Silence Is Liberating: Removing the Handcuffs on Grammatical Expression in the Manual Modality

Susan Goldin-Meadow and David McNeill
University of Chicago

Jenny Singleton
University of Illinois at Urbana-Champaign

Grammatical properties are found in conventional sign languages of the deaf and in unconventional gesture systems created by deaf children lacking language models. However, they do not arise in spontaneous gestures produced along with speech. The authors propose a model explaining when the manual modality will assume grammatical properties and when it will not. The model argues that two grammatical features, segmentation and hierarchical combination, appear in all settings in which one human communicates symbolically with another. These properties are preferentially assumed by speech whenever words are spoken, constraining the manual modality to a global form. However, when the manual modality must carry the full burden of communication, it is freed from the global form it assumes when integrated with speech—only to be constrained by the task of symbolic communication to take on the grammatical properties of segmentation and hierarchical combination.

Perhaps the clearest example of the robustness of language comes from the fact that language is not tied to the mouth and ear but can also be processed by the hand and eye. Sign languages of the deaf have been found to take over all of the functions and to assume the grammatical properties characteristic of spoken languages (e.g., Klima & Bellugi, 1979). Moreover, when exposed to a conventional sign language such as American Sign Language, deaf children acquire the language as effortlessly as hearing children acquire spoken language (Newport & Meier, 1985). Thus, the manual modality can serve as a medium for language, suggesting that the capacity for creating and learning a linguistic system is modality independent. However, communication in the manual modality does not always assume languagelike properties. In particular, when hearing adults and children use their hands to gesture as they speak, those gestures do not take on the grammatical properties characteristic of speech (McNeill, 1992).

In this article, we propose a model that explains when the

manual modality will assume grammatical properties and when it will not. The model argues that when communicators do not attend to their gestures, as when they gesture unwittingly while they speak, the manual modality plays a role in relation to speech, and its form is constrained by that relationship. As a result, when it accompanies speech, manual communication is global and mimetic in form and is not characterized by segmentation and hierarchical combination—two hallmarks of grammatical structure. In contrast, when communicators are forced to rely solely on gesture, that is, when gesture is produced on its own without speech and thus itself must assume the full burden of communication, the manual modality is liberated from the constraints imposed on it by speech—only to be constrained to take on the grammatical properties essential to human language, most notably segmentation and hierarchical combination.

We begin by describing the gestures that accompany speech, focusing on two essential points: (a) The gestures produced along with speech form, with that speech, an integrated system, and (b) the roles that gesture and speech play in this integrated system are distinct, with speech conveying information in a linear and segmented fashion and gesture conveying information in a global and noncompositional manner. We then contrast the gestures that accompany speech with a naturalistic situation in which gesture has assumed the primary burden of communication and (we argue) has consequently taken on the grammatical properties that are essential to human language. Finally, we experimentally test our hypothesis that gesture becomes grammatical only when it assumes the full burden of communication. We asked hearing adults to describe a series of scenes using only their hands (and not their mouths), and we contrasted these self-sufficient gestures with the gestures these same adults spontaneously produced when asked to describe the same scenes verbally. We found that the gestures produced without speech were characterized by the rudiments of a simple syntax—in particular, they were segmented and combined hierar-

chically. These regularities of form, reminiscent of linguistic structure, were not found in the gestures that accompanied speech. Our findings suggest that grammatical expression is not dictated by the modality in which it is produced, nor is it restricted to established languages (signed or spoken) that have evolved over time. Rather, grammatical expression may well be a natural outgrowth of the way humans use symbols to communicate with other humans.

Gestures That Accompany Speech

Acts of speaking are often accompanied by movements of the hands and arms called *gestures* (Kendon, 1972, used the term *gesticulation*). We focus in this section on movements of this type, that is, on movements that occur only during speech and that are difficult to interpret in the absence of speech. As a result, we include gestures that have variously been called *pictograms, ideograms, illustrators,* and *batons* (Ekman & Friesen, 1969), but we exclude another, better known, more stereotypic type of gesture called *emblems* (Ekman & Friesen, 1969). Emblems are gestures that have conventional paraphrases or names and can be used as if they were spoken words. In many uses they are, in fact, unspoken words; for example, the "okay" sign can be used without speech in American culture to mean "things are fine." Our goal here is to describe spontaneous gestures that are not regulated by a conventional social code.

Gestures Convey Meaning

Whereas verbal behavior is assumed to be closely tied to a speaker's thoughts, nonverbal behavior, including gesture, has traditionally been assumed to reflect the speaker's feelings or emotions (Wundt, 1900/1973; for a review of studies focusing on gesture as a reflection of emotion and attitude see Feyereisen & de Lannoy, 1991). More recently, however, researchers who have focused on the hand gestures individuals produce while speaking (e.g., Kendon, 1980; McNeill, 1985, 1987, 1992) have argued that gesture can convey substantive information and, as such, can provide insight into a speaker's mental representation. For example, McNeill (1987) found that adults use hand gestures to portray concrete images (such as the actions or attributes of cartoon characters) as well as abstract concepts (such as mathematical concepts of quotients, factors, or even limits in calculus). Moreover, the gestures that adults produce when speaking have been found to convey the kinds of semantic information that speech conveys (although gesture cannot always be decoded as precisely and reliably as is speech; Krauss, Morrel-Samuels, & Colasante, 1991).

Children, too, often use hand gestures as they speak (Jancovic, Devoe, & Wiener, 1975), gesturing when asked to narrate a story (e.g., McNeill, 1992) or when asked to explain their responses to a problem (e.g., Church & Goldin-Meadow, 1986). The gestures children produce in a problem-solving situation have been used to provide insight into the way children represent those problems. For example, Evans and Rubin (1979) taught children between the ages of 5 and 10 to play a simple board game and then asked them to explain the game to an adult. The children's verbal statements of the rules were routinely accompanied by gestures that conveyed information

about their knowledge of the game. As a second example, Crowder and Newman (1993) found that gestures were an expected mode of communication in a sixth-grade science lesson on the seasons and that the gestures the students produced revealed knowledge that the children possessed about the seasons (e.g., a child discussing the seasons used both hands to produce a symmetrical gesture, laying down temperature bands on either side of the equator, and thus revealing, through her hands, knowledge of the symmetry of the hemispheres).

These examples from both adults and children make it clear that gesture can convey meaning and can offer a view of the speaker's mental representation.

Gesture and Speech Form an Integrated System

Gesture not only conveys meaning but it does so in a manner that is integrated with speech. Several types of evidence lend support to the view that gesture and speech form a single, unified system. First, gestures occur with speech. Whereas emblems may be delivered in utter silence, the gestures that are the focus of attention in this section are almost invariably accompanied by speech. McNeill (1992) found that 90% of gestures were produced when the gesturer was speaking. Thus, acts of speaking and gesturing are bound to each other in time.

Second, gestures and speech are semantically and pragmatically coexpressive. When people speak, they produce a variety of gesture types (iconics, metaphorics, beats, cohesives, deictics; cf. McNeill, 1992), and each type of gesture has a characteristic type of speech with which it occurs. For example, iconic gestures are gestures whose forms have a transparent relationship to the notions they convey. Iconics accompany utterances that depict concrete objects and events and fulfill a narrative function (i.e., they accompany the speech that tells the story). As an example, a speaker produced the following iconic gesture when describing a scene from a comic book in which a character bends a tree back to the ground: The speaker grasped his hand as though gripping something and pulled his hand back. He produced this gesture as he uttered the words "and he bends it way back," a concrete description of an event in the story. In contrast, metaphoric gestures are also pictorial, but the pictorial content presents an abstract idea rather than a concrete object or event. Metaphorics accompany utterances that refer to the structure of the discourse rather than to a particular event in the narrative (i.e., they accompany metanarrative rather than narrative utterances). As an example, a speaker produced the following metaphoric gesture when announcing that what he had just seen and was about to recount was a cartoon: The speaker raised his hands as though he were offering an object to the listener. He produced this gesture as he said "it was a Sylvester and Tweety cartoon," an utterance that set up and introduced the topic of discussion rather than formed part of the storyline. Other gesture types similarly have their own parallels with speech (see McNeill, 1992, chapter 7), suggesting a linked relationship between the two modalities.

Finally, gesture and speech are synchronous and thus form a unified system in this sense. The gesture and the linguistic segment representing the same information as that gesture are cotemporal. Specifically, the gesture movement—the "stroke"—lines up in time with the equivalent linguistic segment; for exam-

ple, in the previous example of an iconic gesture, the speaker produced the stroke of the gesture just as he said "bends it way back" (see Kita, 1993, for more subtle examples of how speech and gesture adjust to each other in timing; Morrel-Samuels & Krauss, 1992, for evidence that the timing of gesture and speech is related to the rated familiarity of the spoken word; and Scoble & Mayberry, 1994, for evidence that gesture and speech are synchronized even when, as in stuttering, the speech production process goes awry). Such synchrony implies that the speaker is presenting the same meaning in both channels at the same moment and that gesture and speech form a single, integrated system.

Gesture and Speech Convey Meaning Differently

Although gesture and speech refer to the same event, they do not always convey precisely the same information about that event. For example, when describing Granny's chase after Sylvester in a cartoon narrative, a speaker said "she chases him out again" while moving her hand as though swinging an object (McNeill, 1992). Speech conveys the ideas of pursuit and recurrence, but gesture conveys the weapon used (an umbrella) during the chase. Both speech and gesture refer to the same event, but each presents a somewhat different aspect of it.[1]

Why are gesture and speech so easily able to take different perspectives on the same event? We suggest that the method used by gesture for conveying meaning is fundamentally different from that of speech.

Speech has the effect of segmenting and linearizing meaning. What might be an instantaneous thought is divided up and strung out through time. A single event, say, somebody sitting down on a chair, must be conveyed in segments: the person, the chair, the movement, the direction, and so forth. These segments are organized into a hierarchically structured string of words. The total effect is to present what had been a single instantaneous picture in the form of a string of segments. Segmentation and linearization to form a hierarchy are essential characteristics of all linguistic systems (including sign languages that are not spoken at all, see below).

De Saussure (1916/1959) argued that the linear-segmented character of spoken language is a property that arises because language is unidimensional but meanings are multidimensional. Language can only vary along the single dimension of time. At all levels (phonemes, words, phrases, sentences, and discourse), language depends on variations along this one axis of time. This restriction forces language to break meaning complexes into segments and to reconstruct multidimensional meanings by combining the segments in time.

The gestures that accompany speech are different in their basic organization. They are themselves free to vary on dimensions of space, time, form, trajectory, and so forth and can present meaning complexes without undergoing segmentation or linearization. Unlike spoken sentences in which lower constituents combine into higher constituents, each gesture is a complete expression of meaning unto itself (McNeill, 1992). For example, in describing an individual running, a speaker moved his hand forward while wiggling his index and middle fingers. The gesture is a symbol in that it represents something other than itself—the hand is not a hand but a character, the movement is not a hand in motion but the character in motion, the

space is not the physical space of the narrator but a narra space, the wiggling fingers are not fingers but running 1 Moreover, the gesture is a symbol the parts of which gain m ing because of the meaning of the whole. The wiggling fin mean "running" only because we know that the gesture, whole, depicts someone running and not because this spe uses wiggling fingers to mean running in any other context deed, in other gestures produced by this same speaker, wigg fingers may well have a very different meaning (e.g., indeci between two alternatives). To argue that the wiggling fingers ture is composed of separately meaningful parts, one wc have to show that each of the three components that comp the gesture—the V hand shape, the wiggling motion, and forward motion—is used for a stable meaning across the spt er's gestural repertoire. The data suggest that there is no s stability in the gestures that accompany speech (McN 1992). Moreover, because the speaker does not consistently the forms that constitute the wiggling fingers gesture for st; meanings, the gesture cannot stand on its own without speec and this is consistent with the principle that speech and ges form an integrated system.

In addition to the fact that the gestures accompanying spe do not appear to be composed of meaningful parts and 1 are not wholes created from parts, these gestures also do themselves combine to create larger wholes. Most of the ti gestures are "one to a clause," that is, a spoken clause is acc panied by a single gesture (McNeill, 1992). Moreover, even (ing the times when more than one gesture occurs within a si clause, the gestures do not combine into a more complex ture "clause." Each gesture depicts the content from a diffe angle, bringing out a different aspect or temporal phase, each is a complete expression of meaning by itself. For exam while uttering the clause "and she grabs the knife," a spe; produced two gestures: The hand first groped in a circle \ the palm facing down and the fingers extended (produced as

---

[1] Instances in which gesture and speech convey different infor tion in a problem-solving situation have been called *mismat* (Church & Goldin-Meadow, 1986). Mismatches occur freque when individuals, particularly children, are asked to describe how solved a task (Alibali & Goldin-Meadow, 1993; Crowder & Newr 1993; Evans & Rubin, 1979; Goldin-Meadow, Alibali, & Chu 1993; Perry, Church, & Goldin-Meadow, 1988). At first glance, ture–speech mismatches might appear to contradict the claim gesture and speech form an integrated system. However, even thc mismatches do convey different information in gesture and in spe they do not violate the principles of gesture–speech integration scribed in the text. In particular, mismatches are semantically c pressive in the sense that both modalities describe the same even beit each taking a different perspective on that event. Consider example, a child in the conservation task who says that the amoui water is different because "the glass is tall" while indicating the w of the glass with her hands. Although this child is indeed expres two different pieces of information in the two modalities, she is ne theless describing the same object in gesture and in speech. More the timing of the mismatch also reflects an integrated system. child produces the width gesture as she says "tall," thus synchrono expressing her two perspectives on the glass (for further evidence mismatches are predicated on an integrated gesture–speech sys see Butcher & Goldin-Meadow, 1995).

word "she" was uttered) and then the hand turned up and closed to a fist as though gripping a knife (produced along with the words "grabs the knife"). The gestures are related but do not combine into a single higher unit characterized by the same properties as a spoken clause. Rather, the gestures present successive snapshots of the scene. The spoken words also describe this scene, but whereas the words—"she," "grabs," and "the knife"—combine to form the clause, the gestures—groping and grabbing—do not combine to form anything resembling a clause. Rather, each gesture represents a predicate unto itself.

In summary, the gestures that accompany speech form an integrated system with that speech and, within the gesture-speech system, convey semantic meaning. Gesture is not only semantically and pragmatically coexpressive with speech, but it is also synchronous with speech. As a result, when gesture is produced along with speech, its form is constrained by the framing that speech provides. In particular, gesture is constrained to take on a global and mimetic form.

Thus, when produced along with speech, gestures do not and—we would argue, because of the constraints imposed by speech—cannot assume the linear and segmented form of the words they accompany. In the next section, however, we will show that the manual modality at times can take on a linear and segmented form, and we will argue that it does so when it is a speaker's sole means of communication.

Gestures That Assume the Primary Burden of Communication

We begin by describing communication systems in the manual modality that are codified and that are used, and learned, as native languages by deaf people—conventional sign languages. We then turn to the "unconventional" and idiosyncratic gesture systems invented by deaf children who have not been exposed to conventional sign language to communicate with the hearing individuals around them.

Conventional Sign Languages

The sign languages of the deaf are autonomous languages that are not based on the spoken languages of hearing cultures (Bellugi & Studdert-Kennedy, 1980; Klima & Bellugi, 1979; Lane & Grosjean, 1980). A sign language such as American Sign Language (ASL) is a primary linguistic system passed down from one generation of deaf people to the next and is a language in the full sense of the word. Like spoken languages, ASL is structured at syntactic (Fischer, 1974; Liddell, 1980; Lillo-Martin, 1986; Padden, 1983), morphological (Fischer, 1973; Fischer & Gough, 1978; Klima & Bellugi, 1979; Newport, 1981; Supalla, 1982, 1986; Supalla & Newport, 1978), and "phonological"[2] (Battison, 1974; Coulter, 1990; Lane, Boyes-Braem, & Bellugi, 1976; Liddell, 1984; Liddell & Johnson, 1986; Padden & Perlmutter, 1987; Sandler, 1986; Stokoe, 1960; Wilbur, 1986) levels.

Thus, unlike the gestures that accompany speech, the signs of ASL combine with one another to create larger wholes, that is, sentences. Like many spoken languages, ASL has a basic or canonical word/sign order, although it has considerable word order flexibility. Subject–verb–object (SVO) is the unmarked order, with other orders possible only when one of the constituents is fronted and marked for topic (Fischer & Gough, 1978; Friedman, 1976; Liddell, 1980; Padden, 1983). Moreover, and despite the fact that the adult language shows considerable word order flexibility, word order plays an important role in the early stages of ASL acquisition. Deaf children learning ASL from their deaf parents use consistent word order as a syntactic device early in development (Hoffmeister, 1978; Newport & Ashbrook, 1977; Newport & Meier, 1985) and their preferred order is SVO, the canonical order of the adult system (Hoffmeister, 1978).

In addition, and again unlike the gestures that accompany speech, the signs of ASL are themselves composed of meaningful components, that is, morphemes. Like spoken languages, ASL has developed grammatical markers that serve as inflectional (e.g., Fischer, 1973; Fischer & Gough, 1978; Klima & Bellugi, 1979; Metlay & Supalla, 1995) and derivational morphemes (e.g., Supalla & Newport, 1978). In other words, there are regular changes in basic form associated with systematic changes in meaning, and these changes occur internal to a sign. For example, the ASL verb "ask both" is composed of two parts simultaneously produced: "ask," which (in its uninflected form) is produced by moving the index finger away from the chest area (neutral space) and bending it as it moves, and "both," which involves producing this same movement twice, once to the left and once to the right. This dual motion is a grammatical morpheme (marking the notion "dual person") that can be added to a set of verbs in ASL and, when added, it consistently contributes the meaning "both" to the verb (Klima & Bellugi, 1979).

Note that the simultaneous occurrence of morphemes within a sign can give an ASL sign an iconic quality. However, signs in ASL are not always iconic. For example, the sign for "slow" is made by moving one hand across the back of the other hand. When the sign is modified to be "very slow," it is made *more rapidly* because this is the particular modification of movement associated with an intensification meaning (Klima & Bellugi, 1979). Thus, modifying the meaning of a sign can reduce its iconicity in ASL simply because the meaning of the sign as a whole is, in rule-governed fashion, made up of the meanings of the components that comprise it.

In contrast, as described above, the gestures that accompany speech are not composed of parts but are instead noncompositional wholes. Because the gesture as a whole must be a good representation of its referent, the addition of semantic information to a spontaneous gesture always increases its iconicity—if something is thought of as very slow, the gesture for it is also very slow (McNeill, 1992). The gesture as a whole represents "very slow" and, although one could, in principle, break up the gesture into two parts (e.g., "slow," a movement across the back of the hand, and "very," an exaggerated and slowed movement), there is no evidence that these particular forms have independent and consistent meaning across a range of gestures—as they would have to if they were part of a morphological system like ASL's.

---

[2] This level is called phonological in sign language despite the absence of sound because it is composed of distinctive units that are meaningless and contrastive, hallmarks of the phonological level in speech.

Thus, conventional sign languages such as ASL, which are used by the deaf as their primary communication systems, exhibit the complexities of linguistic organization found in all spoken languages. Although the manual modality presents one with the opportunity to mimetically depict events as global wholes (as one sees in the gestures that accompany speech), when the modality assumes the full burden of communication, global representation is abandoned in favor of representation by parts that combine to form a system of hierarchically arranged sequences of linear segments (i.e., morphemes combine to form signs, which, in turn, combine to form sentences). We suggest that this type of segmented representation arises in sign language (and not in spontaneous gesture) because, in sign language, the manual modality itself assumes the primary burden of communication.

Sign languages of the deaf, in addition to being primary communication systems, are also systems that have histories and are passed down from one generation of users to the next (Frishberg, 1975). They are codified linguistic systems and, in this sense, are very different from the gestures that accompany speech. In the next section, we examine a situation in which noncodified gesture itself assumes the full burden of communication. We ask whether gestures that take on the function of a primary communication system more closely resemble in form the gestures that accompany speech or the signs of a conventional language.

When Gesture Is the Only Means of Communication Available

As mentioned above, deaf children born to deaf parents and exposed from birth to a conventional sign language such as ASL acquire that language naturally; that is, these children progress through stages in acquiring sign language similar to those of hearing children acquiring a spoken language (Caselli, 1983; Hoffmeister, 1978; Hoffmeister & Wilbur, 1980; Kantor, 1982; Newport & Ashbrook, 1977; Newport & Meier, 1985). Thus, in an appropriate linguistic environment—in this case, a signing environment—deaf children are not handicapped with respect to language learning.

However, 90% of deaf children are not born to deaf parents who could provide early exposure to a conventional sign language. Rather, they are born to hearing parents who, quite naturally, tend to expose their children to speech (Hoffmeister & Wilbur, 1980). Unfortunately, it is extremely uncommon for deaf children with severe to profound hearing losses to acquire the spoken language of their hearing parents spontaneously, that is, without intensive and specialized instruction. Even with instruction, deaf children's acquisition of speech is markedly delayed when compared either with the acquisition of speech by hearing children of hearing parents or with the acquisition of sign by deaf children of deaf parents. By age 5 or 6, and despite intensive early training programs, many profoundly deaf children have only a very reduced oral linguistic capacity (Conrad, 1979; Mayberry, 1992; Meadow, 1968).

In addition, unless hearing parents send their deaf children to a school in which sign language is used, these deaf children are not likely to receive conventional sign language input. Under such inopportune circumstances, these deaf children

might be expected to fail to communicate at all or perhaps to communicate only in nonsymbolic ways. This turns out not to be the case.

Studies of deaf children of hearing parents have shown that these children spontaneously use gestures (referred to as *home signs*) to communicate, even if they are not exposed to a conventional sign language model (Fant, 1972; Lenneberg, 1964; Moores, 1974; Tervoort, 1961). Given a home environment in which family members communicate with each other through many different channels, one might expect that the deaf child would exploit the accessible modality (the manual modality) for the purposes of communication. The question we ask here is whether the gestures the child uses to communicate are structured in languagelike ways or whether they resemble the global, nonsegmented gestures that the hearing individuals in their environment use as they speak.

Gestures are combined into strings. Goldin-Meadow and her colleagues (Feldman, Goldin-Meadow, & Gleitman, 1978; Goldin-Meadow, 1979; Goldin-Meadow & Feldman, 1977; Goldin-Meadow & Mylander, 1984, 1990a) observed the home signs of 6 deaf children of hearing parents in Philadelphia and 4 in Chicago and found that all 10 children used gesture to communicate despite the fact that they were not exposed to a conventional sign language. The children used three major types of gestures: (a) Deictic gestures were typically points that maintained a constant kinesic form in all contexts. These deictics were used to single out objects, people, places, and the like. (b) Characterizing gestures were stylized pantomimes, the iconic forms of which varied with the intended meaning of each sign (e.g., a fist pounded in the air as if someone was hammering). Characterizing gestures were often used as predicates conveying action and attribute information (i.e., as verbs and adjectives, respectively) but could also be used as labels for objects (i.e., as nouns; see Goldin-Meadow, Butcher, Mylander, & Dodge, 1994). (c) Marker gestures were typically head or hand movements (e.g., nod, headshake, or two-handed "flip") that are conventionalized within the child's culture and that the children used as modulators (e.g., to affirm, negate, or doubt).

Unlike the gestures that hearing adults and hearing children produce along with speech, which are rarely combined with one another, the home-signing deaf children's gestures were frequently combined into strings. Motoric criteria were used to isolate strings of gestures. In particular, in the analyses of the deaf children's gestures, two gestures were considered part of the same gesture string if they were connected by a continuous flow of movement. In contrast, two gestures were considered distinct, each comprising its own unit, if the hand retracted or relaxed between the two gestures.

Moreover, the children's gesture strings conveyed the range of semantic relations typically found in early child language (in particular, action and attribute relations). The children combined points with other points (e.g., a point at a toy followed by a point at the table to request that the toy be transferred to the table), points with characterizing gestures (e.g., a point at a bubble jar followed by the characterizing gesture "twist" to comment on the fact that the jar has just been twisted open), and characterizing gestures with other characterizing gestures (as exemplified in the next paragraph).

Strings of characterizing gestures were of two types. The first

type were strings in which each characterizing gesture represented a separate proposition (e.g., a "hit"-characterizing gesture followed by a "fall"-characterizing gesture, used to comment on the fact that the child had hit the tower [act$_1$] and the tower had fallen [act$_2$]). Note that this type of characterizing + characterizing string, conveying as it does two propositions within the bounds of a single string (as defined by the motoric criteria described above), suggests that the deaf child's system is recursive (a system in which a unit can be used to derive, by rules of the system, a longer string that contains the original unit). Recursion is a central component of all natural languages and gives language its generative capacity. The second type were strings in which the two characterizing gestures were part of the same proposition, with one gesture representing the verb of the proposition and the other representing a noun within that same proposition (e.g., an "eat"-characterizing gesture followed by a "give"-characterizing gesture, used to request that a toy grape, which was in fact inedible, be given to the child who then threw the grape rather than eating it). In this instance, the "eat" gesture was used to identify an object and thus appeared to be playing a nounlike role in the string (see Goldin-Meadow et al., 1994, for evidence that noun and verb uses of characterizing gestures are distinguished within the deaf child's gestures).

Gesture strings are structured. The deaf children's gesture strings shared several structural properties with early sentences in child language and, on this basis, warrant the label "sentence." For example, although the children did not produce gestures for all of the possible thematic roles that could be conveyed within a sentence, they were not haphazard in their selection of which roles to convey in gesture. The children were more likely to produce a gesture for the patient (e.g., cheese) in a sentence about eating than to produce a gesture for the actor (e.g., mouse). In addition, the children produced gestures for the intransitive actor (e.g., the mouse in a sentence describing a mouse running to his hole) as often as they produced gestures for the patient (e.g., the cheese in a sentence describing a mouse eating cheese) and far more often than they produced gestures for the transitive actor (e.g., the mouse in a sentence describing a mouse eating cheese). In this way, the likelihood of production served to distinguish among thematic roles.

As a second example of structure, the children also distinguished among the thematic roles they did express by placing the gesture for a given role in a particular position in a gesture string; that is, the gestures the children produced within their strings were not produced in haphazard order but rather appeared to follow a small set of gesture order regularities (Goldin-Meadow, 1979; Goldin-Meadow & Mylander, 1984). Gesture order regularities describe where the gesture for a particular thematic role is likely to appear in a gesture string. For example, gestures for an object playing the patient role tended to precede gestures for the act; for example, to comment on the fact that he intended to throw a toy grape, one child produced the string "grape throw"—a pointing gesture for the patient (grape) followed by a characterizing gesture for the act (throw). As a second example, gestures for the act tended to precede gestures for an object playing the role of recipient or goal; for example, to request that an object be moved to the table, a child produced the string "transfer table"—a characterizing gesture for the act (transfer) followed by a pointing gesture for the re-

cipient (table). Finally, gestures for an object playing the role of patient tended to precede gestures for the object playing the role of recipient; for example, to comment on the fact that a toy duck had been handed to his sister, a child produced the string "duck sister"—a pointing gesture for the patient (duck) followed by a pointing gesture for the recipient (sister).

Thus, when the children produced gestures for the patient, act, or recipient within a single string, they tended to produce these gestures according to the following order:

gesture string → patient–act–recipient.

In summary, the deaf children conjoined the gestures they produced into strings characterized by gesture order regularities as well as by likelihood of production regularities. The gesture strings could therefore be said to conform to a syntax, albeit a simple one.

Gestures are themselves composed of parts. The deaf children's gestures not only formed parts of longer sentence units but they themselves were made up of smaller parts. For example, to request the experimenter to lay a penny down flat on a toy, one deaf child produced a downward motion with his hand shaped like an O. In itself this could be a global gesture presenting the shape and trajectory as an unanalyzed whole. The experimenter pretended not to understand and, after several repetitions, the child factored the gesture into its components: First he statically held up the gesture for a round object (the O hand shape) and then, quite deliberately and with his hand no longer in the O shape but exhibiting a flat palm, made the trajectory for downward movement. The original gesture was thus decomposed into two elements. This example implies the presence of a system of linguistic segments in which the complex meaning of "round + thing + moving + downward" is broken into components and the components combined into a gesture. The experimenter's feigned lack of understanding may have been instrumental in getting the child to factor his gesture into parts at that particular moment. However, the point we stress here is that, when the gesture was broken into parts, those parts conformed to a wider system, one that accounted for the vast majority of gestures produced by this particular child (Goldin-Meadow & Mylander, 1990b). Thus, the deaf child's gestures—unlike the spontaneous gesticulations of hearing adults and children—are themselves composed of simpler gesture elements.[3]

The particular combination described above was one among many in the deaf child's gesture system, for indeed this child, as well as 3 others whose gestures have been analyzed for components, had each devised a morphological system (Goldin-Meadow & Mylander, 1990b; Goldin-Meadow, Mylander, & Butcher, 1995). Systematic compositionality of gestures within

---

[3] It is possible that speakers who gesture in contexts in which speech is present but not fully audible would demonstrate a similar process of factoring their gestures as they speak—for example, in Times Square at the height of the New Year's Eve frenzy. However, we suspect that, because of the tightly timed arrangement that gesture maintains with the speech it accompanies (cf. Kita, 1993; McNeill, 1992), speakers in such a situation would factor their gestures into component parts only if they also factored their speech into parts (or if they gave up and stopped talking altogether).

a system of contrasts is crucial evidence of segmentation and combination.

As an example of how this child combined the components in his gestures to contrast with one another, the fist hand shape (a component representing "narrow, long objects" in the child's system) was frequently combined with a short arc motion (a component meaning "reposition"). The meaning of the fist + short arc combination was a composite of the meanings of its component parts—"reposition a narrow, long object" (used, for example, to describe pulling out a newspaper). When the same fist hand shape was combined with a different motion, an arc to-and-fro (meaning "move back and forth"), the meaning of this new fist + arc to-and-fro combination reflected the changed motion and the constant hand shape—"move back and forth a narrow, long object" (used to describe waving a balloon string back and forth). The child's gestures can therefore be said to conform to a paradigm or system of contrasts.

The gesture systems of 4 children have been analyzed at this level and, although similar in many respects, were sufficiently different to suggest that the children had introduced relatively arbitrary distinctions into their systems. For example, in contrast to the first child who used the fist hand shape to represent narrow and long objects, 2 of the other children used the same hand shape to represent objects that were also narrow but could be of any length (Goldin-Meadow et al., 1995). However, all of the children produced gestures that could be characterized by paradigms of hand shape and motion combinations. Thus, each child was found to use a limited set of discrete hand shape and motion forms—that is, the forms were categorical rather than continuous; to consistently associate each hand shape or motion form with a particular meaning (or set of meanings) throughout the corpus—that is, each form was meaningful (and, in this sense, the children's hand shape and motion units formed part of a morphological, as opposed to a phonological, system); and to produce most of the hand shapes with more than one motion and most of the motions with more than one hand shape—that is, each hand shape and motion was an independent and meaningful morpheme that could combine with other morphemes in the system to create larger meaningful units (the system was combinatorial).

The children's gestures are different from their hearing mothers' gestures. Despite their limited access to a conventional linguistic model, the deaf children were found to elaborate a gestural communication system that was linear and segmented. Nevertheless, even if the deaf child is not making use of a conventional language model, it is possible that the child's hearing parents are responsible for the form of the child's gestures. In an effort to communicate, the hearing parents of these deaf children might have spontaneously generated segmented and combinatorial gestures that their children then learned.

However, in an analysis of the syntactic structure of gestures produced by six hearing mother–deaf child pairs, Goldin-Meadow and Mylander (1983, 1984) found that the hearing mothers produced very few gesture strings and the gesture strings that they did produce either were not structured or were structured differently from their children's gesture strings. To determine whether the deaf children used the spontaneous gestures their hearing parents produced as a model for their own morphological systems, Goldin-Meadow and her

colleagues (Goldin-Meadow & Mylander, 1990b; Goldin-Meadow et al., 1995) coded the gestures each mother produced within the framework of the morphological system developed by her child. Each mother was found to use her gestures in a more restricted way than her child, omitting many of the morphemes that the child produced (or using the ones she did produce more narrowly than the child) and omitting completely many of the hand shape/motion combinations that the child produced. In addition, the mothers' gestures were analyzed with the same tools used to analyze the children's gestures; that is, the mother's gestures were treated as a system unto itself. However, the resulting systems for the mothers did not capture their children's gestures well at all. Moreover, the differences that were found across the children's systems could not be traced to their mothers' systems.

Why were the gestures of mother and child so different? The deaf children's gestures were structured in languagelike ways (both syntactic and morphologic) despite the fact that their hearing mothers' gestures were not. Why was there so little in common between the gesture systems of mother and child? One might have expected that because they interact with one another on a daily basis, mother and child would develop gesture systems that resemble one another. They did not.

We suggest that the hearing mothers' gestures and the deaf children's gestures were structured so differently because the hearing mothers produced gestures for a very different purpose than did the deaf children (cf. Goldin-Meadow, 1993). The deaf children used gesture as their sole means of communication; thus, their gestures were forced to assume the full burden of communication. In contrast, the hearing mothers rarely gestured without speaking (not surprisingly, given that the mothers were committed to teaching their children spoken English); thus, their gestures were all produced along with speech and served a role in relation to that speech, which itself assumed the primary burden of communication. In fact, the gestures that the hearing mothers produced appeared to be no different from the gestures that hearing individuals typically use with speech (Bekken, Goldin-Meadow, & Dymkowski, 1990).

As we have argued above, the gestures that hearing individuals produce along with speech form an integrated system with that speech; it is only when analyzed without speech (i.e., when coded with the sound turned off) that these gestures appear to be unsystematic. Thus, we suggest that because the gestures produced by the deaf children's hearing mothers formed an integrated system with their speech and were constrained by that speech, those gestures were not "free" to assume the languagelike qualities of their deaf children's gestures. One might suspect that if the mothers merely refrained from speaking as they gestured, their gestures might have become more languagelike in structure, assuming the segmented and combinatorial form also found in the children's gestures. The next section constitutes an experimental test of this prediction.

Gesture With and Without Speech: An Experimental Manipulation

We have shown that the manual modality, when it is used to accompany speech, does not take on the segmented and hierarchical combinatorial form that is characteristic of language.

The manual modality can, however, assume this form, and it does so when it serves as a primary communication system, as in the conventional sign languages of the deaf or the idiosyncratic gesture systems used by deaf children of hearing parents who use gesture as their sole means of communication. These findings suggest that combinatorial structure at more than one level (i.e., within-word and across-word structure) is fundamental to human language—so fundamental that it will arise in human languages that are produced in another modality, the manual modality, even in those manual systems developed without access to a conventional language model.[4] Thus, hierarchical combinatorial structure is not forced on language by the modality in which it is produced, nor is it maintained as a universal property of language solely by historical tradition.

We explore here the hypothesis that segmented and hierarchical combinatorial structure will arise spontaneously, at least in the manual modality, when gesture is required to carry the primary burden of communication, but not when gesture is used in relation to speech. To test this hypothesis, we conducted an experiment in which hearing adults were asked to describe a series of scenes with and without speech. We predicted that the gestures the adults would produce without speech would be segmented and hierarchically combinatorial and, as such, would be distinct from the gestures these same adults would produce when they described the scenes verbally.

The Stimulus for a Study of Gesture With and Without Speech

To explore the way in which individuals use gesture with and without speech, we needed a stimulus that would evoke communication under both conditions. We chose to use the Verbs of Motion Production (VMP) test developed by Supalla (1982; see also Supalla et al., in press) to assess knowledge of verbs of motion in native ASL signers. This test is composed of a series of unrelated scenes in which small objects move across space. The test involves no conventional language (i.e., neither sign nor speech) and thus could easily be adapted for use with individuals who are not familiar with ASL, and could be used to elicit gesture without speech as well as gesture with speech. We begin with a description of the structure of verbs of motion in ASL in general, and the VMP test in particular.

Verbs of motion in ASL. Research on stem-formation processes has found ASL to be comparable to those spoken languages that are morphologically quite complex. This research has focused on signs that are highly mimetic in form (as contrasted with the "frozen" signs of ASL that are listed in ASL dictionaries as single-morpheme stems). Mimetic signs in ASL were originally thought to be built on an analog use of movement and space in which movement is mapped in a continuous rather than a discrete fashion (Cohen, Namir, & Schlesinger, 1977; DeMatteo, 1977). In other words, mimetic signs were thought not to be divisible into component parts, but rather were considered unanalyzable lexical items that mapped, as wholes, onto events in the world (and, as such, would have been comparable to the gestures that accompany speech). However, more recent research has shown these mimetic signs to be composed of combinations of a limited set of discrete morphemes (McDonald, 1982; Newport, 1981; Schick, 1987, 1990; Su-

palla, 1982). For example, to describe a drunk's weaving walk down a path, an ASL signer would not represent the idiosyncrasies of the drunk's particular meanderings, but would instead use a conventional morpheme representing random movement (i.e., a side-to-side motion) in conjunction with a conventional morpheme representing change of location.

The VMP test. Newport and Supalla (1992; Newport, 1990) have shown that verbs of motion in ASL contain as many as six simultaneously produced morphemes affixed on a single verb stem. Supalla (1982) designed the VMP test to explore the types of morphemes that constitute the primary morphology of ASL verbs of motion (five motion and location and two hand shape morphemes). Every ASL verb of motion requires at least a central object hand shape morpheme (indicating the class of the object that is moving, i.e., its category [e.g., a human or a vehicle] or its shape [e.g., round or straight]) and a root motion morpheme (indicating the type of path traversed by the moving object; e.g., a linear path, an arced path, or a circle). For example, the root morpheme "linear path" (representing change of location along a straight path) can be combined with one of many possible central object morphemes representing the moving object (e.g., bent V = a small animal; thumb pointing up with the index and middle fingers extended = a vehicle). These combinations create a set of signs for which the meanings are predictable from the meanings of the individual motion and hand shape morphemes (i.e., a small animal moves along a straight path or a vehicle moves along a straight path). Along with the root and central object morphemes that are requisite in every verb of motion, the verb may also contain a morpheme for a secondary object. For example, if the vehicle hand shape were added to the root and central object morphemes in the previous example, the resulting meaning would be "a small animal moves along a straight path in relation to a vehicle."

The VMP test is composed of 120 short-filmed events of toy people and objects that move in varying paths and manners of motion; for example, a doll jumping into a hoop or a robot moving past a motorcycle. The test items are each constructed to elicit a single verb of motion.

---

[4] Although the deaf children in Goldin-Meadow's studies were developing their gesture systems without direct access to a usable conventional language model, they were using their gestures to communicate with hearing individuals who did have knowledge of a conventional language. The hearing parents of the deaf children might have selectively reinforced particular gestural combinations that their children began using on a trial-and-error basis. If so, the structure found in the deaf children's gestures would be attributable to the children's hearing parents (who did have a conventional language) and not to the children (who did not have such a model). However, in a study of six mother-child dyads, Goldin-Meadow and Mylander (1983, 1984) found that the hearing mothers did not respond to their deaf children's gesture strings on the basis of the order in which the gestures in those strings were produced. The mothers were as likely to respond to strings that conformed to the child's predominant gesture order pattern as they were to respond to strings that did not conform to this pattern. These findings suggest that maternal responsiveness, and thus a conventional language model by proxy, was not shaping the structural patterns found in the deaf children's gestures.

42   GOLDIN-MEADOW, McNEILL, AND SINGLETON

Method

Participants and procedure. Sixteen hearing adults participated in the study. None of the hearing participants had knowledge of ASL or any other sign language. The adults were all students at the University of Chicago and were recruited through sign-up sheets distributed in psychology classes or posted in various campus buildings. We presented each participant with 40 segments from the VMP test selected so that half contained a single moving object and half contained one object moving in relation to a second object. The test was administered by the same experimenter (Jenny Singleton) for all participants in a quiet room at the University of Chicago and was videotaped. Each participant was first asked to view the segments and describe them; no mention was made of gesture for this first pass through the segments, although it was anticipated that gesturing would occur spontaneously (the gesture + speech condition). The participant was then asked to view the segments again, this time using gesture and no speech to depict what happened in each segment (the gesture condition).

Coding the form of individual gestures. The vast majority of gestures produced in both conditions were iconic. The occasional metaphoric gesture or beat produced in the gesture + speech condition was described but not included in further analyses. Each of the gestures produced in the two conditions was described along two dimensions using categories established by Supalla (1982): the shape of the hand and the trajectory of the motion. For example, to describe a circle moving diagonally across the video screen, one participant shaped his index finger and thumb into a circle and moved the circle from the left, at waist level, up and across to the right, at eye level. The hand shape was considered an O and the motion was considered a linear path.

We relied on changes in the shape of the hand or in the trajectory of the motion to determine whether a particular hand shape and motion combination formed a single gesture (as in the preceding example) or two gestures. For example, when describing the circle moving across the videoscreen, if the participant held his O-shaped hand in space but, before moving his hand in a linear path, changed his hand shape from an O to an index finger point, we would attribute to the participant two gestures (the O-hand held in space and the point moved in a linear path) rather than one.

Coding strings of gestures. If more than one gesture was produced to describe a particular scene, we determined whether those gestures were connected into a string using the following criterion. Two gestures were considered part of the same gesture string if they were connected by a continuous flow of movement. Two gestures were considered distinct, each comprising its own unit, if the hand retracted or relaxed or if there was a pause longer than one second between the two gestures. For example, if the participant held an O hand shape in place at chest level and then, without dropping his hand or pausing, moved a pointing hand diagonally from left to right, the two gestures would be considered to constitute a gesture string. If, however, the participant relaxed his hand before producing the diagonal gesture or paused for longer than one second the two gestures would be considered separate units. The same criteria were used to determine gesture connectedness in both conditions. As a result, whether or not two gestures were produced along with a single spoken clause in the gesture + speech condition did not contribute at all to our decision as to whether those two gestures were connected; indeed, we made these decisions with the sound turned off.

Coding the meaning of individual gestures. In addition to coding decisions involving the form of the gesture, we also made coding decisions involving meanings and attributed meaning to each of the iconic gestures produced in the two conditions. We coded the gestures in the gesture + speech condition with the sound turned off, but, in both conditions, we knew which scene the participant was describing and coded the gestures in relation to that scene. Three semantic elements appeared in the scenes and could, in principle, be conveyed gesturally—the action, the moving object, and the stationary object.

If the participant produced a movement that captured any aspect of the motion portrayed in the scene, the gesture was assumed to represent the action in the scene and was consequently labeled *action.* For example, if, to describe a doll jumping into a hoop, the participant arced her hand from Point 1 to Point 2, the gesture was considered a symbol for the motion and thus an action gesture.

As described above, half of the scenes shown to each participant involved a single object moving in space (e.g., a wind-up toy bird rolls across the screen), and half involved two objects, one moving in relation to the other (e.g., a robot moves past a motorcycle). For both the one-object and the two-object scenes, if the participant produced a gesture that captured an aspect of the object that was moving in the scene, it was assumed to represent that object and was consequently labeled *moving object.* For example, if, to describe a doll with curly hair walking across the screen, the participant moved both hands in curllike shapes at the side of her head, that gesture was considered to be a symbol for the doll, and thus a moving object gesture.

For the two-object scenes, if the participant produced a gesture that captured an aspect of the object that did not move in the scene, it was assumed to represent that object and was consequently labeled *stationary object.* For example, if, to describe the doll jumping into the hoop, the participant shaped his index finger and thumb into a circle, holding the shape at chest level, that gesture was considered to be a symbol for the hoop, and thus a stationary object gesture.

Coding object information incorporated into the action. At times, the hand shape of the action gesture portrayed an aspect of either the moving object or the stationary object. For example, to describe a circle moving diagonally across the screen, one participant shaped her index finger and thumb into a circle and moved the circle from left to right and upward. Because it conveyed the movement in the scene, the gesture was considered to be an action. However, the gesture did incorporate information about the circle, the moving object, and was consequently classified as an action that had moving object information incorporated into its form. As an example of incorporation of the stationary object, to describe the doll jumping into a hoop, one participant moved her right index finger across space and into her left hand, shaped as a circle. The gesture was considered to be an action but, because it conveyed information about the circle, this time playing the role of the stationary object, the action gesture was classified as incorporating the stationary object.

Reliability of the coding decisions. Coders were trained to transcribe the form and meaning of the gestures according to a system adapted from two sources—Supalla's (1982) system for coding signed languages and McNeill's (1992) system for coding gesticulation. Intercoder reliability was calculated between a pair of coders blind to the hypotheses of the study on a subset of the data. Reliability was assessed separately for the gesture condition and the gesture + speech condition. For describing the form of individual gestures (including decisions about whether a handshape/motion combination formed one or two gestures), reliability was 85% agreement in the gesture condition and 80% in the gesture + speech condition. For deciding whether two gestures were connected into a string, reliability was 100% in the gesture condition and 93% in the gesture + speech condition. For classifying a gesture as either an action, a moving object, or a stationary object, reliability was 85% in the gesture condition and 100% in the gesture + speech condition. Finally, for deciding whether an action incorporated information about an object into its form, reliability was 98% in the gesture condition and 100% in the gesture + speech condition. The data were analyzed using paired *t* tests, comparing responses in the gesture condition with those in the gesture + speech condition.

Results

Not surprisingly, participants were more likely to produce more gestures in the gesture condition, in which they were spe-
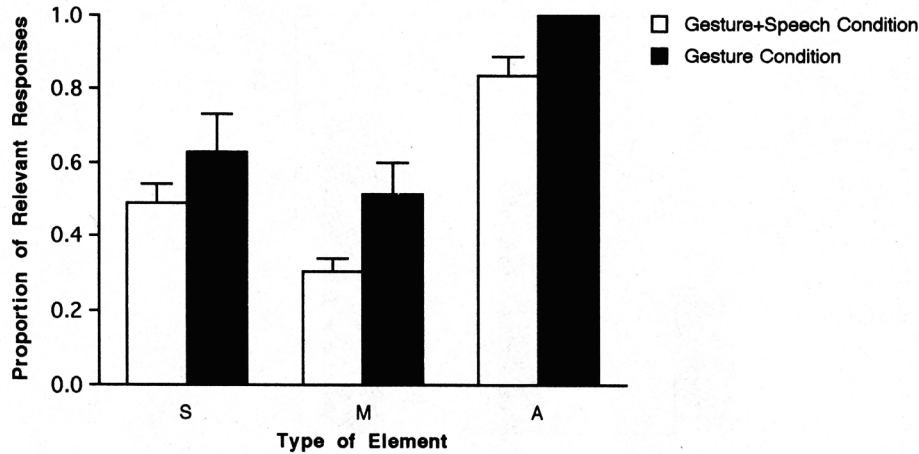
Figure 1. The mean proportion of semantic elements conveyed in the gesture + speech condition and the gesture condition. The proportion for each semantic element reflects the number of responses that contained a gesture for that semantic element, divided by the total number of gestured responses in which the element could have been conveyed. In both conditions, participants produced gestures for a sizable proportion of each of the three elements. S = stationary object; M = moving object; A = action. The error bars reflect the standard error.

cifically instructed to do so, than in the gesture + speech condition, in which gesture was not mentioned in the task instructions. The participants produced gestures for all 40 of the scenes in the gesture condition, but for only an average of 19.4 of the scenes in the gesture + speech condition (range from 8–37, SD = 10.22). As a result, the total number of gestures produced in the gesture condition (M = 96.9, SD = 47.6) was greater than the total number produced in the gesture + speech condition (M = 38.6, SD = 29.7). However, in both the gesture and the gesture + speech conditions, when participants did produce gestures, they tended to produce more than one gesture (as opposed to just one) per response; on average, participants produced 2.4 (SD = 1.2) gestures per response in the gesture condition and 1.8 gestures per response (SD = 0.54) in the gesture + speech condition. Because participants did not produce gestures for every scene in the gesture + speech condition, and because we are particularly interested in comparisons of how participants used the gestures they did produce, we analyzed the data in the gesture versus gesture + speech conditions as proportions of the responses each participant produced that included gestures (see below).

Semantic elements conveyed in gesture. We begin by examining how often participants conveyed in gesture each of the three semantic elements—the action, the moving object, and the stationary object—in the two conditions. Figure 1 presents the mean proportion of gestures for actions, moving objects, and stationary objects produced by participants in the gesture + speech condition and the gesture condition. The proportion for each semantic element reflects the number of responses that contained a gesture for that semantic element, divided by the total number of responses in which the element could have been conveyed. Thus, for the gesture condition in which all participants produced at least some gestures in all 40 scenes, the denominator was 40 for the moving object and the action (because both are involved in all 40 of the scenes—the one-object scenes both are involved in all 40 of the scenes—the one-object scenes

as well as the two-object scenes) and 20 for the stationary object (because they were involved only in the two-object scenes). For the gesture + speech condition, the denominator for each individual differed and depended on the particular number of one-object and two-object scenes for which that participant produced any gestures at all.

As can be seen in Figure 1, participants were very likely to represent the action in their gestures in both conditions; all of the responses in the gesture condition contained actions, as did 84% of the gestured responses in the gesture + speech condition. The objects involved in the action were less frequently represented in gesture in both conditions, but still found fairly often; 52% and 63% of the responses in the gesture condition contained gestures for moving objects and stationary objects, respectively, as did 30% and 49% of gestured responses in the gesture + speech condition, respectively. Each of the three semantic elements was more likely to be gestured in the gesture condition than in the gesture + speech condition. This difference was significant for actions, $t(15)$ = 3.53, p = .003, and moving objects, $t(15)$ = 2.98, p = .009, but not for stationary objects, $t(15)$ = 1.14, p = .27, ns. However, what is particularly impressive is not the differences between the two conditions but the similarities. Participants in both the gesture and the gesture + speech conditions used gesture to convey substantive information about the semantic elements in the scenes and, in both conditions, they did so relatively frequently. Moreover, the pattern of production was the same in the two conditions. Participants in both the gesture and the gesture + speech conditions produced more gestures for actions than for stationary objects and more gestures for stationary objects than for moving objects. We next consider the form that the gestures for these semantic elements took in the two conditions.

Semantic elements conveyed in gesture strings. In the gesture + speech condition, participants tended to produce the gestures for their semantic elements as single units; that is, they
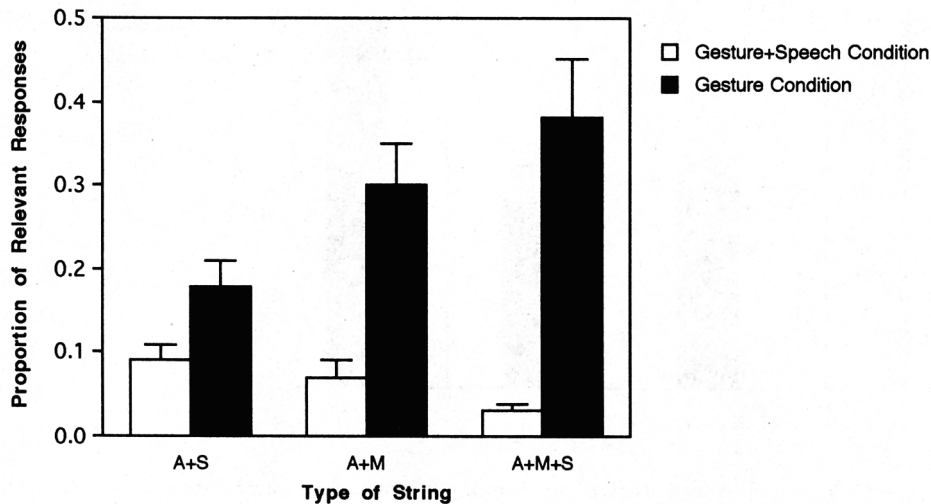
Figure 2. The types of elements combined within a single gesture string in the gesture + speech condition and the gesture condition. The figure presents the mean number of responses containing a particular combination of elements within a string, as a proportion of the total number of relevant responses containing gestures. Relevant responses included the gestured responses to both the one-object and the two-object scenes for A + M combinations, but only the gestured responses to the two-object scenes for the A + S and the A + M + S combinations. Participants were more likely to produce their gestures for semantic elements in strings (i.e., concatenated into units larger than the gesture itself) in the gesture condition than in the gesture + speech condition. S = stationary object; M = moving object; A = action. The error bars reflect the standard error.

rarely connected the gesture for a semantic element with another gesture in a gesture string. Eighty-two percent of the semantic elements (actions, moving objects, and stationary objects) that were conveyed in gesture in the gesture + speech condition (i.e., the semantic elements that were gestured in the gesture + speech condition of Figure 1) were conveyed as single gestures, unconnected to any other gestures. In contrast, in the gesture condition, only 36% of the semantic elements that were conveyed in gesture were conveyed as single, unconnected gestures, $t(15) = 5.32, p < .0001$.

This finding—that participants were significantly less likely to conjoin their gestures for semantic elements into strings in the gesture + speech condition than in the gesture condition—is particularly striking because as described above, participants often produced more than one gesture per response in the gesture + speech condition and thus did have the opportunity to conjoin their gestures for semantic elements into strings. Indeed, and most telling, even if we restrict our analysis only to those responses in which two or more elements were gestured, we still find that participants were less likely to conjoin their gestures into strings in the gesture + speech condition than in the gesture condition; 53% of the responses that had more than one semantic element contained *no* conjoined strings in the gesture + speech condition versus 3% in the gesture condition, $t(15) = 5.74, p < .0001$. Thus, although participants produced almost as many gestures for actions, moving objects, and stationary objects in the gesture + speech condition as in the gesture condition (cf. Figure 1), they infrequently produced those elements in gesture strings in the gesture + speech condition.

Figure 2 describes the particular combinations of elements found in the gesture strings the participants produced in each of the two conditions. The figure presents the number of responses containing a particular combination of elements in a string, as a proportion of the total number of relevant responses containing gestures that were produced in the two conditions. The denominator for the gesture condition in Figure 2 was 40 for strings containing the moving object and the action (because both are involved in all 40 of the scenes—the one-object scenes as well as the two-object scenes) and 20 for strings containing the stationary object and the action, and for strings containing the stationary object, the moving object, and the action (because the stationary object was involved only in the two-object scenes). For the gesture + speech condition, the denominator for each individual differed and depended on the particular number of one-object and two-object scenes for which that participant produced any gestures at all. Participants produced significantly more strings in the gesture condition than in the gesture + speech condition for all three types of string combinations—strings containing three elements, the action, the moving object, and the stationary object (38% in the gesture condition vs. 3% in the gesture + speech condition), $t(15) = 5.05, p < .0001$—and both types of strings containing two elements, the action and the moving object (30% vs. 7%), $t(15) = 4.63, p < .0001$, and the action and the stationary object (18% vs. 9%), $t(15) = 2.24, p = .04$.

The order of semantic elements in gesture strings. As Figure 2 indicates, participants tended to produce their gestures in strings relatively often, particularly in the gesture condition. We next asked whether they were likely to place gestures for particular semantic elements in particular positions within their strings; that is, could their gesture strings be characterized by a simple syntax? Figure 3 presents the number of strings contain-
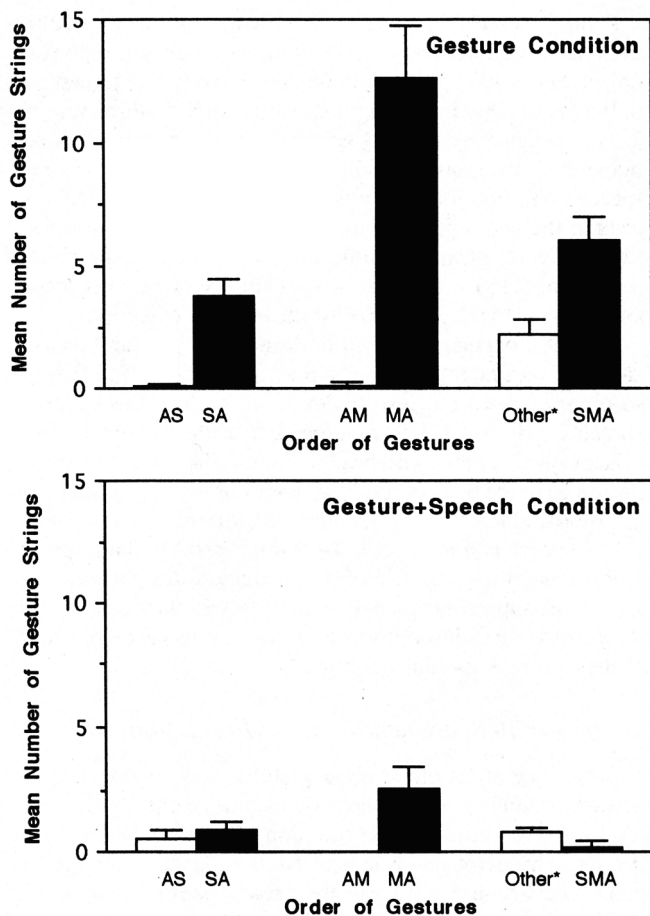
Figure 3. The mean number of strings produced in the gesture condition (upper graph) and the gesture + speech condition (lower graph), categorized according to the order of elements produced. Participants were more likely to produce their gestures in a consistent order in the gesture condition than in the gesture + speech condition. Moreover, the consistent gesture orders found in the gesture condition did not conform to canonical word order in English. S = stationary object; M = moving object; A = action. The "Other" category, which is marked by an asterisk in both graphs, includes the other five orders in which these three elements could have been produced (i.e., MSA, SAM, MAS, ASM, and AMS). The error bars reflect the standard error.

ing particular elements categorized according to the order of those elements for participants in the gesture condition (upper graph) and for participants in the gesture + speech condition (lower graph).

Looking first at the gesture condition, we found that participants produced more strings following the moving object–action (MA) order than following the action–moving object (AM) order (12.67 vs. 0.13), $t(14) = 5.94, p < .0001$ (center bars of the upper graph in Figure 3). Indeed, 13 of the 15 participants who produced strings containing actions and moving objects produced some strings with an MA order and none with an AM order; 1 participant produced 10 strings with an MA order and only 1 with an AM order. The last of the 15 participants produced one exemplar in each order. These data suggest that participants were following an all-or-none pattern

when ordering the elements of their gesture strings rather than a probabilistic one. Such all-or-none patterns are characteristic of natural language.

Participants also produced more strings following the stationary object–action (SA) order than following the action–stationary object (AS) order (3.73 vs. 0.07), $t(14) = 4.99, p < .0001$ (bars on the left side of the upper graph in Figure 3). Fourteen of the 15 participants who produced strings containing actions and stationary objects produced some strings with an SA order and none with an AS order; the last participant produced one string with an AS order and none with an SA order. Again, participants appeared to follow an all-or-none rather than a probabilistic pattern.

Finally, participants produced more strings following the stationary object–moving object–action (SMA) order than strings following any of the other five possible orders of these three elements (6.08 vs. 2.17), $t(11) = 3.33, p = .007$ (bars on the right side of the upper graph in Figure 3). Eleven of the 12 participants who produced strings containing actions, moving objects, and stationary objects produced more strings with an SMA order than with any of the other five possible orders; in fact, 6 produced no strings or no more than one string in any of the other five orders. The last of the 12 participants produced five strings with an SMA order but seven with an MSA order. Thus, when participants produced strings of gestures in the gesture condition, they produced those gestures in particular orders according to the semantic element represented by each gesture.

Note that one of the gesture orders followed by participants in the gesture condition conformed to an English pattern—the MA order that is reminiscent of what would be a subject–verb order in English (e.g., "the *girl jumps* into the hoop"). Given this fact, it may not be surprising that all of the few strings containing actions and moving objects that participants produced in the gesture + speech condition (which were produced along with spoken English) also conformed to the English-like MA order rather than the non-English AM order (2.56 vs. 0.00), $t(8) = 3.00, p = .017$ (see the center bars of the lower graph in Figure 3).

However, the other two orders that participants produced in the gesture condition (i.e., SA: "hoop jump," and SMA: "hoop girl jump") do not conform to a canonical order in English. Perhaps not unrelatedly, participants produced very few strings containing these combinations of elements in the gesture + speech condition, and the few they did produce were not ordered as in the gesture condition. In the gesture + speech condition, participants produced a small number of strings containing actions and stationary objects (cf. Figure 2) and, in total, 7 participants produced six strings conforming to the SA order and four conforming to the AS order (bars on the left side of the lower graph in Figure 3). They also produced very few strings containing all three elements in the gesture + speech condition (cf. Figure 2) and, in total, 5 participants produced only one string in the SMA order and four in the MAS order (the order most like English, e.g., "the girl jumps into the hoop"; bars on the right side of the lower graph in Figure 3).

In summary, participants in the gesture + speech condition rarely concatenated their gestures into strings and, when they did, their strings conformed to a canonical order in English. In contrast, participants in the gesture condition frequently con-

catenated their gestures into strings, and the elements of those strings were produced in the following rule-governed order:

$$\text{gesture string} \rightarrow (S)(M)A,$$

where S = stationary object, M = moving object, A = action, and the parentheses indicate an optional element. This order does not conform to a canonical order in English. In addition, it is worth noting that this particular order distinguishes object-referring gestures (akin to nouns) from action-referring gestures (akin to verbs)—actions occurred at the end of the string, and stationary and moving objects consistently preceded the action. A distinction between nouns and verbs is one of the few that has traditionally been accepted as a linguistic universal (e.g., Robins, 1952; Sapir, 1921) and whose status as a universal continues to be uncontested (e.g., Givon, 1979; Hawkins, 1988; Hopper & Thompson, 1984, 1988; Schachter, 1985; Thompson, 1988). It is a distinction that Sapir (1921, p. 119) considered to be essential to the "life of language" and thus one that might be expected to appear in the gestures of participants in the gesture condition, if indeed they were using their gestures in a languagelike manner.

Incorporating a moving object or stationary object into the action. Besides producing a separate lexical gesture (i.e., lexicalization, a syntactic device based on regularities across gestures within a string), the participants used one other device to represent the moving or stationary object in a gesture string— they incorporated a hand shape that captured aspects of the object into the gesture for the action. This device is reminiscent of morphological structure, based on regularities within the gesture itself. For example, to describe a circle moving diagonally across the video screen, several participants produced a diagonal movement across space (an action) using a circle hand shape rather than an index finger, thus "incorporating" the shape of the moving object into the action gesture (the action could either occur alone or conjoined with other gestures in a string). Participants were more likely to incorporate information about an object into the action in the gesture condition than in the gesture + speech condition; 84% of the objects that could be mentioned in the gesture condition were conveyed by incorporation, compared with 35% in the gesture + speech condition, $t(15) = 5.99, p < .0001$.

Moreover, in the gesture condition, participants appeared to use the device of incorporation into actions, in conjunction with the lexicalization device, to ensure that all of the semantic elements that could be conveyed in a string were in fact conveyed. Limiting our analysis to responses containing action gestures, Figure 4 shows the proportion of action gestures describing one-object scenes in which the moving object was explicitly conveyed in gesture for both the gesture + speech and the gesture conditions. In this figure, an object could be conveyed either by a separate lexical gesture produced in a string along with the action gesture (lexicalization) or by incorporating hand shape information about the object into the action gesture (incorporation). The figure also presents, for the two conditions, the proportion of action gestures describing two-object scenes in which both the moving object and the stationary object were explicitly conveyed in gesture (either by lexicalization or by incorporation).

Figure 4 reveals that, using either incorporation or lexicalization, participants conveyed 99% of the single objects in the one-object scenes and 95% of both objects in the two-objects scenes in the gesture condition. In other words, almost all relevant objects were conveyed in every gesture response produced by participants in the gesture condition. In contrast, in the gesture + speech condition, participants conveyed 52% of the single objects in the one-object scenes and only 12% of both objects in the two-object scenes, significantly fewer than in the gesture condition, $t(15) = 5.16, p < .0001$ for the one-object scenes, and $t(15) = 16.02, p < .0001$ for the two-object scenes.

At times, participants used both incorporation and lexicalization to redundantly convey a particular object, and they did so more often in the gesture condition than in the gesture + speech condition. Thirty-nine percent of the moving and stationary objects that could be conveyed in the gesture responses produced in the gesture condition were conveyed by both incorporation and lexicalization, compared with only 7% in the gesture + speech condition, $t(15) = 5.00, p < .0001$. Thus, participants tended to make use of redundancy, an important property in all natural languages, in the gestures they produced in the gesture condition but rarely in the gestures they produced in the gesture + speech condition.[5]

Gesture With Speech and Without It: A Summary

Before beginning our study of gesture produced alone versus gesture produced with speech, we examined the contrast in a naturalistic communication situation. We found that deaf children who have not yet been exposed to a conventional sign language and who use gesture as their sole means of communication develop a system of gestures characterized by hierarchical levels of organization, with structure both *across* and *within* gestures. In contrast, the hearing mothers of these deaf children, who also use gesture but always in conjunction with speech, do not develop such structures; their gestures represent events globally and in a noncombinatoric fashion.

We then tested the robustness of this contrast by studying the gestures created by naive adults responding to videotaped stimuli under experimental conditions. The gestures of the hearing adults in our study were found to resemble the gestures of the deaf children when the adults were instructed to use only gesture in their responses (the gesture condition). In contrast, the gestures of the adults resembled the gestures of the deaf children's hearing mothers and other hearing adults when the adults were asked to respond with speech and no mention was made of gesture (the gesture + speech condition).

Specifically, in the gesture condition, the adults frequently combined their gestures into strings and those strings were reliably ordered, with gestures for certain semantic elements occurring in particular positions in the string (i.e., structure across gestures at the sentence level). In addition, the verblike action gestures the adults produced in the gesture condition

---

[5] Note, however, that the information conveyed in the gestures produced in the gesture + speech condition could be (and frequently was) redundant with information expressed in speech; that is, in the gesture + speech condition, redundancy could be found either within speech or across gesture and speech, but not within gesture.
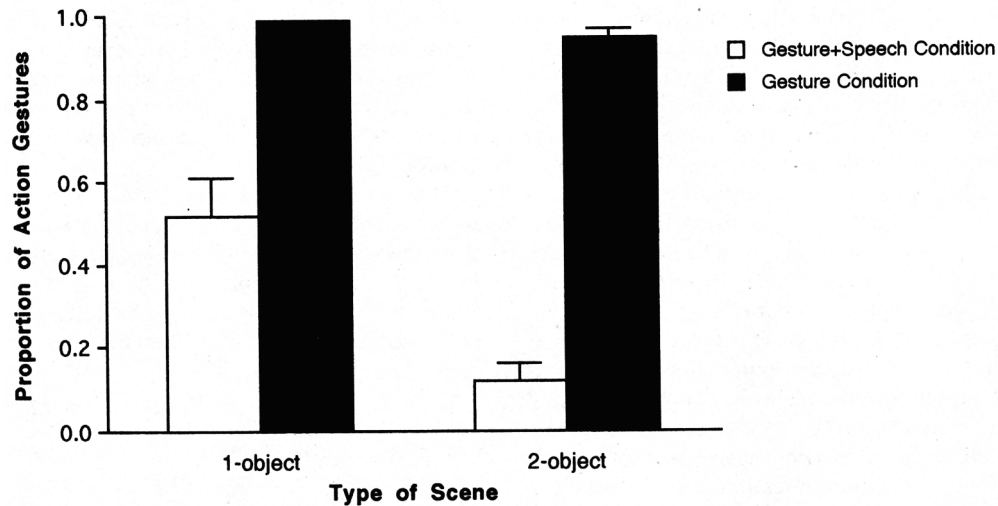
Figure 4. The mean proportion of action gestures for which relevant objects were explicitly conveyed in gesture in the gesture + speech condition and in the gesture condition. Relevant objects included both the moving object and the stationary object for the actions describing two-object scenes (i.e., two relevant objects) but only the moving object for the actions describing one-object scenes (i.e., one relevant object). An object was considered to be conveyed in gesture if a separate gesture for the object was produced within the gesture string containing the action gesture (lexicalization) or if hand shape information representing the object was incorporated into the action gesture (incorporation). Participants consistently conveyed all of the relevant objects within a gesture string (either by lexicalization or by incorporation) in the gesture condition but not in the gesture + speech condition. The error bars reflect the standard error.

could be said to be divisible into hand shape and motion parts, with the hand shape of the action frequently conveying information about the objects in its semantic frame (i.e., structure within the gesture at the word level). By using the two devices they developed for conveying the objects in the scenes—lexicalization (in which a separate gesture is produced to convey the object, akin to a syntactic device) and incorporation (in which a part of the action gesture, the hand shape, is used to convey the object, akin to a morphological device)—the adults conveyed essentially all of the relevant objects in the scenes in the gesture condition. Indeed, fairly often in this condition, they conveyed an object redundantly, using both lexicalization and incorporation.

In contrast, although the adults did produce gestures for actions and objects in the gesture + speech condition—and did so almost as often as they did in the gesture condition (cf. Figure 1)—they rarely combined the gestures for these elements into strings; 82% of the elements they conveyed were conveyed as single gestures. In other words, they did not use their gestures as building blocks for a larger unit. In addition, when they did use strings in the gesture + speech condition, the adults produced the gestures in those strings in an order reminiscent of canonical English order. They did not order their gestures in non-English patterns—patterns that were found in the gesture orders of the strings they produced in the gesture condition. Finally, the adults used incorporation to convey objects less often in the gesture + speech condition than in the gesture condition, almost never using incorporation redundantly with lexicalization to convey an object, and rarely using the two devices in concert to convey both objects in the two-object scenes. Thus, the parts of their gestures were not used systematically with other aspects of

the gestures to convey information, suggesting that these parts may not have functioned as independent units in this condition.

In summary, the gestures produced by the hearing adults in the gesture condition of our study and those produced by the deaf children in Goldin-Meadow's studies were characterized by the properties of segmentation and hierarchical combination. In contrast, the gestures produced by the hearing adults in the gesture + speech condition of our study and those produced by the hearing mothers of the deaf children along with their speech were not. We explore the properties of segmentation and combination in the next section, focusing on the conditions that foster their appearance.

The Resilience of Grammatical Properties: Segmentation and Combination in Symbolic Human Communication

The Resilience of Segmentation and Hierarchical Combination

In previous work, Goldin-Meadow (1982) identified segmentation and hierarchical combination as "resilient properties" of linguistic systems—properties that appear even when the conditions of language acquisition vary greatly from the conditions children typically experience. Segmentation and combination are found in a child's communications even when the child is lacking a conventional language model (Goldin-Meadow & Mylander, 1990a), when the language model that the child does have is degenerate (Singleton & Newport, 1992), or when the child is first exposed to a language model after puberty, well beyond the age language acquisition typically begins (Curtiss,

1977; Goldin-Meadow, 1978; Newport, 1990). These properties thus appear to be overdetermined in child communication, with a variety of trajectories leading to their development.

The findings of the present study indicate that segmentation and hierarchical combination also appear in the gestures created by hearing adults asked to abandon speech and use only their hands to communicate. The appearance of these properties in the hearing adults' gestures is particularly striking because they are not properties typically found in the gestures hearing adults use on a daily basis. Even more impressive is the fact that these properties appeared *immediately* in the adults' gestures. With essentially no time for reflection on what might be fundamental to languagelike communication, when asked to use only their hands to communicate, the adults in our study produced gestures characterized by segmentation and hierarchical combination. These properties thus appear to be resilient in all symbolic human communication, and not just child systems.

What is the source of the linguistic properties that appear in the gestures used by the adults in the gesture condition (but not the gesture + speech condition) in our experiment and in the gestures used by the deaf children in Goldin-Meadow's studies? The modality itself obviously cannot explain these properties because the modality is held constant across all of these situations, even the gesture + speech condition. In addition, transfers from speech (i.e., modeling the gestures directly after properties in speech) do not explain the properties, at least not all of them, because the ordering pattern used by the adults in the gesture condition did not conform to a canonical English order and thus could not have been readily derived from English. Moreover, the deaf children of hearing parents have no usable spoken language source for their gesture systems, but still develop the properties of segmentation and hierarchical combination in those systems (see Footnote 4).

Are Segmentation and Hierarchical Combination Imposed by the Structure of Thought?

One possibility is that the properties of segmentation and hierarchical combination reflect the nature of human thought. In other words, when asked to describe a scene, humans might think of the scene in terms of discrete entities arranged in hierarchical combination and, as a result, would be forced to describe that scene in this form and this form only.

We think this possibility unlikely simply because we know from the gestures that people spontaneously produce along with their spoken narratives that human thought need not be segmented and hierarchically arranged (McNeill, 1992). As described earlier, the gestures that adults and children spontaneously produce along with their speech express ideas but express them in a different form from speech—an unsegmented form in which a single gesture can convey an entire, global proposition. For example, consider how a speaker might describe the spatial relationship of Scotland to Ireland in words and in gesture. The verbal description will contain a series of words syntactically organized and is likely to mention each landmass in its own phrase or clause. In contrast, the gesture is likely to be a single, global display of two spaces in a certain relationship, capturing in nonsegmented form the information that is in

speech and possibly going beyond speech in showing the relative positions in space. Indeed, it is because spontaneous gesture takes on a very different form from speech that McNeill (1992) proposed it as a unique window through which the thoughts of a speaker can be observed (see also Goldin-Meadow, Alibali, & Church, 1993).

Thus, humans can and do think about a scene as a global whole, as reflected in the spontaneous gestures that accompany their speech. The structure of thought cannot, then, be solely responsible for the segmented and combinatorial structure seen in the gestures of the hearing adults in the gesture condition or in the gestures of the deaf children in Goldin-Meadow's studies.

The Drive Toward Segmentation: Establishing Reference

The task that faced the adults in both conditions of our experiment was to describe a simple scene to another person. The first step in such a task is to make it clear what the scene is about, that is, to establish reference. This step was easy to accomplish in the gesture + speech condition; participants used speech to focus attention on the objects in the scene, either by introducing the objects in separate clauses ("There's a toy girl with a circle standing next to her; she jumps into it") or by referring to the objects with separate nouns ("A toy girl jumps into a circle"). A gesture might accompany the speech (e.g., the participant might jump his hand, loosely held, from one point to another as he says, "a toy girl jumps into a circle"). Note that the gesture can convey mimetic aspects of the scene that are not conveyed in speech (e.g., how high the girl jumped, how softly she landed, etc.). The information conveyed by gesture is interpretable to the listener primarily because the gesture is framed by the speech that it accompanies; that is, speech supplies a socially constituted form of the same or related content that is provided in imagistic form by the gesture.

Imagine now how the same participant might establish reference in the gesture condition. He could produce the same gesture that he produced in the gesture + speech condition—jumping his loosely held hand from one point to another—but such a gesture, no longer framed by speech, would not indicate who is doing the jumping, nor would it indicate the end point of the jump. To meet these requirements, participants in the gesture condition could alter the gestures they produced in the gesture + speech condition in one (or both) of two ways: (a) They could produce a series of gestures to refer to the scene (e.g., they might form curls on their own heads to indicate the girl, hold a C-shaped hand to indicate the circle, and then produce the "jump" gesture). (b) They could add elements to the gesture itself that serve to identify the objects (e.g., an inverted-V hand shape indicating the legs of the girl might be "jumped" into a C-shaped hand indicating the circle). Interestingly, both of these strategies have been identified in a study similar to our own (Dufour, 1992) in which hearing adults were asked to use gesture rather than speech to convey a story.[6]

Thus, the requirement that reference be established appears

---

[6] Dufour did not include in his study a condition in which the participant was asked to tell the story in speech with no mention made of gesture; as a result, the design of his study did not permit him to determine whether participants' gestures differed with and without speech.

to lead naturally to a process of segmenting the scene, breaking it up so that each element is identified. Note that the process of establishing reference through symbolic means is inherently social—the goal is not just to reconstruct the scene in one's own mind but to reconstruct it in the mind of another. Freyd (1983) suggested that the very act of sharing information across minds will force that information to take a discrete and segmented form. Freyd argued that even if knowledge is truly represented in a continuous form, in the process of sharing that knowledge with another individual, the knowledge must go through a discrete filter, and, as a result, its representation ends up looking discrete.

The Drive Toward Combination: Relating Referents to One Another

Once a scene has been segmented into elements, the elements need to be brought together into a whole in order for the scene to be adequately described. If left uncombined, the gestures do serve to refer to the elements in the scene but they do not convey how those elements function together within the scene. To do so, the gestures for the elements need to be brought together in such a way that the relationship among them is clear. One strategy for accomplishing this goal is to conjoin the gestures for the elements in a scene into a single gesture string and to position each gesture within that string according to the role played by the element the gesture represents. This is, in fact, the strategy followed by participants in the gesture condition, who tended to produce gestures for the elements in a scene within a single gesture string (i.e., without breaking the flow of movement between the gestures). Moreover, participants in this condition positioned the gestures within the string in a systematic fashion: The gesture for the stationary object preceded the gesture for the moving object, which, in turn, preceded the gesture for the action. This particular order, in fact, follows an intuitively natural progression in terms of laying out the scene for the listener—it first sets the scene for the listener (with the stationary object), then introduces the focus or topic of the scene (the moving object), and finally comments on that topic (the action). Interestingly, this particular order is one that is routinely found in ASL for scenes of this type (T. Supalla, personal communication, January 20, 1994) and is reminiscent of an order found in a gesture system developed by a woman who was the only deaf person in a hearing Amerindian reservation (Yau, 1985). The order thus appears to be one that is naturally exploited in manual communication systems—in conventional sign languages handed down from generation to generation as well as in spontaneous gesture systems created by individuals.

In summary, the communicative requirement to make clear how the elements in a scene relate to one another appears to lead naturally to the process of systematic combination. In this regard, it is important to recall that, in the gesture + speech condition, participants did produce gestures for the elements. Indeed, a look at Figure 1 shows that participants conveyed elements in this condition almost as frequently as in the gesture condition (although the intent may have been to convey a global rather than a segmented meaning).[7] The crucial difference is that, in the gesture + speech condition, gestures referring to elements were each produced as a separate unit (i.e., there was

a break in the flow of movement between the gestures), whereas in the gesture condition the gestures were conjoined within a single string (with no break in the movement). Thus, the difference between the two conditions was not primarily a matter of referring to the elements but of referring to them as *elements within a proposition*. In the gesture condition, this task is assumed entirely by gesture but, in the gesture + speech condition, the task of recombining elements to relate them to one another is naturally taken over by speech.

We now gain insight into how silence "liberates" grammatical properties in the manual modality. Segmented and combinatorial representation is the form that symbolic human communication must take. Speech performs these required functions in normal discourse, with gesture forming an integrated system with that speech and conveying information in a global and mimetic form. The need for segmented and combinatorial representation is not created by speech but rather by the demands of symbolic human communication. When speech is removed, the need for segmented and combinatorial representation remains, and the requirement is filled, for both the deaf children in Goldin-Meadow's studies and the adults in our study, by gesture. Thus, when gesture alone carries the full burden of communication, it no longer has the global and largely mimetic form that it is constrained to assume when it forms an integrated system with speech—it must now take over the

<hr>

[7] The fact that participants generated gestures for elements within a scene almost as often in the gesture + speech condition as in the gesture condition suggests that segmentation may occur in gesture even when it does not carry the full burden of communication. However, it is important to point out that segmentation is not independent of combination. It is only when two gestures are combined within the bounds of a single string that there is good evidence that these gestures function as a unit, with each gesture reflecting a piece of the proposition—that is, that the gestures truly represent segmented elements within a proposition. For example, a gesture coded as representing a rolling movement could either stand for the semantic element "rolling" or for the whole "ball-rolling" proposition. Imagine that a speaker produces this rolling gesture concatenated with another gesture resembling the shape of the ball. Because the two gestures form a unit and contrast with one another within that unit, we assume that each gesture represents a piece of the proposition, with the rolling gesture representing the act ("roll") and the ball gesture representing the moving object ("ball"). In contrast, if the rolling gesture is produced on its own and not concatenated with any others, the gesture could easily be assumed to stand for "ball rolling" as a whole.

Because the gestures in the gesture + speech condition were almost always produced on their own (i.e., not combined and therefore not contrasted with other gestures), there was no reason to assume that these gestures were segmented elements. Rather than semantic pieces of a scene, the gestures in the gesture + speech condition appear to be snapshots of full propositions taken from different angles—the angle of the object, the angle of the action, and so forth—presented along with speech. As snapshots of the whole, the gestures in the gesture + speech condition did not need to be combined with other gestures to convey the proposition—each gesture did so on its own. Thus, although in the gesture condition there was reason (specifically, the frequent concatenation of gestures into strings) to assume that each gesture referred to a particular element and that segmentation had taken place, there was no reason to make this assumption in the gesture + speech condition, and indeed some reason not to make it in this condition (the lack of concatenation despite gestures apparently referring to parts of the scene).

essential role and assume the segmented and combinatorial form required for symbolic human communication. By breaking the bond between gesture and speech, silence frees gesture from the constraints imposed by being part of an integrated system with speech—only to force gesture into the constraints imposed by carrying the full burden of communication, that is, the constraints of a languagelike system.

Grammatical Properties Beyond Segmentation and Combination

Internal Consistency: Going Beyond Segmentation and Combination to Form a Linguistic System

The adults in the gesture condition produced gestures that had the properties of segmentation and combination in common with the gestures produced by the deaf children in Goldin-Meadow's studies. However, the gestures that the adults produced did not share all of the languagelike properties found in the deaf children's gesture systems. In particular, the adults' gestures were not systematically organized into a system of internal contrasts, whereas the deaf children's gestures were. Thus, unlike the deaf children, the adults displayed hierarchical but not necessarily consistent use of constituents in their gestures. Singleton, Morford, and Goldin-Meadow (1993) gave the VMP test to one of the deaf children in Goldin-Meadow's sample. The child described each scene in the test using his own gesture system, and his gestures were compared with those produced by the 16 hearing adults in the gesture condition in this study and with gestures produced by 8 hearing children. Singleton et al. found that, when incorporating hand shape information into their action gestures, the hearing adults and children rarely used the same hand shape to represent an object each time it occurred on the test, whereas the deaf child frequently did. In other words, there was consistency across hand shapes in this deaf child's gestures (as well as in the gestures of each of the 3 other deaf children shown to have morphological structure in their spontaneous communications; Goldin-Meadow et al., 1995), but not in the gestures of each hearing adult or child. Thus, there was evidence for a system of contrasts in the deaf children's gestures but not in the gestures of the hearing adults and children.

Singleton et al. (1993) suggested that the hand shape differences between the deaf child and the hearing gesturers reflect a fundamental difference in the way participants generated gestures. When the hearing gesturers generated a gesture, their goal was to produce a hand shape that adequately represented the object, and their choice of hand shapes appeared to be constrained only by their imaginations and the physical limitations imposed by the hands themselves. For example, one of the child-hearing gesturers produced a different hand shape each of the five times she represented an airplane on the test. Each hand shape captured an idiosyncratic property (often the differently shaped wings) of the airplane pictured in that event: (a) thumb, index, and middle fingers extended; (b) two palms forming a V; (c) two palms crossed in an X; (d) thumb and pinky extended; and (e) O-hand used to represent the way the airplane, made of paper, might have been thrown (the airplane was not actually thrown in the event). In contrast, when the deaf children gener-

ated a gesture, their choice of hand shapes was guided not only by how well a hand shape captured the features of an object, but also by how well that hand shape fit into the set of hand shapes allowed in their individual gesture systems. The deaf child who took the VMP test used two different hand shapes on the five airplane segments—a flat-palm hand shape and a thumb-and-pinky-extended hand shape—both of which formed part of this child's morphological system and could be used to represent airplanes in that system (Singleton et al., 1993).

Thus, the gestures that the adults produced in the gesture condition were organized only in relation to their referents; that is, the form of the gesture was constrained by the object or action it was used to represent. In contrast, the deaf children's gestures, in addition to being constrained by form-to-referent relationships, were also organized in relation to one another; that is, the form of the gesture was constrained not only by what it represented but also by how it contrasted with the other gesture forms in the system—form-to-form relationships. The deaf children's gestures can therefore be said to have the crucial aspect of linguistic systems that de Saussure (1916/1959) called linguistic "value"—an aspect that appears to be lacking in the adults' gestures. According to de Saussure, a language is a system in which all the elements fit together and in which the value of any one element depends on the simultaneous existence of all of the others. The fact that linguistic value is a characteristic of the deaf children's gesture systems but not the adults' gesture systems suggests that the deaf children—but not the hearing adults—had begun to consider their gestures as a "problem space" (Karmiloff-Smith, 1979) in their own right.

Conditions Fostering a System of Contrasts

Why might the deaf children go beyond form-to-referent relations to establish form-to-form relations in their gesture systems, whereas the adults failed to do so? One possibility is that the child gesturers were children and therefore more gifted at language creation than the adults (as a critical period hypothesis—cf. Lenneberg, 1967—might predict). Another possibility is that the deaf children developed their gesture systems over a period of years, whereas the adults generated their gestures "on the spot" (see Singleton et al., 1993, for further discussion of this point). Indeed, there is evidence that the deaf children first created gestures with an eye to how well the gesture captured aspects of the intended referent. Only later in development did the children provide evidence that they created their gestures, not only to be good representations, but also to contrast meaningfully with the other gestures in their systems (Goldin-Meadow & Mylander, 1990b; Goldin-Meadow et al., 1995).[8] Thus, it may be necessary first to create a set of gestures that are adequate representations of objects and actions. The next step, which presumably requires time and experience with the gestures, may be to pull back and consider the set of gestures as a problem space, isolating the components that recur across the

---

[8] It is important to point out that throughout this developmental period, the deaf children tended to use the same gesture to represent a given referent; that is, their gestures formed a stable lexicon throughout the period during which these other developmental changes were occurring (see Goldin-Meadow, Butcher, Mylander, & Dodge, 1994).

gestures and organizing the set around those recurring components. It is this second step that results in the de Saussurian property of linguistic value and that the adult participants in our study had not yet taken.

Is there evidence that the deaf children considered their own gestures as a problem space in its own right? Although it is difficult to obtain direct evidence, there are indications that the deaf children were cognizant of their gestures as a symbolic system. For example, one of the deaf children used gesture to refer to his own gestures. To request a Donald Duck toy that the experimenter held behind her back, the child pursed his lips to imitate Donald Duck's bill and then pointed at his own gesture, the pursed lips (Goldin-Meadow, 1993). Thus, the child was able to use the pointing gesture metalinguistically, suggesting that he could distance himself from his own gestures and treat them as objects to be reflected on and referred to. As a second example, this child was observed to "correct" the gestures his hearing sister produced when she took the VMP test (Singleton et al., 1993). When the sister used a hand shape that, in the deaf child's system, was not an acceptable representation for the object in a particular vignette, the deaf child first mocked his sister's hand shape choice and then showed her the hand shape that did conform to his system. The child thus had a well-developed and articulated sense of what counts as an acceptable gesture in his system. He not only produced gestures that adhered to his standards, but he also imposed his standards on the gestures of another. The distance that this deaf child achieved from his gesture system may not be possible in the short-term creation situation that the adults in our study experienced. Moreover, it may be just this type of distance that is essential for a speaker to introduce the important de Saussurian characteristic of linguistic value into a communication system.

Grammatical Properties and the Conditions of Language Creation

In summary, the emergence of segmentation and combination in the experimental paradigm that we used in our study highlights the resilience of these grammatical properties in symbolic human communication.[9] With no time for reflection, the adults in our study constructed a set of gestures characterized by segmentation and combination. However, our simple experimental paradigm was not sufficient to support the emergence of other grammatical properties thought equally central to human language—in particular, a system of contrasts. These findings suggest that it may be possible to classify grammatical properties according to the types of conditions that support their creation. By altering aspects of our experimental paradigm, we may be able to explore the effects of various circumstances on the emergence of grammatical properties.

For example, using the experimental paradigm as is, we discovered that a system of contrasts in which the form of a symbol is constrained by its relationship to other symbols in the system (as well as by its relationship to its intended referent) is not an immediate consequence of symbolically communicating information to another human. The data on the deaf children suggest that continued experience with a set of gestures may be required for a system of contrasts to emerge in those gestures. However, it is also possible that even with continued experience with their

own gestures, adults may not be able to develop a system of contrasts within those gestures. In other words, it may be that the creator must be a child in order for a system of contrasts to emerge. Our experimental paradigm can be adapted to address this issue. Adults and children may be asked to repeat the procedure in the gesture condition over an extended period, thus allowing each creator the time that appears to be necessary (but may not be sufficient) to develop a system of contrasts.

Alternatively, it may be that what is crucial for a system of contrasts to evolve in a symbolic communication system has more to do with the recipient than the creator. Because the adults in our experiment received no feedback from their recipient (who was the experimenter) as to how well they communicated the characteristics of each vignette, the grammatical properties that appeared in their gestures had to reflect the creators' intuitive sense of what their recipient would find comprehensible—not what the recipient actually understood. To explore the role of the recipient, both gesturer and receiver must be participants in the study. Our experimental paradigm can be modified to include a naive recipient, and the nature (e.g., child vs. adult) or number (e.g., one vs. many) of recipients can be systematically varied. For example, an adult asked to communicate with the same recipient might, over time, generate a set of arbitrary symbols; these symbols may or may not form a system of contrasts. In contrast, an adult asked to communicate with a variety of recipients might, over the same period, generate a relatively iconic set of gestures but one that becomes systematized in response to the variability among receivers.

Along these lines, it is important to note that, although both the deaf children in Goldin-Meadow's study and the adults in our study used their gestures to communicate with others, neither of these gesture systems can be considered a shared system. Both the hearing adults and the deaf children *produced* gestures that were received by others, but neither the adults nor the deaf children *received* their own gestures as input. The hearing adults received no gestural input when they participated in the experiment. The deaf children saw the spontaneous gestures of their hearing parents; however, the parents' gestures were structured like all gestures that accompany speech—they were global and mimetic, rarely combined with one another and showing no evidence of internal components (Goldin-Meadow & Mylander, 1984, 1990b). Thus, the gestures the deaf children received as input were distinct from the gestures they themselves produced as output. "Communication," as we have used the term, implies only that the gesturer intends his message to be received by another. It is, however, interesting to ask whether a communication system that is both produced and received will take on grammatical properties that are distinct from the properties found in a system that is only produced and not received. Our experimental paradigm can again be adapted to explore this issue. Two adults may be asked to alternate in producing and receiving gestural descriptions in the gesture condition, a condition that could well evolve toward a system that is truly shared, and one that may differ substantially from a system used only in production.

_____

[9] Casey and Kluender (in press) recently argued that segmentation and combination may have been intermediate morphosyntactic forms in the evolution of language; if so, these grammatical properties would be resilient in an evolutionary as well as a developmental sense.

As a final example, it is worth noting that our study is a single generation deep. Passing a conventional communication system down from generation to generation tends to alter the structure of that system, creating language change over historical time. This issue can be explored over a shorter period by again adapting our experimental paradigm along a different dimension. After two adults have developed a gesture system, they may be asked to share that system with a third person (an adult or perhaps a child) who is new to the task. The third person, experiencing the system for the first time, may introduce changes into the system, changes that may come about only when a novice views the system as a whole (cf. Kegl, 1994; Singleton & Newport, 1992).

Thus, our experimental paradigm may be adapted to probe language creation over the short-term, providing us with a technique to explore the effects of various environments on the structure of symbolic human communication. Note, however, that our gesture-creation paradigm, although useful for exploring aspects of language creation, does not simulate the conditions that existed when language was created for the first time. Nevertheless, our findings provide us with grounds for speculation about language evolution, and we turn to this topic in the final section.

Gesture and Linguistic Evolution

Signed languages (as well as our own gesture condition) make it clear that the manual modality can assume a segmented and combinatorial form. Why then did language become the province of the oral modality? Why is speech the most common form of linguistic behavior in human cultures when it could just as easily have been sign? We speculate that having segmented structure in the oral modality as we currently do leaves the manual modality free to co-occur with speech and to capture the mimetic aspects of communication *along with* speech. Thus, our current arrangement allows us to retain, along with a segmented representation and in a single stream of communication, the imagistic aspects of the mimetic that are so vital to communication (cf. McNeill, 1992).[10] Note that the alternative arrangement—in which the manual modality would serve languagelike functions and the oral modality would serve the mimetic functions—has the disadvantage of forcing the oral modality to be unnaturally imagistic in form (although see Haiman, 1985, for evidence that the oral modality does exhibit some iconic properties). As an example, Huttenlocher (1973, 1976) pointed out that a verbal description of the shape of the East Coast of the United States is likely not only to be very cumbersome, but also to leave out important information about the coastline—information that we suggest could easily be captured in a mimetic gesture tracing the outline of the coast. Because gesture allows one to represent an image as a whole without breaking it into parts, gesture offers a better vehicle for encoding imagistic information than does speech (Goldin-Meadow et al., 1993). The manual modality is therefore the natural choice to encode mimetic information, leaving information that is better captured in a discrete and segmented form to the oral modality. Under this scenario, speech became the predominant medium of human language not because it is so well suited to the linear and segmented requirements of sym-

bolic human communication (the manual modality is equally suited to the job), but rather because it is not particularly good at capturing the imagistic components of human communication (a task at which the manual modality excels).[11]

In summary, when both the manual and oral modalities are accessible, communication typically involves the hand as well as the mouth (this is so even in the congenitally blind, who have no model for gesturing; Iverson & Goldin-Meadow, 1995). Gesture and speech form an integrated system, with speech assuming a segmented and combinatorial form and gesture assuming a global and mimetic form. The question we asked in this study was what happens when speech is removed. Will gesture continue to assume the global and mimetic form that it typically assumes when produced along with speech, or will it take over the segmented and combinatorial form that is typically the province of speech? We found that gesture does not remain global and mimetic but rather immediately switches its form. Gesture assumes the segmented and combinatorial form of speech, although it does not take on all of the grammatical properties found in speech, nor even all of the grammatical properties found in gesture systems developed over a period of years by deaf children who use gesture as their sole means of communication (cf. Goldin-Meadow & Mylander, 1990a).

Our findings suggest that a segmented and combinatorial form is essential to symbolic human communication, so essential that, if prevented from coming out of the mouth, it will characterize what comes out of the hands. We speculate that the properties of segmentation and combination arise inevitably, not from how humans think, but from the task of symbolically communicating their thoughts to others—the need to establish reference and relate those referents to one another. It is not the communicative task alone that fosters these properties, because segmentation and combination do not appear in communication systems used by other animals. Rather, segmentation and combination appear to be inevitable conse-

---

[10] One might argue that the imagistic aspects of the mimetic are not so vital to symbolic human communication on the grounds that we do communicate when gesturing is not possible (e.g., on the telephone). However, as far as we know, no one has attempted to evaluate whether communication in such situations is as effective as face-to-face communication (indeed, many speakers find themselves gesturing on the telephone, presumably feeling a need to augment their talk in some perhaps mimetic way; cf. Rime, 1982). Moreover, there is much work suggesting that the gestures produced along with speech do indeed serve an essential role in communication (Goldin-Meadow, Wein, & Chang, 1992; Kendon, 1994; McNeill, Cassell, & McCullough, 1994).

[11] This speculation about the importance of maintaining a vehicle for mimetic representation along with speech raises an interesting question with respect to sign language. In sign, it is the manual modality that assumes the segmented and combinatorial form essential to human language. Can the manual modality at the same time also be used for global and mimetic expression? In other words, do signers gesture along with their signs and, if not, how is the global and mimetic function filled? One possibility is that sounds or mouth movements might assume the mimetic function for signers. Although such movements have frequently been observed in fluent signers, as far as we know, no work has been conducted to investigate whether these behaviors (or any others, for that matter) serve for sign the mimetic function that gesture serves for speech.

quences of the human mind grappling with the task of shared symbolic communication.

We end by emphasizing that, although the imagistic information that gesture is constrained to convey when it accompanies speech is an important part of human communication (cf., McNeill, 1992), it can be compromised when gesture is called on to carry the full burden of communication. When gesture is the only modality available, it is no longer purely driven by imagery but instead assumes the segmented and combinatorial form required for symbolic human communication. Thus, when the bond between gesture and speech is broken, gesture is liberated—no longer bound by the constraints imposed on it in the integrated system it shares with speech. However, this freedom is short-lived as, once liberated, gesture is constrained again—forced to take on the forms imposed by carrying the full burden of symbolic human communication.

References

Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture–speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology, 25,* 468–523.

Battison, R. (1974). Phonological deletion in American Sign Language. *Sign Language Studies, 5,* 1–19.

Bekken, K., Goldin-Meadow, S., & Dymkowski, T. (1990, October). *Dissociation of maternal speech and gesture to deaf children of hearing parents.* Paper presented at the 15th Annual Boston University Conference on Language Development, Boston.

Bellugi, U., & Studdert-Kennedy, M. (Eds.). (1980). *Signed and spoken language: Biological constraints on linguistic form.* Deerfield Beach, FL: Verlag Chemie.

Butcher, C., & Goldin-Meadow, S. (1995). *Gesture and the transition from one- to two-word speech: When hand and mouth come together.* Manuscript submitted for publication.

Caselli, M. C. (1983). Communication to language: Deaf children's and hearing children's development compared. *Sign Language Studies, 39,* 113–144.

Casey, S., & Kluender, R. (in press). Evidence for intermediate forms in the evolution of language. In A. Dainora, R. Hemphill, B. Luka, B. Need, & S. Pargman (Eds.), *Proceedings of the 31st Regional Meeting of the Chicago Linguistic Society.* Chicago: Chicago Linguistic Society.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23,* 43–71.

Cohen, E., Namir, L., & Schlesinger, I. M. (1977). *A new dictionary of sign language.* The Hague, Netherlands: Mouton.

Conrad, R. (1979). *The deaf child.* London: Harper & Row.

Coulter, G. R. (1990). Emphatic stress in ASL. In S. D. Fisher and P. Siple (Eds.), *Theoretical issues in sign language research: Vol. 1. Linguistics* (pp. 109–125). Chicago: University of Chicago Press.

Crowder, E. M., & Newman, D. (1993). Telling what they know: The role of gesture and language in children's science explanations. *Pragmatics and Cognition, 1,* 341–376.

Curtiss, S. (1977). *Genie: A psycholinguistic study of a modern-day "wild child."* New York: Academic Press.

DeMatteo, A. (1977). Visual imagery and visual analogues in American Sign Language. In L. Friedman (Ed.), *On the other hand: New perspectives on American Sign Language* (pp. 109–136). New York: Academic Press.

de Saussure, F. (1959). *Course in general linguistics* (W. Baskin, Trans.). New York: Philosophical Library. (Original work published 1916)

Dufour, R. (1992). *The use of gestures for communicative purposes: Can gestures become grammatical?* Unpublished doctoral dissertation, University of Illinois, Urbana-Champaign.

Ekman, P., & Friesen, W. (1969). The repertoire of nonverbal behavioral categories. *Semiotica, 1,* 49–98.

Evans, M. A., & Rubin, K. H. (1979). Hand gestures as a communicative mode in school-aged children. *The Journal of Genetic Psychology, 135,* 189–196.

Fant, L. J. (1972). *Ameslan: An introduction to American Sign Language.* Silver Spring, MD: National Association of the Deaf.

Feldman, H., Goldin-Meadow, S., & Gleitman, L. (1978). Beyond Herodotus: The creation of language by linguistically deprived deaf children. In A. Lock (Ed.), *Action, symbol, and gesture: The emergence of language* (pp. 351–414). New York: Academic Press.

Feyereisen, P., & de Lannoy, J.-D. (1991). *Gestures and speech: Psychological investigations.* Cambridge, England: Cambridge University Press.

Fischer, S. (1973). Two processes of reduplication in American Sign Language. *Foundations of Language, 9,* 469–480.

Fischer, S. (1974). Sign language and linguistic universals. In C. Rohrer & N. Ruwet (Eds.), *Actes du colloque Franco-Allemand de grammaire transformationelle: Vol. 2. Etudes de Sémantique et autres* [*Proceedings of the Franco-German conference on French transformational grammar*] (pp. 187–204). Tubingen, Germany: Max Niemeyer Verlag.

Fischer, S., & Gough, B. (1978). Verbs in American Sign Language. *Sign Language Studies, 18,* 17–48.

Freyd, J. J. (1983). Shareability: The social psychology of epistemology. *Cognitive Science, 7,* 191–210.

Friedman, L. A. (1976). The manifestation of subject, object, and topic in American Sign Language. In C. Li (Ed.), *Subject and topic* (pp. 125–148). New York: Academic Press.

Frishberg, N. (1975). Arbitrariness and iconicity: Historical change in American Sign Language. *Language, 51,* 696–719.

Givon, T. (1979). *On understanding grammar.* New York: Academic Press.

Goldin-Meadow, S. (1978). A study in human capacities [Review of the book *Genie: A psycholinguistic study of a modern-day "wild child"*]. *Science, 200,* 649–651.

Goldin-Meadow, S. (1979). Structure in a manual communication system developed without a conventional language model: Language without a helping hand. In H. Whitaker & H. A. Whitaker (Eds.), *Studies in neurolinguistics* (Vol. 4, pp. 125–209). New York: Academic Press.

Goldin-Meadow, S. (1982). The resilience of recursion: A study of a communication system developed without a conventional language model. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 51–77). New York: Cambridge University Press.

Goldin-Meadow, S. (1993). When does gesture become language? A study of gesture used as a primary communication system by deaf children of hearing parents. In K. R. Gibson & T. Ingold (Eds.), *Tools, language and cognition in human evolution* (pp. 63–85). New York: Cambridge University Press.

Goldin-Meadow, S., Alibali, M. W., & Church, R. B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review, 100,* 279–297.

Goldin-Meadow, S., Butcher, C., Mylander, C., & Dodge, M. (1994). Nouns and verbs in a self-styled gesture system: What's in a name? *Cognitive Psychology, 27,* 259–319.

Goldin-Meadow, S., & Feldman, H. (1977). The development of language-like communication without a language model. *Science, 197,* 401–403.

Goldin-Meadow, S., & Mylander, C. (1983). Gestural communication

in deaf children: The non-effects of parental input on language development. *Science, 221*, 372–374.

Goldin-Meadow, S., & Mylander, C. (1984). Gestural communication in deaf children: The effects and non-effects of parental input on early language development. *Monographs of the Society for Research in Child Development, 49*, (3–4, Serial No. 207).

Goldin-Meadow, S., & Mylander, C. (1990a). Beyond the input given: The child's role in the acquisition of language. *Language, 66*, 323–355.

Goldin-Meadow, S., & Mylander, C. (1990b). The role of parental input in the development of a morphological system. *Journal of Child Language, 17*, 527–563.

Goldin-Meadow, S., Mylander, C., & Butcher, C. (1995). The resilience of combinatorial structure at the word level: Morphology in self-styled gesture systems. *Cognition, 56*, 195–262.

Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing knowledge through gesture: Using children's hands to read their minds. *Cognition and Instruction, 9*, 201–219.

Haiman, J. (1985). *Iconicity in syntax.* Amsterdam: John Benjamins.

Hawkins, J. A. (1988). Explaining language universals. In J. A. Hawkins (Ed.), *Explaining language universals* (pp. 3–28). Cambridge, MA: Basil Blackwell.

Hoffmeister, R. (1978). *The development of demonstrative pronouns, locatives and personal pronouns in the acquisition of American Sign Language by deaf children of deaf parents.* Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

Hoffmeister, R., & Wilbur, R. (1980). Developmental: The acquisition of sign language. In H. Lane & F. Grosjean (Eds.), *Recent perspectives on American Sign Language* (pp. 61–78). Hillsdale, NJ: Erlbaum.

Hopper, P. J., & Thompson, S. A. (1984). The iconicity of the universal categories "noun" and "verbs." In J. Haiman (Ed.), *Iconicity in syntax* (pp. 151–183). Philadelphia: John Benjamins.

Hopper, P. J., & Thompson, S. A. (1988). The discourse basis for lexical categories in universal grammar. *Language, 60*, 703–752.

Huttenlocher, J. (1973). Language and thought. In G. A. Miller (Ed.), *Communication, language and meaning: Psychological perspectives* (pp. 172–184). New York: Basic Books.

Huttenlocher, J. (1976). Language and intelligence. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 261–281). Hillsdale, NJ: Erlbaum.

Iverson, J., & Goldin-Meadow, S. (in press). *What's communication got to do with it: Gesture in blind children. Developmental Psychology.*

Jancovic, M. A., Devoe, S., & Wiener, M. (1975). Age-related changes in hand and arm movements as nonverbal communication: Some conceptualizations and an empirical exploration. *Child Development, 46*, 922–928.

Kantor, R. (1982). Communicative interaction: Mother modification and child acquisition of American Sign Language. *Sign Language Studies, 36*, 233–282.

Karmiloff-Smith, A. (1979). *A functional approach to child language: A study of determiners and reference.* Cambridge, England: Cambridge University Press.

Kegl, J. (1994). The Nicaraguan Sign Language Project: An overview. *Signpost, 7*, 24–31.

Kendon, A. (1972). Some relationships between body motion and speech. In A. Siegman & B. Pope (Ed.), *Studies in dyadic communication* (pp. 177–210). New York: Pergamon Press.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *Relationship of the verbal and nonverbal communication* (pp. 207–228). The Hague, Netherlands: Mouton.

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

Kita, S. (1993). *Language and thought interface: A study of spontaneous*

gestures and Japanese mimetics. Unpublished doctoral dissertation, University of Chicago.

Klima, E., & Bellugi, U. (1979). *The signs of language.* Cambridge, MA: Harvard 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.

Lane, H., Boyes-Braem, P., & Bellugi, U. (1976). Preliminaries to a distinctive feature analysis of handshapes in American Sign Language. *Cognitive Psychology, 8*, 263–289.

Lane, H., & Grosjean, F. (1980). *Recent perspectives on American Sign Language.* Hillsdale, NJ: Erlbaum.

Lenneberg, E. H. (1964). Capacity for language acquisition. In J. A. Fodor & J. J. Katz (Eds.), *The structure of language: Readings in the philosophy of language* (pp. 579–603). Englewood Cliffs, NJ: Prentice Hall.

Lenneberg, E. H. (1967). *Biological foundations of language.* New York: Wiley.

Liddell, S. (1980). *American Sign Language syntax.* The Hague, Netherlands: Mouton.

Liddell, S. (1984). "Think" and "believe": Sequentiality in American Sign Language. *Language, 60*, 372–399.

Liddell, S., & Johnson, R. (1986). American Sign Language compound formation processes, lexicalization, and phonological remnants. *Natural Language and Linguistic Theory, 4*, 445–513.

Lillo-Martin, D. (1986). Two kinds of null arguments in American Sign Language. *Natural Language and Linguistic Theory, 4*, 415–444.

Mayberry, R. I. (1992). The cognitive development of deaf children: Recent insights. In F. Boller & J. Graffman (Series Eds.) & S. Segalowitz & I. Rapin (Vol. Eds.), *Child neuropsychology: Vol. 7. Handbook of neuropsychology* (pp. 51–68). Amsterdam: Elsevier.

McDonald, B. (1982). *Aspects of the American Sign Language predicate system.* Unpublished doctoral dissertation, University of Buffalo, Buffalo, NY.

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

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

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

McNeill, D., Cassell, J., & McCullough, K.-E. (1994). Communicative effects of speech-mismatched gestures. *Research on Language and Social Interaction, 27*, 223–238.

Meadow, K. (1968). Early manual communication in relation to the deaf child's intellectual, social, and communicative functioning. *American Annals of the Deaf, 113*, 29–41.

Metlay, D. S., & Supalla, T. (1995). Morpho-syntactic structure of aspect and number inflections in ASL. In K. Emmorey & J. S. Reilly (Eds.), *Language, gesture, and space* (pp. 255–284). Hillsdale, NJ: Erlbaum.

Moores, D. F. (1974). Nonvocal systems of verbal behavior. In R. L. Schiefelbusch & L. L. Lloyd (Eds.), *Language perspectives: Acquisition, retardation, and intervention* (pp. 377–417). Baltimore: University Park Press.

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*, 615–622.

Newport, E. L. (1981). Constraints on structure: Evidence from American Sign Language and language learning. In W. A. Collins (Ed.), *Minnesota Symposium on Child Psychology* (Vol. 14, pp. 93–124). Hillsdale, NJ: Erlbaum.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11–28.

Newport, E. L., & Ashbrook, E. F. (1977). The emergence of semantic

relations in American Sign Language. *Papers and Reports on Child Language Development, 13,* 16–21.

Newport, E. L., & Meier, R. P. (1985). The acquisition of American Sign Language. In D. I. Slobin (Ed.), *The cross-linguistic study of language acquisition, Vol. 1: The data* (pp. 881–938). Hillsdale, NJ: Erlbaum.

Newport, E. L., & Supalla, T. (1992). *Critical period effects in the acquisition of a primary language: I. The influence of maturational state on the acquisition of complex morphology in American Sign Language.* Unpublished manuscript, University of Rochester, Rochester, NY.

Padden, C. (1983). *Interaction of morphology and syntax in American Sign Language.* Unpublished doctoral dissertation, University of California, San Diego.

Padden, C., & Perlmutter, D. (1987). American Sign Language and the architecture of phonological theory. *Natural Language and Linguistic Theory, 5,* 335–375.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3,* 359–400.

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

Robins, R. H. (1952). Noun and verb in universal grammar. *Language, 28,* 289–298.

Sandler, W. (1986). The spreading hand autosegment of American Sign Language. *Sign Language Studies, 50,* 1–28.

Sapir, E. (1921). *Language: An introduction to the study of speech.* New York: Harcourt Brace Jovanovich.

Schachter, P. (1985). Parts-of-speech systems. In T. Shopen (Ed.), *Language typology and syntactic description: Clause structure* (Vol. 1, pp. 3–61). Cambridge, England: Cambridge University Press.

Schick, B. S. (1987). *The acquisition of classifier predicates in American Sign Language.* Unpublished doctoral dissertation, Purdue University, Lafayette, IN.

Schick, B. S. (1990). The effects of morphological complexity on phonological simplification in ASL. *Sign Language Studies, 66,* 25–41.

Scoble, J., & Mayberry, R. (1994). *Stuttering stops the flow of gesture too: The gesture-speech relationship in chronic stuttering.* Manuscript in preparation, McGill University, Montreal, Quebec, Canada.

Singleton, J. L., Morford, J. P., & Goldin-Meadow, S. (1993). Once is not enough: Standards of well-formedness in manual communication created over three different timespans. *Language, 69,* 683–715.

Singleton, J. L., & Newport, E. L. (1992). *When learners surpass their models: The acquisition of American Sign Language from impoverished input.* Unpublished manuscript, University of Illinois, Urbana-Champaign.

Stokoe, W. C. (1960). Sign language structure: An outline of the visual communications systems. *Studies in Linguistics, Occasional Papers, 8.* Buffalo, NY: University of Buffalo, Department of Anthropology and Linguistics.

Supalla, T. (1982). *Structure and acquisition of verbs of motion and location in American Sign Language.* Unpublished doctoral dissertation, University of California, San Diego.

Supalla, T. (1986). The classifier system in American Sign Language. In C. Craig (Ed.), *Noun classes and categorization* (pp. 181–215). Philadelphia: John Benjamins.

Supalla, T., & Newport, E. (1978). How many seats in a chair? The derivation of nouns and verbs in American Sign Language. In P. Siple (Ed.), *Understanding language through sign language research* (pp. 91–132). New York: Academic Press.

Supalla, T., Newport, E. L., Singleton, J. L., Supalla, S., Metlay, D., & Coulter, G. (in press). *Test battery for American Sign Language morphology and syntax.* San Diego, CA: Dawn Press.

Tervoort, B. T. (1961). Esoteric symbolism in the communication behavior of young deaf children. *American Annals of the Deaf, 106,* 436–480.

Thompson, S. A. (1988). A discourse approach to the cross-linguistic category "adjective." In J. A. Hawkins (Ed.), *Explaining language universals* (pp. 167–185). Cambridge, MA: Basil Blackwell.

Wilbur, R. (1986). Interaction of linguistic theory and sign language research. In P. C. Bjarkman & V. Raskin (Eds.), *The real world linguist: Linguistic applications for the 1980's* (pp. 166–182). Norwood, NJ: Ablex.

Wundt, W. (1973). *The language of gestures.* The Hague, Netherlands: Mouton. (Original work published 1900)

Yau, S. (1985). Sociological and cognitive factors in the creation of a sign language by an isolated member of a hearing community. In W. Stokoe & V. Volterra (Eds.), *SLR '83: Sign language research* (pp. 299–306). Silver Spring, MD: Linstok Press.

New Editor Appointed

The Publications and Communications Board of the American Psychological Association announces the appointment of Kevin R. Murphy, PhD, as editor of the *Journal of Applied Psychology* for a six-year term beginning in 1997.

As of March 1, 1996, submit manuscripts to Kevin R. Murphy, PhD, Department of Psychology, Colorado State University, Fort Collins, CO 80523-1876.