Kendon gives us, it seems reasonable to assume that the particular cake the speaker referred to in the example was large and round. Of course, 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.

Apropos our earlier discussion of communication, a central theoretical question 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. From his discussion, it is clear that Kendon (1980) assumes that it was. Our assumption is that it was not. Below we will consider some of the implications of this assumption.

We follow Levelt in assuming that inputs from the conceptualizing stage to the formulating stage of the speech processor must be in propositional form. However, the knowledge that is accessed from memory and becomes incorporated into the communicative intention may be multiply encoded in propositional and nonpropositional formats, or even encoded exclusively in nonpropositional formats. In order for nonpropositionally-encoded knowledge to be reflected in speech, it must be 'translated' into propositional form.

Insert Fig. 2 about here

How do these nonpropositionally-represented features come to be reflected gesturally? The model is illustrated in Fig. 2.   Like the speech production system, the gesture production system has three stages. In the first stage, a Spatial/Dynamic Feature Selector takes representations that have been activated in spatial or visual working memory, selects out elementary spatial and dynamic features, and renders them as a set of *spatial/dynamic specifications.* These specifications are essentially abstract features of movements—velocity, direction, contour, and the like. This set of abstract features serves as input to the Motor Planner, which translates them into a *motor program* —a set of instructions for executing the lexical gesture.   These instructions are then executed by the Motor System in the form of a gestural movement. Thus thinking of a particular cake that happened to have been round would activate the spatial feature ROUND, which is translated into a circling motion by the Motor Planner. A single gesture may reflect one or more spatial/dynamic feature. For example, the circling gesture is, at least in principle, capable of representing both shape and circumference. However, it may be difficult or impossible to simultaneously incorporate some combinations of features into a single gesture; for example, if both ROUND and THICK were activated features of the remembered cake, the gesture might reflect one or the other, but probably not both.

Once they have been produced, gestures are monitored kinesthetically in much the same way that the output of the speech production system is monitored auditorily. We will leave the destination of the kinesthetic monitor's output to be specified later.

It's not difficult to see in general terms how such a system could generate gestures from conceptual content that is concrete and spatial (e.g., *book, arc)* or

involves movement (e.g., *twist, lift)*, but gesturing also accompanies talk whose conceptual content is abstract and static. What is the origin of these gestures? In the first place, it's important to know that gesturing is strongly associated with speech having spatial content. Speakers describing animated action cartoons were nearly five times more likely to gesture during clauses containing spatial prepositions than they were elsewhere (Rauscher et al. 1996). Zhang (reported in Krauss et al. in preparation) measured the proportion of time speakers spent gesturing as they defined a variety of common words. Ratings of word-concepts' 'spatiality' accounted for more than 50% of the variability in the amount of time speakers spent gesturing while defining them.

Although the evidence suggests that gesturing derives primarily from spatial (and, we believe, dynamic) features of concepts, some gesturing accompanies speech that has no apparent spatial content. In Zhang's study, subjects gestured more than twice as much when defining 'adjacent' and 'cube' than they did defining 'thought' and 'devotion;' nevertheless, gestures accompanied about 17% of the time spent defining the latter two words. Since neither term has explicit spatial or dynamic content, the fact that their definitions were accompanied by gesturing seems inconsistent with our model. In understanding how the model accounts for such gestures, it's important to bear in mind that it assumes gestures to be products of memorial representations rather then of communicative intentions. What we believe to be involved are inter-connected systems containing concepts, lemmas, long-term visuo-spatial representations and motor schemata so arranged that the activation of any concept can result in the activation of a loosely connected motor schema. The outline of such a system has only recently begun to emerge, describing such connections between visuo-spatial and motor schema mediated by the

increasingly popular notion of 'embodiment' (Ballard et al. 1997), lemma and visuo-spatial representations (Bierwisch 1996) and the figurative context of lexical semantics (Gibbs 1997). These ideas, of course, require much more detailed development to connect them to specific gestural phenomena.

Gestural Facilitation of Speech

Our contention is that lexical gestures facilitate lexical retrieval. The process by which this is accomplished is illustrated in Fig. 3, in which the gesture production system and the speech production system are connected. We will first describe the general form of the model and then examine more closely some specific issues it raises. In Fig. 3, the spatio-dynamic information the gesture encodes is fed via the kinesic monitor to the formulator, where it facilitates lexical retrieval. Facilitation is achieved through cross-modal priming, in which gesturally-represented features of the concept in memory participate in lexical retrieval. Of course, it is possible to locate the site of gestural input more precisely (e.g., the grammatical encoder or the phonological encoder).

Fig. 3 shows the gesture production system affecting the speech production system. Not shown is a path by which the speech production system provides input to the gesture production system. Some such a link may be necessary to tell the gesture production system when to terminate a gesture.

Insert Fig. 3 about here

FURTHER SPECIFICATION OF THE MODEL

The model illustrated in Fig. 3 describes a set of structures and a general process flow. For some aspects of the model there is considerable empirical support, while other aspects are quite speculative. Models of the gesture production system and the process by which it and the speech production system interact must of necessity be speculative

because there is relatively little reliable data available to constrain them. Not surprisingly, researchers have made quite different assumptions about the processes involved. In this section we will consider alternative ways of specifying the process of gesture production and gesture-speech interaction, and examine the relevant data.

Gesture Origins

In the gesture literature, the term 'origin' has been used to refer to two different, but related, aspects of the gesture production system. One usage refers to the source of input to the gesture production system; this is the usage we will favor. The other usage of 'origin' refers to the process that triggers or activates the gesture. We will refer to this process as gesture *initiation,* and discuss it, along with gesture termination, in the next section

In Fig. 3, the origin of gesture is shown to be the spatial-dynamic representations in working memory that activate the feature-selection component of the gesture production system (see Fig. 2); in our view, gestures always involve processes that precede the formulation of a communicative intention. Others have made different assumptions. For example, in a model quite similar to ours in other respects, de Ruiter (in press) designates the conceptualizer as the origin of the gesture production system's input. This is consistent with his assumption that lexical gestures are communicatively intended. We have already discussed what we see as problematic with that assumption. Specifying the conceptualizer as the origin of gestures raises the additional problem of how such gestures could aid in lexical access. If the conceptualizer's input to the Gesture Planner in de Ruiter's model contains the same information as the input to the formulator, it would be difficult to see how gestural information could facilitate lexical retrieval, or why preventing gesturing should make lexical retrieval more difficult.

The idea that lexical gestures have an early origin is consistent with the well-established finding that lexical gestures precede their lexical affiliates (Butterworth and Beattie 1978; Schegloff 1984; Morrel-Samuels and Krauss 1992). Morrel-Samuels and Krauss, for example, examined 60 carefully selected lexical gestures and found the gesture-speech 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 been initiated. The product-moment correlation between gestural duration and asynchrony is +0.71.

The idea that lexical gestures originate in short-term memory contrasts with the position taken by McNeill (1992), who argued for multiple links between the speech and the gesture systems, consistent with a connectionist cognitive architecture. We find such an architecture contributes little to explicating the relationship of gesture and speech because it is insufficiently constrained and, therefore, does not produce sufficiently specific predictions.

Employing more modular architectures, Butterworth and Beattie (1978), Butterworth and Hadar (1989) and Hadar and Butterworth (1997) have argued in favor of two different gestural origins, one in short term memory and the other later in the speech production process. Their argument hinges upon a distinction between iconic gesture and gestures that are indefinite, in the sense that they cannot be affiliated with a specific lexical item. They hypothesize that the origin of indefinite gestures is short term memory, while iconic gestures are directly activated by lexical processes. However, the available data do not support this idea. By the logic of their argument, iconic gestures should tend to

start during hesitation pauses, while indefinite gestures should not. By the same argument, the gesture-speech asynchrony (i.e., the interval between the initiation of the gesture and the articulation of the lexical affiliate) should be smaller for iconic gestures that start during hesitation pauses than it is for iconic gestures that do not start during hesitation pauses. In neither case is the available evidence supportive (Hadar et al. 1998a, 1998b). The Butterworth and Hadar argument brings into focus the important issues of gesture initiation and termination, which we address in the next section.

The Initiation and Termination of Gestures

What causes a speaker to gesture? Krauss et al. (1995; in press) assume that the early conceptual processes that produce the input to the conceptualizer routinely implicate non-propositional representations. Some of these derive from spatial or dynamic properties of the processed concepts, and this particular subset of non-propositional representations is linked with the gesture production system: its activation activates the spatial/dynamic features selector. On this account, whenever a spatial representation is activated, a gesture is triggered. Krauss et al. (1995, in press) present two kinds of evidence in support of their model: First, speech with spatial content is considerably more likely to be accompanied by gestures than speech with other kinds of content, although this was not found in a study by Hadar and Krauss (in press). Second, immobilizing gesture selectively impairs speech with spatial content. Although the latter finding has not yet been independently replicated, the account as a whole seems plausible and consistent with the available data. However, it certainly is possible that the subset of representations linked up with gesture production is not specifically spatial, but visual, as McNeill (1992) suggests, or visuo-spatial, as Hadar et al. (1998a) suggest.

Krauss et al. (1995, in press) explain the tendency of gestures to be associated with hesitations by assuming that lexical selection switches off the gesture production system. On this account, if the set of features that activated the gesture is realized in lexical selection, the gesture production process is aborted. Consequently, many gestures are activated but not executed; difficulties encountered in lexical selection may simply allows sufficient time for the gesture to reach execution.

Butterworth and Hadar have proposed a more complex dual mechanism for gesture initiation in which some gestures are activated directly from short-term memory, while others are initiated by failures of lexical retrieval (Butterworth and Hadar 1989). Hadar and Butterworth 1997). They contend that such failures often initiate a re-run of lexical selection, and that during such re-runs, the formulator attempts to gather more cues for lexical selection by activating non-propositional representations related to the sought-for lexical entry. These non-propositional representations, in turn, activate a gesture. We accept this possibility, but stress that the available evidence suggests that in these cases the loop of the re-run must be 'deep' enough to activate early representations, and it is these early representations which activate the gesture. We also note that the lexical loop may, at best, apply only to those gestures that are associated with hesitation. On our count, these amount to about 70% of gestures in aphasic patients with primarily lexical retrieval problems, but in healthy subjects they amount to about 30% of lexical gestures (Hadar et al. 1998b). A different kind of mechanism must be hypothesized to account for gestures that are not associated with hesitation.

The Input from Gesture to Speech

In order to affect lexical retrieval, gesture-related information must enter the speech system. There are a number of possible entry points. Fig. 3 shows the output of the kinesic monitor feeding into the formulator. This is the simplest inference from the hypothesis of lexical facilitation: gesture-related information acts as input to lexical selection either in the form of additional cues (Hadar and Butterworth 1997) or in the form of cross-modal priming (Krauss et al. in press).

Hadar and Butterworth (1997) have suggested that gestural information might be input to the conceptualizer. On this account, gesture-related information would contribute to the construction of the speaker's communicative intention and affect lexical retrieval only indirectly. The available evidence, although far from definitive, is not supportive of this view. Rauscher et al. (1996) found that preventing speakers from gesturing increased the proportion of nonjuncture pauses in their speech. Unlike juncture pauses, which can result from a variety of causes (including conceptualizing), nonjuncture pauses seem mainly to reflect problems in lexical retrieval (see Krauss et al. 1996 for a discussion). Hence, the fact that preventing gestures increases the relative frequency of hesitations is consistent with the proposition that the gestures are involved with lexical retrieval. However, the subjects' task in this study (describing the plots of animated action cartoons) may have made minimal conceptual demands. Rauscher et al. found that making lexical retrieval more difficult increased the impact of preventing gesturing; similarly, varying the conceptual complexity of the speaker's task might reveal that what we are calling lexical gestures also function at the conceptual level. Research by Goldin-Meadow et al. (1993) and observations by McNeill (1992) seem to indicate that

gesturing is associated with conceptual activity, but their specific role is far from clear.

Within the formulator, gesture-related information could serve as an input to either the grammatical encoder or the phonological encoder, or to both. Lexical retrieval proceeds in two separate stages--lemma selection and word-form selection--and problems at either stage could result in the kinds of dysfluencies observed by Rauscher et al. It is reasonable to assume that gesture-related information enters the speech system at the point at which facilitation occurs, so examining facilitation effects may help us decide this issue. Unfortunately, the empirical evidence is contradictory.

Some findings support the idea of semantic facilitation, suggesting entry via the lemma system. For example, aphasic patients whose have problems naming objects tend to produce more lexical gestures if their difficulties involve retrieval of the lemma rather than retrieval of the word form (Hadar et al. 1998b). Also, native speakers of Hebrew with good facility in English produced more lexical gestures accompanying self-generated descriptions in both English and Hebrew than they did while translating from Hebrew to English or vice versa (Teitelman 1997). If one assumes that hesitation in a second language derives from problems in accessing word forms (Kroll and Stewart 1994), then the dearth of gestures contra-indicates phonological retrieval as the beneficiary of gesture. However, the same study found more lexical gestures accompanying self-generated descriptions in English than in Hebrew, which is consistent with the word-form hypothesis. To further complicate matters, Dushay (1991) found that subjects in a referential communication task gestured less often when describing abstract figures and synthesized sounds in L_2 than in L_1. Dushay's subjects, students taking second year Spanish, were considerably less fluent than

Teitelman's, and unpublished data collected by Melissa Lau suggests that frequency of gesturing in L_2 may be a function of the speaker's fluency; the more fluent her English-Cantonese bilinguals were in Cantonese, the more frequently they gestured while speaking it. Like Teitelman, Lau's subjects gestured more when speaking spontaneously than they did when translating, either from Cantonese to English or from English to Cantonese. However, unlike Teitelman's subjects, hers gestured more frequently overall when speaking English (L_1) than Cantonese (L_2). At this point, it is not clear what can be concluded about gestural input to the speech production system from the gestural behavior of bilinguals.

Finally, since there is no systematic relationship between the semantic features that are part of the lemma and the phonological features that make up the word form, if lexical facilitation is achieved through priming, as Krauss et al. (1995) suggest, it seems likely that facilitation occurs at the level of lemma selection.

On the other hand, findings from studies using the 'tip of the tongue' (TOT) paradigm are consistent with the view that gestures facilitate retrieval at the word form rather than the lemma level. It is well accepted that TOT retrieval failures in normal subjects tend to be phonological rather than semantic (Brown and McNeill 1966; Jones and Langford 1987; Kohn et al. 1987; Jones 1989; Brown 1991; Meyer and Bock 1992), and there is some evidence that preventing gesturing increases retrieval failures in the TOT situation (Frick-Horbury and Guttentag, in press). In the same vein, Broca's aphasics tend to produce very high proportions of lexical gestures (Cicone et al. 1979; McNeill 1992), and their ability to name also seems to benefit from intentionally performing a gesture prior to naming (Hanlon et al. 1990). However, there is some disagreement on

the nature of naming problems in Broca's aphasia. While some researchers (e.g., Brown 1982) consider these patients' difficulties to be similar to TOTs, others (e.g., Williams and Canter 1987) argue that their retrieval failures primarily involve verbs, and therefore originate in lemma selection.

In sum, there is considerable evidence to indicate that the gesture production system's output affects the formulator in the speech production system. Certainly it is possible that gesture-related information also affects the conceptualizer, but the evidence for this is largely anecdotal. Within the formulator, gesture-related information could influence either grammatical or the phonological encoding, and there is indirect evidence consistent with both possibilities. Although it seems reasonably clear that information from the gesture-production system can affect speech production, we are not yet in a position to specify the locus or loci of these effects.

The Output from Gesture to Speech

At what point does the gesture-related information leave the gesture production system and enter the speech system? As before, there are several possibilities, all underconstrained by the available data. Butterworth and Hadar (1989; Hadar and Butterworth 1997) have suggested that the information leaves the gesture production system from its origin, that is, prior to actually generating the gesture. In their account, gesturing might be considered an artifact of the activation of the direct origin, the real purpose of which is to re-run the word selection process. They offer no data to support their claim, but rather adduce it as an inference from considerations of conceptual processing. In their view, gesture acts only to maintain activation of the non-propositional representation long enough for the word selection process to develop. The actual production of gesture, then, can contribute to facilitation, but does so indirectly.

Krauss et al. (1995, in press) hold a contrary view—that the gesture must actually be performed for facilitation to occur. In their model, information contained in the gesture, consisting of kinesthetic and proprioceptive representations, is extracted by the kinesic monitor. It is this information, inputted to the formulator, that facilitates retrieval. They conclude this largely from the finding that limitation of gesturing impairs word retrieval in normal subjects. Note that some impairment can be inferred from the Hadar and Butterworth model as well, but the two accounts differ as to how readily subjects should be able to compensate for gestural immobilization. According to Hadar and Butterworth (1997), but not Krauss et al. (1995), compensation should be fairly easy.

CONCLUDING COMMENT

We have described the general outlines of a model for the production of lexical gestures. We also have examined in some detail a number of alternative ways of formulating the model, and considered their strengths and vulnerabilities. One conclusion seems reasonably clear. As things currently stand, there is so little reliable data to constrain theory on gesture production that any processing model must be both tentative and highly speculative. Nevertheless, we do not believe that model building in such circumstances is a waste of time. Models provide a convenient way of systematizing available data. They also compel theorists to make explicit the assumptions that underlie their formulations, thus making it easier to assess in what ways, and to what extent, apparently different theories actually differ. Finally, and arguably most importantly, models guide investigators' efforts, and lead them to collect data that will confirm or disconfirm one or another model.

Our account has relied primarily on data from experiments. Experimentation is, of course, a powerful method for generating certain kinds of data, but it also has serious limitations, and Kendon (1994) has remarked on discrepancies between the conclusions reached by investigators who rely on experimental findings and those whose data derive mainly from natural observation. Observational studies have enhanced our understanding of what gestures accomplish, and the recent addition of neuropsychological observations should provide further insight into the gesture production system. The conclusions of careful and seasoned observers certainly deserve to be taken seriously. At the same time, we are uncomfortable with some investigators' excessive reliance upon observers' impressions of gestural form or meaning, especially when the observers are aware of the contents of the accompanying speech. Without the proper controls, such impressions provide a weak foundation for theory and, we believe, are more usefully thought of as a source of hypotheses than as data in their own right.

We have few illusions that we have considered every possible implementation of the model, or that all of the assumptions we have made will ultimately prove to have been justified. Indeed, our own ideas about the process by which lexical gestures are generated have changed considerably over the last several years, and we would be surprised if they did not continue to change as data accumulate. Although much of the data currently available is equivocal, and much more remains to be collected, we believe that the process in which models are produced and data (both experimental and observational) are collected to confirm or disconfirm them will ultimately result in a genuine understanding of how gestures are produced and how they are related to speech.
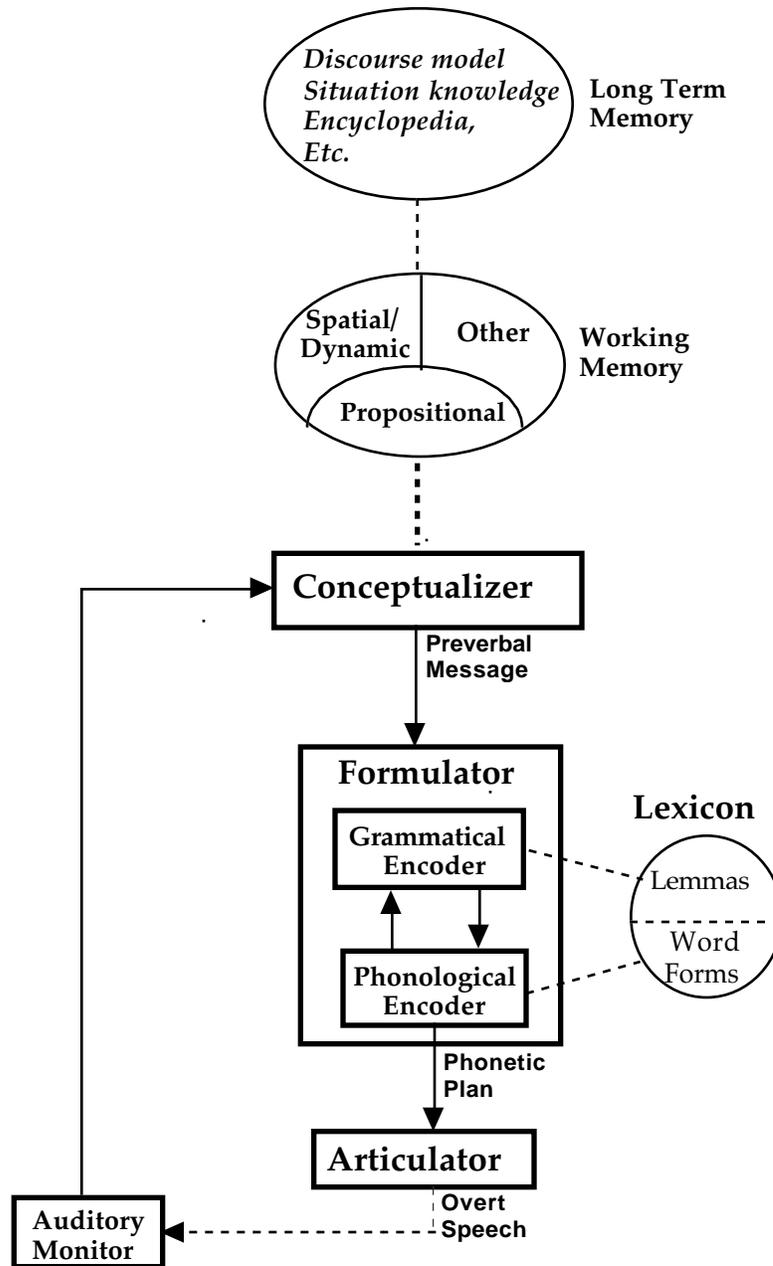
Fig. 1. A cognitive architecture for the speech production process
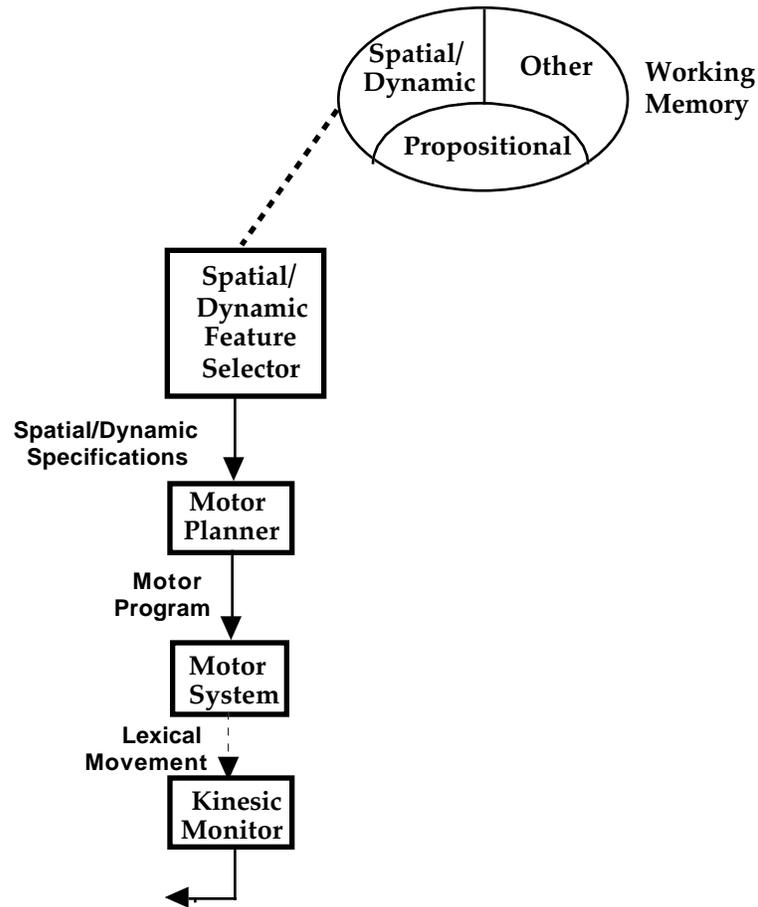(based on Levelt 1989).

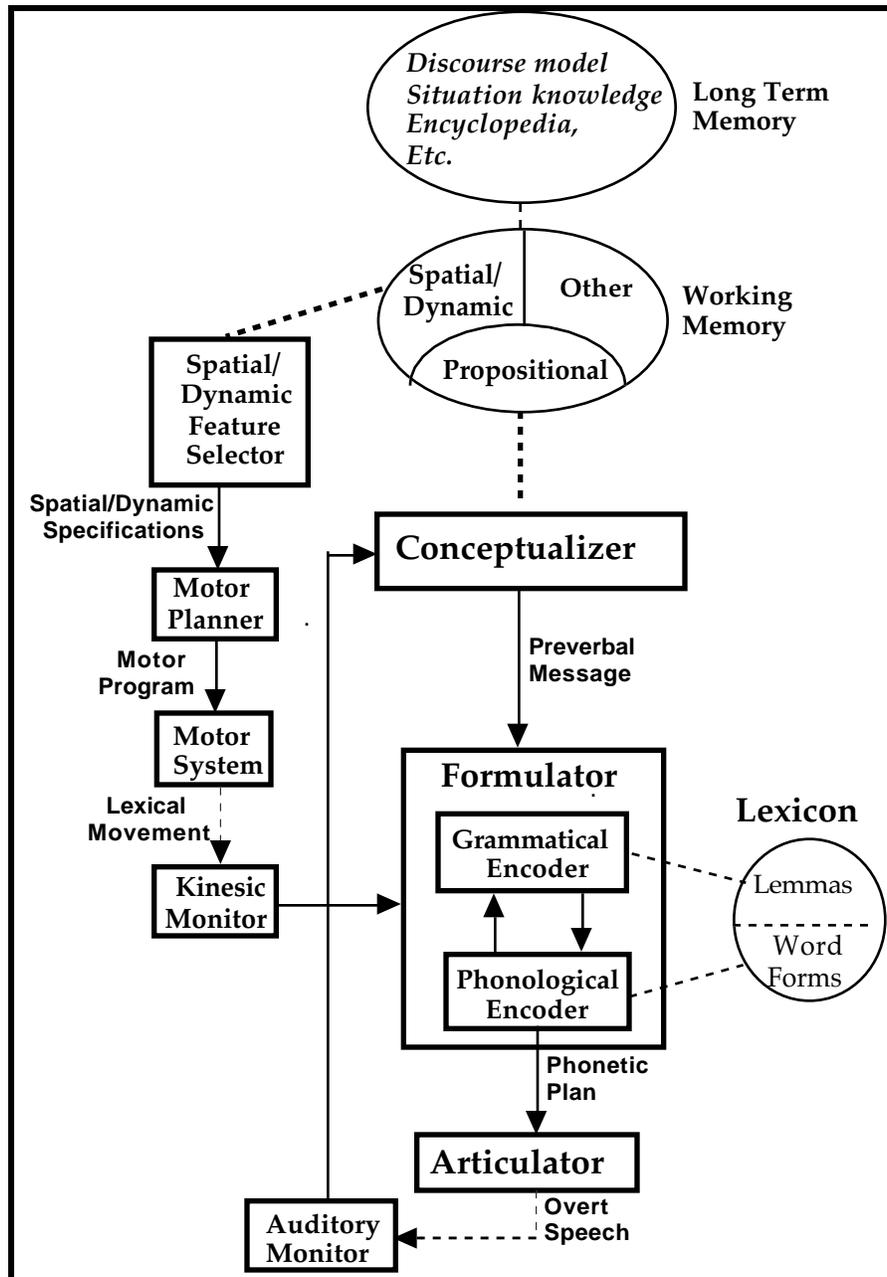Fig.2. A cognitive architecture for the gesture production system.

Fig. 3. Interaction of the speech and gesture production systems.

REFERENCES

Ballard, D.H., Hayhoe, M.M., Pook, P.K. and Rao, R.P. (1997) Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, **20**, 723-67.

Bavelas, J. B., Chovil, N., Lawrie, D. A., and Wade, A. (1992). Interactive gestures. *Discourse Processes,* **15**, 469-89.

Bierwisch, M. (1996) How much space gets into language? In *Language and Space,* (ed. P. Bloom, M.A. Peterson, L. Nadel, and M.F. Garrett, Cambridge, MIT Press.

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

Brennan, S.E. (1991). *Seeking and providing evidence for mutual understanding.* Unpublished PhD. dissertation, Stanford University.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin*, **109**, 204-23.

Brown, J.W. (1982). Hierarchy and evolution in neurolinguistics. In *Neural models of language processes*. (ed. M.A. Arbib, D. Caplan and J.C. Marshall), New York, Academic Press.

Brown, R., and McNeill, D. (1966). The 'tip of the tongue' phenomenon. *Journal of Verbal Learning and Verbal Behavior*, **4**, 325-37.

Butterworth, B.L. (1989). Lexical access in speech production. In *Lexical representation and process,* (ed. W. Marslen-Wilson), pp. 108-35. Cambridg, MIT Press.

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

Butterworth, B., and Hadar, U. (1989). Gesture, speech and computational stages: A reply to McNeill. *Psychological Review*, **96**, 168-74.

Butterworth, B., Swallow, J. and Grimston, M. (1981). Gestures and lexical processes in jargon aphasia. In *Jargonaphasia,*(ed. J.W. Brown), pp. 113-24, New York, Academic Press.

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

Christenfeld, N., Schachter, S., and Bilous, F. (1991). Filled pauses and gestures: It's not coincidence. *Journal of Psycholinguistic Research,* **20**, 1-10.

Cicone, M., Wapner, W., Foldi, N., Zurif, E. and Gardner, H. (1979) The relation between gesture and language in aphasic communication. *Brain and Language,* **8**, 324-49.

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

Clark, H. H., and Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition,* **22**, 1-39.

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

Cohen, A. A., and Harrison, R. P. (1972). Intentionality in the use of hand illustrators in face-to-face communication situations. *Journal of Personality and Social Psychology*, **28**, 276-79.

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

Dell, G.S. (1986) A spreading activation theory of retrieval in lamguage production. *Psychological Review*, **93**, 283-321.

de Ruiter, J. P. (in press). The production of gesture and speech. In *Language and Gesture: Window into thought and action* (ed. D. McNeill), New York, Cambridge University Press.

Dittmann, A. T., and Llewelyn, L. G. (1969). Body movement and speech rhythm in social conversation. *Journal of Personality and Social Psychology,* **23**, 283-92.

Dobrogaev, S. M. (1929). Ucnenie o reflekse v problemakh iazykovedeniia [Observations on reflexes and issues in language study]. *Iazykovedenie i Materializm,*, 105-73.

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

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

Ekman, P., and Friesen, W. V. (1972). Hand movements. *Journal of Communication*, **22**, 353-74.

Feyereisen, P., and deLannoy, J.-D. (1991). *Gesture and speech: Psychological investigations*. Cambridge, Cambridge University Press.

Feyereisen, P., Van de Wiele, M., and 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 *Studies in dyadic communication* (ed. A. W. Siegman and B. Pope), pp. 153-75. New York, Pergamon.

Freedman, N., and Hoffman, S. (1967). Kinetic behavior in altered clinical states: Approach to objective analysis of motor behavior during clinical interviews. *Perceptual and Motor Skills*, **24**, 527-39.

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

Garrett, M.F. (1984) The organization of processing structure for language production: Applications to aphasic speech. In *Biological Perspectives on Language*, (ed. D. Caplan, A.R. Lecours and A. Smith), pp. 172-93, Cambridge, MIT Press.

Gibbs, R.W. (1997) How language reflects the embodied nature of creative cognition. In *Creative thought: An Investigation of conceptual structures and processes* (ed. T.B. Ward, S.M. Smith, and J. Vaid), Washington, DC, APA Publications.

Glucksberg, S. (1991). Beyond literal meanings: The psychology of allusion. *Psychological Science*, **2**, 146-52.

Glucksberg, S. (in press). How metaphors work. In *Metaphor and thought* (2nd edn.) (ed. A. Ortony), Cambridge, Cambridge University Press.

Goldblum, M. C. (1978). Les trouble des gestes d'accompagnement du langage au cours des lesions corticales uniliaterales. In *Du Controle Moteur a l'Organisation du Geste*, (ed. H. Hecaen and M. Jeannerod), pp. 383-95, Paris, Masson.

Graham, J. A., and Argyle, M. (1975). A cross-cultural study of the communication of extra-verbal meaning by gestures. *International Journal of Psychology*, **10**, 57-67.

Graham, J. A., and Heywood, S. (1975). The effects of elimination of hand gestures and of verbal codability on speech performance. *European Journal of Social Psychology*, **5**, 185-95.

Grice, H. P. (1957). Meaning. *Philosophical Review,* **64**, 377-88.

Hadar, U. (1989). Two types of gesture and their role in speech production. *Journal of Language and Social Psychology*, **8**, 221-28.

Hadar, U., and Butterworth, B. (1997). Iconic gestures, imagery and word retrieval in speech. *Semiotica*, **115**, 147-72.

Hadar, U., Burstein, A., Krauss, R. M., and Soroker, N. (1998a). Ideational gestures and speech: A neurolinguistic investigation. *Language and Cognitive Processes,* **13**, 59-76.

Hadar, U., Wenkert-Olenik, D., Krauss, R.M. and Soroker, N. (1998b). Gesture and the processing of speech: Neuropsychological evidence. *Brain and Language,* **62**, 107-26.

Hadar. U., and Krauss, R.M. (in press) Iconic gestures: The grammatical categories of lexical affiliates. *Journal of Neurolinguistics*.

Hanlon, R.E., Brown, J.W. and Gerstman, L.J. (1990) Enhancement of naming in nonfluent aphasia through gesture. *Brain and Language*, **38**, 298-314.

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., and Langford, S. (1987). Phonological blocking in the tip of the tongue state. *Cognition,* **26**, 115-22.

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

Kendon, A. (1983). Gesture and speech: How they interact. In J. M. Weimann and 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., et al. (1987). Lexical retrieval: The tip-of-the-tongue phenomenon. *Applied Psycholinguistics,* **8**, 245-66.

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

Krauss, R. M., Morrel-Samuels, P., and Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology*, **61**, 743-54.

Krauss, R. M., Dushay, R. A., Chen, Y., and Rauscher, F. (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology*, **31**, 533-52.

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

Krauss, R.M., Chen, Y., and Gottesman, R.F. (in press) Lexical gestures and lexical access: A process model. In In *Language and Gesture: Window into thought and action* (ed. D. McNeill), New York, Cambridge University Press.

Krauss, R. M., Gottesman, R. F., Zhang, F. F., and Y., C. (in preparation). What are speakers saying when they gesture? Gramatical and conceptual properties of gestural lexical affiliates.

Kroll, J.F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connection between bilingual memory representation.*Journal of Memory and Language*, **33**, 149-74

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

Levinson, S. C. (1983). *Pragmatics*. Cambridge, Cambridge University 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**, 350-71.

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

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

McNeill, D., Cassell, J., and McCollough, K.-E. (1994). Communicative effects of speech-mismatched gestures. *Language and Social Interaction,* **27**, 223-37.

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

Meyer, A. S., and Bock, K. (1992). The tip-of-the-tongue phenomenon: Blocking or partial activation?. *Memory and Cognition,* **20**, 715-26.

Morrel-Samuels, P., and Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **18**, 615-23.

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

Ragsdale, J.D. and Silvia, C.F. (1982) Distribution of hesitation kinesic phenomena in spontaneous speech. *Language and Speech*, **25**, 185-90.

Rauscher, F. B., Krauss, R. M., and Chen, Y. (1996). Gesture, speech and lexical access: The role of lexical movements in speech production. *Psychological Science, 7*, 226-31.

Rimé, B. (1982). The elimination of visible behaviour from social interactions: Effects on verbal, nonverbal and interpersonal behaviour. *European Journal of Social Psychology*, **12**, 113-29.

Rimé, B. and Schiaratura, L. (1992). Gesture and speech. In *Fundamentals of Nonverbal Behavior* (ed. R.S. Feldman and B. Rimé), pp. 239-84, Cambridge, Cambridge University Press.

Riseborough, M. G. (1981). Physiographic gestures as decoding facilitators: Three experiments exploring a neglected facet of communication. *Journal of Nonverbal Behavior*, **5**, 172-83.

Rogers, W. T. (1978). The contribution of kinesic illustrators toward the comprehension of verbal behaviors within utterances. *Human Communication Research*, **5**, 54-62.

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

Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition, 47*, 1-24.

Schober, M. F. (1995). Speakers, addressees, and frames of reference: Whose effort is minimized in conversations about locations? *Discourse Processes, 20*, 219-47.

Sperber, D., and Wilson, D. (1986) *Relevance: Communication and cognition*. Cambridge, Harvard University Press.

Sumby, W. H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America, 26*, 212-15.

Teitelman, A. (1997). *Coverbal gesture in spontaneous speech and simultaneous translation from first to second, and from second to first, language.* Unpublished MA Thesis, Tel Aviv University.

Werner, H., and Kaplan, B. (1963). *Symbol formation*. New York, Wiley.

Williams, S., and Canter, G. (1987). Action naming performance in four syndromes of aphasia. *Brain and Language, 32*, 124-36.

ACKNOWLEDGMENTS