Chapter 5

Pointing Toward Two-Word Speech in Young Children

Susan Goldin-Meadow
Cynthia Butcher
University of Chicago

The first two-word combinations that a child produces are significant for two reasons. First, they reflect the child's developing ability to express propositional information within a single communicative act. Rather than produce *baby* and *drink* in separate utterances, the child can now conjoin them within a single sentence, *baby drink*, thus explicitly signaling that there is a relationship between the two elements. Second, two-word combinations are the child's first step into syntax. Independent of the language they are learning, children across the globe tend to produce the words that comprise their sentences in a consistent order. The particular orders they use mirror the orders provided by the language models they experience—*baby drink* rather than *drink baby* for an English-learning child. Even when the language a child is learning has relatively free word order, the child tends to adhere to a consistent pattern based on a frequently occurring adult pattern.

These two features of early two-word combinations are robust. They are found in the first two-sign combinations produced by deaf children acquiring a conventional sign language from their deaf parents (Newport & Meier, 1985), and even in the first two-gesture combinations invented by deaf children not yet exposed to conventional sign language by their hearing parents (Goldin-Meadow, 1997, 2002a).

Children begin to produce two-word sentences at approximately 18 months of age. They have, however, been able to produce isolated words since 12 months. Why is there a delay between the onset of words and the onset of word combinations? Combining two words into a single communicative act requires a number of skills. Children not only need to be able to intend to convey a proposition; they must also be able to segment that proposition into elements, label those elements with words, and combine the words into a single string.

Until children actually produce two words in a single combination, there is no explicit evidence in their talk that they intend to convey a proposition. There is, however, evidence that children in the one-word period can produce two elements of a proposition in one communicative act—but only if one looks across modalities. One-word children can utter a word—*drink*— and indicate the object of that action through their gestures—a point at a bottle. Assuming that gesture and speech are functioning as a unit, the two modalities together convey, to the observant listener, two elements of a single proposition.

In this chapter, we explore whether combinations in which gesture and speech convey different but complementary information are a transitional bridge between one- and two-word speech. We require two lines of evidence to support this hypothesis. First, in order for this type of combination to be a stepping-stone, gesture and speech must be functioning as a unified system. We therefore begin by exploring the onset of this type of combination in relation to the moment when gesture and speech come together into a well-integrated system—an event that takes place sometime during the oneword period (Butcher & Goldin-Meadow, 2000).

Next, we ask if integrated gesture-speech combinations are a harbinger of two-word combinations. Specifically, we explore the onset of combinations in which gesture and speech convey different information in relation to the onset of two-word speech. Whether these gesture–speech combinations precede or co-occur with two-word speech can provide insight into the conditions needed to combine words within a single sentence. If all that is needed for two-word combinations is the cognitive ability to convey two elements within a single communicative act, then gesture–speech combinations of this sort ought to co-occur with, and not precede with any regularity, the onset of two-word speech. Alternatively, if additional languagespecific skills are required for the onset of two-word combinations, then gesture–speech combinations in which the two modalities convey different information might be expected to reliably precede the onset of two-word speech.

Our goal here is to situate the onset of combinations in which gesture and speech convey different

information relative to two events: (a) the onset of gesture–speech integration during the one-word period, and (b) the onset of two-word combinations.

THE CHILDREN AND THE CODING CATEGORIES

The subjects for this study were six children, three boys and three girls, also described in Butcher and Goldin-Meadow (2000). The children were videotaped in their homes over a period of months, beginning when each child was in the one-word period of language development and continuing until the child began producing two-word combinations. All of the data were collected in spontaneous and unstructured play situations during which the children interacted with their primary caregivers and/or the experimenter around a standard set of toys and books. Four of the six children were seen approximately every 2 weeks; the remaining two subjects were seen approximately every 6 to 8 weeks. Table 5.1 reports the age range during which each child was observed and the number of videotaped sessions conducted during this period.

Table 5.1 also presents the age at which each child first produced a meaningful word (with or without a gesture) on our videotapes, and the age at which the child first produced a two-word combination on the videotapes. Because the videotaped sessions necessarily represent a small sample of each child's communications, the onset ages listed in Table 5.1 may inflate the actual onset ages for these children. Four of the children (Beth, Emily, Nicholas, and Joseph) were already producing meaningful words during their first observation sessions; the remaining two (Ann and Christopher) were not and produced their first meaningful words on the videotapes at ages 16.5 and 13 months, respectively. The ages at which the children began producing two-word combinations on our videotapes ranged from 18 to 26.5 months, an age span that falls within the range typically reported for the onset of two-word speech (cf. Bloom & Capatides, 1987; Bowerman, 1973; Braine, 1976).

TABLE 5.1
Subject Information

We focused in this study on gesture and speech that were used communicatively. All of the communicative gestures and words produced by each child during a half hour of videotape were transcribed and coded. If that half hour did not yield 100 communicative behaviors, additional tape was coded until 100 behaviors were transcribed. A *communicative behavior* was defined as either speech on its own, gesture on its own, or gesture and speech produced together. Children were given credit for having produced a spoken word if the vocalization sounded like an actual English word (e.g., *dog*, *cat*, *duck*, *hot*, *walking*) or if a sound was used consistently to refer to a specific object or action (e.g., *bah* used consistently to refer to a bottle).

Children were given credit for having produced a gesture if their hand movements were directed toward another person whose attention they had (i.e., if the hand movements were communicative) and if those movements were not themselves a direct manipulation of some relevant person or object (i.e., if the movements were symbolic). All acts that were done on objects were excluded, with one exception—if a child held up an object to bring it to another's attention, an act that serves the same function as the pointing gesture, it was counted as a gesture. The form of each gesture was described in terms of the shape of the hand, type of movement, and place of articulation. The vast majority of gestures that the children produced were deictics, either pointing at an object with the index finger or loose palm or holding up an object to call attention to it. The children also produced a few iconic gestures, in addition to nods, side-to-side shakes, and hand flips.

Gestures produced without speech and speech produced without gesture were identified and categorized but coded no further. Two additional coding decisions were made for gestures produced in combination with speech.

1. The timing of a gesture with respect to the speech it accompanied was coded to the nearest frame (1/30 sec) for the gesture–speech combinations produced by each child. Following McNeill (1992) and Kendon (1972, 1980), gesture–speech combinations were considered synchronous if the vocalization occurred on the stroke of the gesture or at the peak of the gesture (the farthest extension before the hand began to retract).
2. The relationship between the information conveyed in speech and the information conveyed in gesture in each gesture–speech combination was coded. Gesture–speech combinations were divided into two

types: (a) combinations in which gesture conveyed the same referent as did speech— for example, a point at a dog accompanied by the word *dog*; and (b) combinations in which gesture conveyed a different referent from the referent conveyed in speech—for example, a point at a pair of glasses accompanied by the word *mommy*. Note that the child's intent may be to make a statement with a gesture–speech combination (e.g., "those are mommy's glasses"), to make a request ("give me mommy's glasses"), or to ask a question ("are those mommy's glasses?"). For our purposes here, we focus only on the fact that *mommy* and *glasses*, the objects indicated in gesture and speech, are part of the utterance, whatever its function.

It is important to stress that gesture never conveys information that is completely redundant with the information conveyed in speech (McNeill & Duncan, 2000). For example, a point at a dog serves to draw an observer's attention to the object; it does not identify the object as a member of a category, as labeling the object with the word *dog* would. Nevertheless, pointing gestures do single out objects for attention, and it is in this loose sense that we say they "convey information" about those objects. If the object indicated by a point is also referred to in speech, we consider the point to convey the same information as the speech it accompanies; if not, we consider it to convey different information.

Reliability between two independent coders was assessed on a subset of the videotaped sessions and ranged between 84% and 100% agreement between the two coders, depending on the coding category (see Butcher & Goldin-Meadow, 2000, for further details on coding procedures and reliability).

GESTURE–SPEECH COMBINATIONS AND THE ONSET OF INTEGRATION ACROSS THE MODALITIES

The spontaneous gestures that adults produce as they speak can convey substantive information and, as a result, provide insight into a speaker's mental representations. Thus, gesture conveys meaning, although it does so differently from speech. Speech conveys meaning by rule-governed combinations of discrete units, codified according to the norms of that language. In contrast, gesture conveys meaning mimetically and idiosyncratically through continuously varying forms (Goldin-Meadow, 2002b; McNeill, 1992).[1]

Despite the fact that gesture and speech represent meaning in different ways, the two modalities form a single, integrated system in adults. Gesture and speech are integrated both semantically and temporally. For 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 (McNeill, 1992): The speaker grasped his hand as if gripping something and pulled the hand back. He produced this gesture as he uttered the words *and he bends it way back*. The gesture was an iconic representation of the event described in speech, and thus contributed to a semantically coherent picture of a single scene. In addition, the speaker produced the "stroke" of the pulling-back gesture just as he said "bends it way back," synchronizing the gesture with speech (see also Kendon, 1980).

The Onset of Gesture–Speech Integration

When do young children's gestures become integrated with the speech they accompany? At a time when children are limited in what they can say, they gesture (Bates, 1976; Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Petitto, 1988). The earliest gestures children use, typically beginning around 10 months, are deictics, gestures whose referential meaning is given entirely by the context and not by their form (McNeill, 1992). For example, a child may hold up an object to draw an adult's attention to that object or, later in development, point at the object. Children also occasionally use iconic gestures (McNeill, 1992). Unlike deictics, the form of an iconic gesture captures aspects of its intended referent. For example, a child might open and close her mouth to represent a fish or flap her hands to represent a bird (Iverson, Capirci, & Caselli, 1994; see also Acredolo & Goodwyn, 1985, 1988; Goodwyn & Acredolo, 1998).

Most of the gesture–speech combinations that children in the one-word period produce contain gestures that convey information redundant with the information conveyed in speech—for example, pointing at an object while naming it (de Laguna, 1927; Greenfield & Smith, 1976; Guillaume, 1927; Leopold, 1949). However, as described earlier, one-word speakers also produce gesture–speech combinations in which gesture conveys information that is different from the information conveyed in speech—for example,

gesturing at an object while describing the action to be done on the object in speech (pointing to an apple and saying “give”) or gesturing at an object and describing the owner of that object in speech (pointing at a toy and saying “mine”; Goldin-Meadow & Morford, 1985; Greenfield & Smith, 1976; Masur, 1982, 1983; Morford & Goldin-Meadow, 1992; Zinober & Martlew, 1985).

This second type of gesture–speech combination allows a child to express two elements of a sentence, one in gesture and one in speech. A child who produces such gesture–speech combinations can therefore be considered to have explicitly conveyed a proposition within a single communicative act—assuming, of course, that gesture and speech are working together as a unified system at this point in the child’s development. We explore this assumption here and ask when children begin to produce combinations in which gesture and speech convey different information relative to when they have integrated gesture and speech into a unified system.

Butcher and Goldin-Meadow (2000) investigated the onset of gesture– speech integration during the transition from one- to two-word speech in the six children who participated in this study. They noted first that the proportion of each child’s communications containing gesture remained relatively constant over the observation period (there were individual differences in the proportion of communications containing gesture, ranging from 20% to 40%, but each child’s proportion remained constant within him or herself). What changed over the observation sessions was the relationship gesture held to speech. In the earliest sessions, gesture did not appear to be well integrated with speech in three senses:

1. Gesture tended to appear without speech. In five of the six children, over 60% of the child’s communications containing gesture were produced without any accompanying sounds at all.

2. Gesture and speech did not form a temporally unified system. Only a small proportion of the few gesture–speech combinations that the five children produced at this early period were synchronous—that is, in most of their combinations, speech did not occur on the stroke or the peak of the gesture (cf. McNeill, 1992).

3. Gesture and speech did not form a semantically unified system. All of the few gestures that these five children produced with speech at this point in time were combined with meaningless sounds, despite the fact that all but one of these children were able to produce meaningful words. In other words, the children’s gestures were either produced alone or, less often, with meaningless sounds. The meaningful words they produced were uttered without gesture.

The relationship between gesture and speech changed when the five children began to produce gesture in combination with a meaningful word that conveyed the same information as the gesture (e.g., point at bottle + “bottle”). Figure 5.1 presents three pieces of data for these six children superimposed on a single graph: (a) the proportion of gesture-alone communications, which declined over time; (b) the proportion of synchronized gesture–speech combinations, which increased over time; and (c) the onset of combinations containing gesture plus words conveying the same information, shown as a vertical line on each graph. Note that for each of the five children who began to produce gesture + word combinations conveying the same information during our observation sessions, the three events converge: Gesture-alone combinations began to decline and synchronous gesture– speech combinations began to increase at just the moment when gesture was first combined in the same utterance with a word conveying the same meaning (the sixth child, Nicholas, had presumably gone through this convergence point prior to our observations).[2]

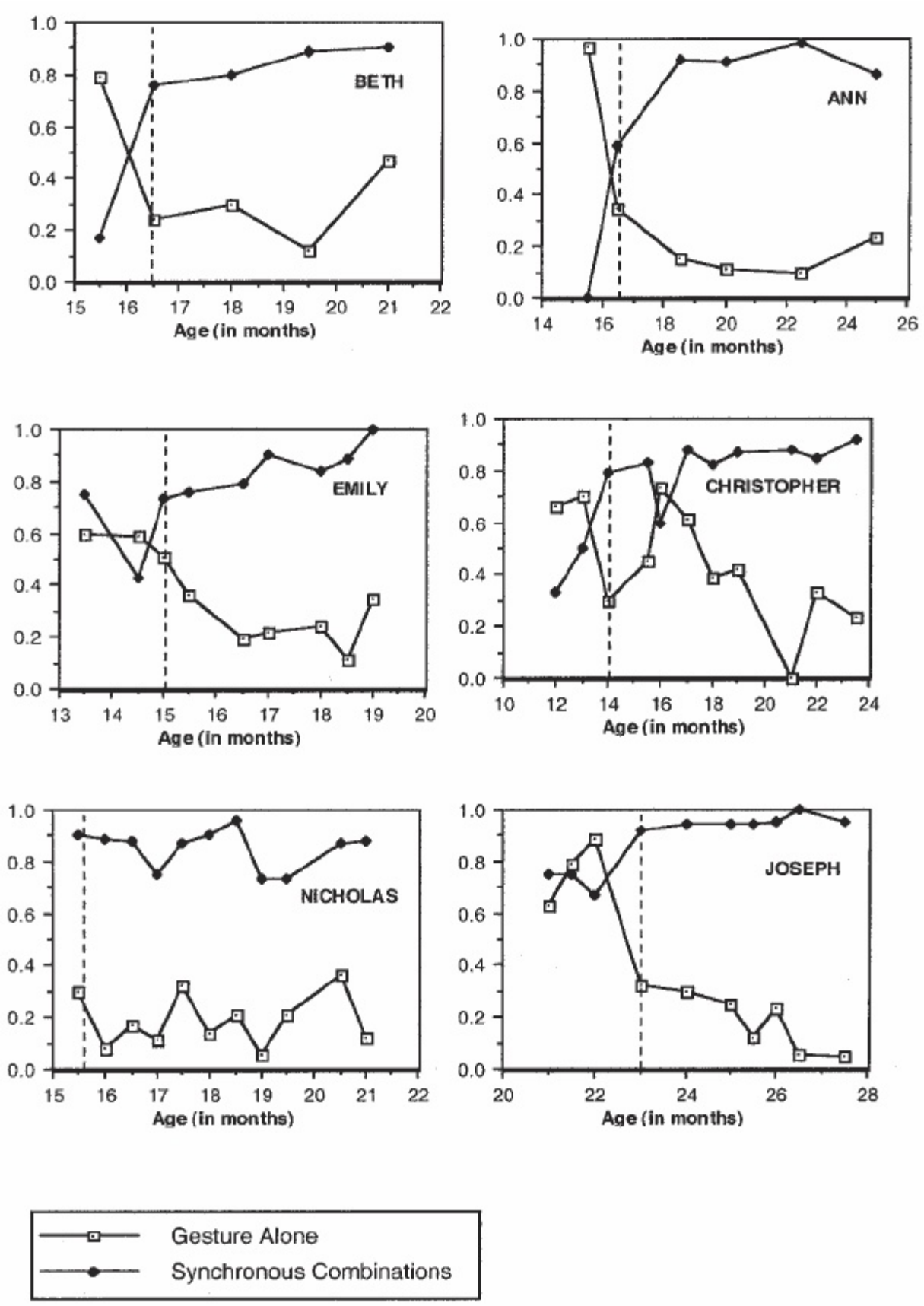

FIG. 5.1. Gesture–speech integration: the proportion of gesture-alone combinations (black diamonds) and the proportion of synchronous gesture–speech combinations (white squares) for each child. The vertical line demarcates the age at which each child first produced gestures in combination with words conveying the

same information. The convergence point of the three marks the integration of gesture and speech in terms of temporal synchrony and semantic coherence. Adapted from Butcher and Goldin-Meadow (2000), with permission.

Thus, the age at which each of the children began to produce communicative gestures in combination with words conveying the same information was precisely the age when timing began to improve dramatically in each child's gesture–speech combinations. Butcher and Goldin-Meadow (2000) took this point to be the onset of gesture–speech integration in these children, the moment when gesture and speech become unified into a system characterized by both semantic and temporal coherence.

The Onset of Combinations in Which Gesture and Speech Convey Different Information

If combinations in which gesture and speech convey different information are a product of two modalities operating independently, we would expect these combinations to occur prior to the session at which gesture and speech begin to merge into a single system—that is, prior to the onset of combinations in which gesture and speech convey the same information. Alternatively, if such combinations are the product of a single system, we would expect them to follow, or at the least co-occur with, the onset of combinations in which gesture and speech convey the same information. The left-hand and middle columns of Table 5.2 present the ages at which each of the six children first produced gesture–speech combinations in which the two modalities conveyed the same information (left) versus different information (middle).[3] None of the six children produced combinations in which the modalities conveyed different information before they produced combinations in which the modalities conveyed the same information.

These onset data suggest that combinations in which gesture and speech convey different information are a product of a unified system in which the two modalities work together. Further evidence for this hypothesis comes from the fact that this type of gesture–speech combination displays the two properties that are the hallmark of gesture–speech integration: semantic coherence and temporal synchrony.

TABLE 5.2
Age (in Months) of the First Appearance of Three Types of Combinations

Semantic Coherence in Combinations in Which Gesture and Speech Convey Different Information

In adult speakers, gesture and speech form an integrated system in the sense that the two modalities convey related, if not identical, information within an utterance. Thus, we asked whether the combinations that the children produced in which gesture conveyed different information from speech were semantically coherent.

In general, the gesture in combinations of this sort tended to indicate an object playing a semantic role in a proposition while the speech referred to another element of that same proposition. Table 5.3 presents examples of the types of these combinations that the six children produced. The most common gesture–speech combination of this type, the only one found in all six children, contained a word representing an action paired with a gesture, either a pointing gesture or a hold-up gesture, indicating the object involved in the action. Examples 1 through 6 in Table 5.3 are of this variety. The children also produced several other types of words combined with pointing or hold-up gestures: words representing the sound made by the object indicated in gesture (Examples 7–9), words representing an attribute of the object indicated in gesture (Examples 10–11), words representing other objects that were either the location (Example 12) or the owner (Example 13) of the object indicated in gesture, and negative words either rejecting the object indicated in gesture (Example 14) or noting its inappropriateness or lack of fit (Example 15).

TABLE 5.3
Examples of Combinations in Which Gesture and Speech Convey Different Information

In addition to pointing and hold-up gestures, the children also produced a few instances of other types of gestures combined with words that conveyed different information: side-to-side head shakes (Example 16), hand flips (Example 17), and iconic gestures (Examples 18–19).

Thus, the combinations that the children produced in which gesture and speech conveyed different information tended to express two elements that held some semantic relationship with respect to one another. Indeed, the number of combinations in which gesture and speech did not convey elements of a single underlying proposition was quite small (e.g., Nicholas pointed at the moon while saying "mama," two referents for which we could determine no propositional relationship[4]). Nicholas produced five, Emily two, Joseph one, and Ann, Beth, and Christopher produced no such combinations at all. As a result, in most of the children's combinations in which gesture conveyed different information from speech, gesture and speech each conveyed an element of the same proposition, suggesting that these combinations were the product of a unified gesture–speech system.

Temporal Synchrony in Combinations in Which Gesture and Speech Convey Different Information

In adult speakers, when gesture conveys information that is different from the information conveyed in speech, that gesture nevertheless has a systematic temporal relation with the speech it accompanies (McNeill & Duncan, 2000). For example, one adult produced a downward gesture while saying "goes through the drainpipe" (thus conveying the direction of motion in gesture and the location where that motion took place in speech). Despite the fact that the information conveyed in the two modalities was not identical, the word "drainpipe" was produced in synchrony with the stroke of the gesture. As other examples, Kita (1993) described subtle cases in which speech and gesture adjust to each other in timing in adults; Morrel-Samuels and Krauss (1992) provided evidence that the timing of gesture and speech is related to the rated familiarity of the spoken word; and Mayberry, Jaques, and DeDe (1998) provided evidence that gesture and speech are synchronized even when, as in stuttering, the speech production process goes awry.

To explore whether combinations in which gesture and speech conveyed different information were synchronous, we examined the timing in these combinations and compared it to the timing in combinations in which gesture and speech conveyed the same information. Because the children were all one-word speakers, we defined a synchronous combination as one in which the vocalization occurred on the stroke or peak (the farthest extension) of the gesture. Figure 5.2 presents the proportion of each type of gesture–speech combination that was synchronous for each of the six children. There were no significant differences in timing between the two types of gesture–speech combinations [Ann $X^2(1) = .20$, n.s.; Beth$X^2(1) = 0$, n.s.; Emily $X^2(1) = .23$, n.s.; Christopher $X^2(1) = .12$, n.s.; Nicholas $X^2(1) = 1.34$, n.s.; Joseph produced too few combinations in which gesture and speech conveyed different information to conduct a statistical analysis]. Thus, once gesture and speech become integrated, gesture–speech combinations tend to be synchronous, whether or not the information conveyed across the two modalities was the same.

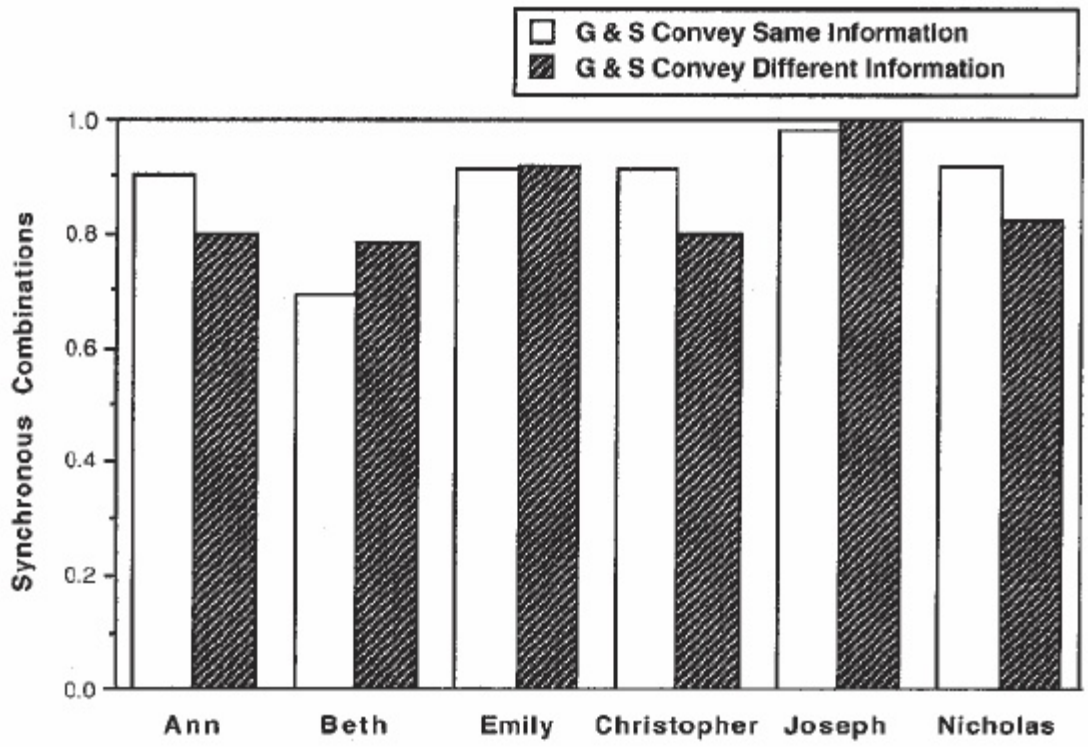

FIG. 5.2. Temporal synchrony in gesture–speech combinations: proportion of synchronous combinations in two types of gesture–speech combinations: combinations in which gesture and speech conveyed the same information (white bars), and combinations in which gesture and speech conveyed different information (black bars). The proportion of synchronous combinations was the same, and high, in both types of combinations.

GESTURE–SPEECH COMBINATIONS AND THE ONSET OF TWO-WORD SPEECH

We have shown that combinations in which gesture and speech convey different information are not produced until gesture and speech are well integrated into a unified system, suggesting that the two pieces of information reflect two elements of a single proposition rather than two unrelated elements. Children who can produce this type of gesture–speech combination clearly have the ability to convey two elements of a proposition within a single communicative act and should be well on the way toward being able to produce two-word utterances. Indeed, if the lack of this skill is the only ability preventing children from producing two-word combinations, they should begin to put two words together in a single string as soon as they begin to produce combinations in which gesture and speech convey different information. If, however, this skill is only one among many necessary for children to produce two-word combinations, they might be expected to produce two words together some time after they begin to produce combinations in which gesture and speech convey different information.

The middle and right-hand columns of Table 5.2 present the ages at which each of the six children in our study began producing gesture–speech combinations in which each modality conveys a different semantic element (middle) versus two-word combinations in which each word conveys a different semantic element (right). All six of the children began producing combinations in which gesture and speech conveyed different information prior to first producing two-word combinations. More impressive is the fact that, across the six children, the correlation between the age of onset of this type of gesture–speech combination and the age of onset of twoword combinations was quite high and reliable (rs = .90, $p < .05$, Spearman rank correlation coefficient, correcting for ties). The top graph in FIG. 5.3 displays the age at which each child began producing two-word combinations (y axis) as a function of the age at which that child began to produce combinations in which gesture and speech conveyed different information (x axis). Note that Joseph, who was substantially older than the other child and at the top of the scale in both measures, is an outlier inflating

the correlation. However, even without Joseph, the correlation was .82 (correcting for ties). In other words, the children who were first to produce combinations in which gesture and speech conveyed different information were also first to produce two-word combinations.

It is important to note that the correlation between gesture–speech combinations and two-word speech is specific to combinations in which gesture and speech conveyed different information. The bottom graph in Fig. 5.3 displays the age at which each child began producing two-word combinations (*y* axis) as a function of the age at which that child began to produce combinations in which gesture and speech conveyed the same (single) semantic element (*x* axis). Although all six children began producing combinations in which gesture and speech conveyed the same information prior to first producing two-word combinations, the correlation between the age of onset of this type of gesture–speech combination and the age of onset of two-word combinations was relatively low (r_s = .46, n.s., Spearman rank correlation coefficient, correcting for ties) and without Joseph, the outlier, dropped to .03 (correcting for ties).

Thus, the onset of combinations in which gesture and speech convey different information not only precedes the onset of two-word speech, it does so in a predictable fashion. The skills involved in producing this type of gesture–speech combination thus appear to be necessary for two-word speech but not sufficient to guarantee its onset.

DISCUSSION

This study explored the relationship between communicative symbolic gesture and speech in young children at the beginning stages of language development. It has two main findings. First, we found that combinations in which gesture and speech convey different information, despite the fact that they might intuitively be thought to reflect two separate systems operating independently, begin to be produced after gesture and speech become integrated into a single system and thus appear to be a product of this unified system integrated across the modalities. Second, we found that the onset of combinations in which gesture and speech convey different information predicts the onset of two-word speech. We consider each of these findings in turn.

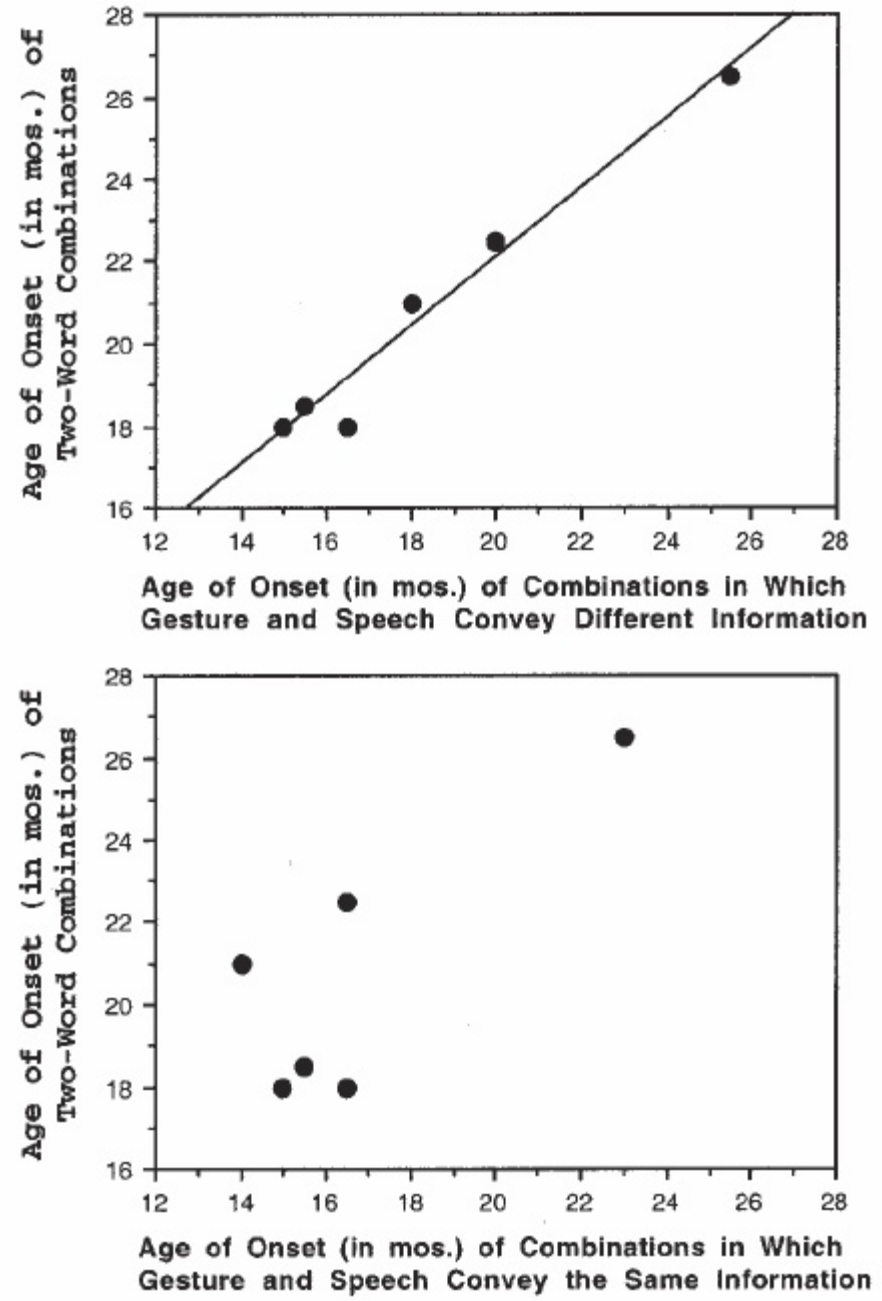

FIG. 5.3. Onset of two-word combinations: age at which each child first produced two-word combinations as a function of the age at which that child began producing combinations in which gesture and speech conveyed different information (top graph), and age at which that child began producing combinations in which gesture and speech conveyed the same information (bottom graph). The onset of two-word combinations is reliably related to the onset of the first type of gesture–speech combination but not the second.

Combinations in Which Gesture and Speech Convey Different Information Reflect a System Integrated Across Modalities

All of the children in our study produced combinations in which gesture conveyed information that was different from the information conveyed in speech. At first glance, combinations of this sort might be thought to be a product of the two modalities operating independently and in parallel. However, the timing of these combinations suggests that they are indeed a product of a single system integrated across the two modalities.

Combinations in which gesture and speech conveyed different information were first produced at the same age as, or after (but never before), combinations in which gesture and speech conveyed the same information (Table 5.2)—the point in development that can be taken to be the first sign of gesture– speech integration in the child's communications (Butcher & Goldin-Meadow, 2000; see Fig. 5.1). Combinations in which gesture and speech conveyed different information would have been expected to occur throughout development—in particular, before the onset of gesture–speech integration as well as after it—if they had, in fact, been a product of the two modalities operating independently of one another.

Moreover, combinations in which gesture and speech convey different information are characterized by the two properties that signal gesturespeech integration:

1. *Semantic coherence.* Despite the fact that gesture conveyed different information from speech, the information conveyed in the two modalities almost always was related and reflected a single idea unit (Table 5.3).
2. *Temporal synchrony.* Gestures that conveyed different information from the information conveyed in speech were nevertheless synchronously timed with respect to that speech (Fig. 5.2).

What happens over time to combinations in which gesture and speech convey different information? Do they disappear as the child begins combining words into longer strings and developing further cognitive skills? Although the particular type of gesture–speech combination that we have described here is likely to drop out as the child becomes more and more comfortable with two-word combinations, utterances in which gesture conveys information that is different from the information conveyed in speech continue to be found throughout development. For example, a child asked to solve a liquid quantity conservation task *says* that the transformed object is different from the original because "this one is taller than this one" but, in the same response, produces a *gesture* reflecting an awareness of the widths of the objects; specifically, she indicates with her hands the skinny diameter of the original object and the wider diameter of the transformed object, thus revealing knowledge of the widths of the task objects that was not evident in her speech (Church & Goldin-Meadow, 1986).

Instances in which gesture and speech convey different information in a problem-solving situation have been called "mismatches" (Church & Goldin-Meadow, 1986) and have been observed when individuals, both adults and children, are asked to describe their reasoning about a task—in toddlers going through a vocabulary spurt (Gershkoff-Stowe & Smith, 1997; in preschool and elementary school children reasoning about a board game (Evans & Rubin, 1979), Piagetian conservation tasks (Church & Goldin-Meadow, 1986), mathematical equivalence problems (Alibali & Goldin-Meadow, 1993; Perry, Church, & Goldin-Meadow, 1988), and seasonal change problems (Crowder & Newman, 1993); in adolescents reasoning about Piagetian bending rods tasks (Stone, Webb, & Mahootian, 1991); in adults reasoning about problems involving gears (Perry & Elder, 1996) and problems involving constant change (Alibali, Bassok, Solomon, Syc, & Goldin-Meadow, 1999); and in individuals of many ages reasoning about moral dilemmas (Church, Schonert-Reichl, Goodman, Kelly, & Ayman- Nolley, 1995) and Tower of Hanoi puzzles (Garber & Goldin-Meadow, 2002). Thus, the behavior that we have observed in our study of the oneword period in language development—combinations in which gesture conveys different information from speech—is one that continues throughout the developmental life span.

Combinations in Which Gesture and Speech Convey Different Information Predict the Onset of Two-Word Speech

We found, in six English-learning children, that the onset of combinations in which gesture and speech convey different information predicts the onset of two-word speech (Fig. 5.3). Similar findings have been reported in a group of Italian-learning children observed during the transition from oneto two-word speech (Volterra & Iverson, 1995). Thus, at the earliest stages of language learning, the relationship between gesture and speech within a single utterance appears to herald change in the child's linguistic system.

There are, in fact, other areas in which the relationship between gesture and speech heralds change. For example, children who produce a relatively large number of gesture-speech mismatches in their explanations of a particular task (e.g., a conservation task or a mathematical equivalence task) are more likely to benefit from instruction in that task than children who produce few mismatches (Alibali & Goldin-Meadow, 1993;

Church & Goldin-Meadow, 1986; Goldin-Meadow, Alibali, & Church, 1993; Perry et al., 1988, 1992). Thus, gesture–speech mismatch signals that the child is in a transitional state with respect to a task, and is therefore ready to make progress on that task if given appropriate input.

Are the situations in which gesture–speech mismatch signals the child's readiness to learn a given task comparable to the situation we have described, in which one-word speakers made the transition to two-word speech? In both instances, it is the relationship between gesture and speech that predicts the likelihood of cognitive change. The question is whether the relationship between gesture and speech is comparable in the two situations. We argue that it is—precisely because, in both situations, gesture and speech convey pieces of information that would, if considered together, form a single notion. The one-word speaker produces one semantic element in speech (e.g., the action "go") and a second element within the same proposition in gesture (e.g., the actor, point at a turtle). If considered together, these two pieces of information convey a single proposition (i.e., actor–action, turtle go). Similarly, the nonconserver describes one dimension of the conservation task in speech (e.g., the height of the containers) and another dimension also relevant to the task in gesture (e.g., the width of the containers). If considered together, these two pieces of information convey a single conservation rationale (i.e., compensation, the fact that even though the glass is taller than the dish, it is also skinnier than the dish).[5]

At some level, the child knows all of the information needed to express a proposition (in the case of the one-word speaker) or a compensation rationale (in the case of the nonconserver), and can even express that information within the bounds of a single communicative act. However, the child does not yet appear to be able to express the information within a single modality in that communicative act. A new level of understanding is apparently reached when the child combines the two pieces of the problem within a single modality, that is, when the one-word speaker expresses both elements of the proposition in a two-word utterance, or when the nonconserver expresses both dimensions of the compensation rationale in speech. Thus, in both situations, knowledge that is not yet sufficiently developed to be expressed within a single (spoken) modality may be expressed across two modalities.

Our data suggest that a one-word speaker who produces combinations in which gesture and speech together convey a single proposition fails to produce two-word combinations *not* because of a cognitive inability to grasp a proposition. Rather, the limitation appears to involve the ability to explicitly express a proposition entirely within the spoken modality. Although it may be the articulatory difficulty involved in producing two words in succession that prevents the child from expressing a two-element proposition in speech,[6] we suspect that the difficulty is deeper: that a child whose understanding of a proposition is tied in part to the manual modality cannot manipulate that proposition as well as one who can express the proposition entirely within the linguistic modality—in other words, that the child's understanding of the proposition is tied to a restricted range of uses and thus remains at an implicit level (cf. Goldin-Meadow & Alibali, 1994; Karmiloff-Smith, 1992). If so, making the transition to two-word speech would involve transforming this implicitly represented knowledge into an explicit form. By transposing information completely to the linguistic modality, the child generates a new level of representation, one that is a problem space in its own right and that can be worked on and improved as a modality of understanding (Karmiloff-Smith, 1979).

In summary, in both the young one-word speaker and the older child, a difference—or mismatch—between the information conveyed in gesture and the information conveyed in speech appears to signal readiness for cognitive growth. It is an open question as to whether the actual production of gesture–speech mismatch contributes to cognitive growth—that is, does the act of producing two different pieces of information across modalities but within a single communicative act improve the child's ability to transpose that knowledge to a new level and thus produce those pieces of information within a single modality? Our future work will investigate whether the act of producing gesture–speech mismatches itself facilitates transition (see Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001).

Even if it turns out that the production of gesture–speech mismatches has little role to play in facilitating cognitive change, mismatch remains a reliable marker of the speaker's potential for cognitive growth. As such, the relationship between gesture and speech may prove to be useful in clinical populations. For example, there is some evidence that children who are delayed in the onset of two-word speech fall naturally into two groups: children who eventually achieve two-word speech, albeit later in development than the norm (i.e., late bloomers), and children who continue to have serious difficulties with spoken language and may

never be able to combine words into a single string (Feldman, Holland, Kemp, & Janosky, 1992; Thal, Tobias, & Morrison, 1991). Indeed, in preliminary analyses of children with unilateral brain damage, we have found that some children display gesturespeech patterns comparable to those in Fig. 5.1 and Table 5.2 during their one-word period, whereas others display atypical gesture–speech patterns (e.g., gesture alone communications stay the same or even increase during this period, or gesture–speech synchrony decreases; Stare, 1996). Additional work is needed to determine whether atypical gesture–speech patterns early in development predict later language learning. If so, the relationship between spontaneous gestures and the speech they accompany may prove to be a useful clinical tool for distinguishing, at a relatively young age, children who will be late bloomers from those who will have difficulty mastering spoken language without intervention.

We have explored gesture's role in the transition from one- to two-word speech. The convergence of the semantic union and the temporal union of gesture and speech marks the beginning of gesture–speech integration in the one-word speaker. This integration sets the stage for the onset of gesture– speech combinations in which gesture conveys different (but related) information from the information that is conveyed in speech. These combinations, in turn, herald the onset of two-word speech. Thus, gesture provides the child with an important vehicle for information that is not yet expressed in speech, and, as such, it provides the listener (as well as the experimenter) with a unique window into the child's mind.

ACKNOWLEDGMENTS

This work was supported by grants from the March of Dimes Foundation and the National Institute on Deafness and other Communication Disorders (RO1 DC00491) to Goldin-Meadow, and by funding from the Home Health Care Foundation of Chicago through the section of Neonatology, Department of Pediatrics at the University of Chicago. We thank Samar Ali, Vera Joanna Burton, and Beth Stare for their help in coding the videotapes and establishing reliability; and Janellen Huttenlocher, Susan Levine, David McNeill, and William Meadow for their intellectual contributions throughout the project.

REFERENCES

Acredolo, L. P., & Goodwyn, S. W. (1985). Symbolic gesture in language development: A case study. *Human Development, 28,* 40–49.

Acredolo, L. P., & Goodwyn, S. W. (1988). Symbolic gesturing in normal infants. *Child Development, 59,* 450–466.

Alibali, M. W., Bassok, M., Solomon, K. O., Syc, S. E., & Goldin-Meadow, S. (1999). Illuminating mental representations through speech and gesture. *Psychological Science, 10,* 327–333.

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.

Bates, E. (1976). *Language and context*. New York: Academic Press.

Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. New York: Academic Press.

Bloom, L., & Capatides, J. B. (1987). Expression of affect and the emergence of language. *Child Development, 58,* 1513–1522.

Bowerman, M. (1973). *Early syntactic development: A cross-linguistic study with special reference to Finnish*. New York: Cambridge University Press.

Braine, M. D. S. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development, 41,* 164.

Butcher, C., & Goldin-Meadow, S. (2000). Gesture and the transition from one- to two-word speech: When hand and mouth come together. In D. McNeill (Ed.), *Language and gesture* (pp. 235–257). New York: Cambridge University Press.

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

Church, R. B., Schonert-Reichl, K., Goodman, N., Kelly, S. D., & Ayman-Nolley, S. (1995). The role of gesture and speech communication as reflections of cognitive understanding. *Journal of Contemporary Legal Issues, 6,* 123–154.

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.
de Laguna, G. (1927). *Speech: Its function and development*. Bloomington: Indiana University Press.
Evans, M. A., & Rubin, K. H. (1979). Hand gestures as a communicative mode in school-aged children. *Journal of Genetic Psychology, 135*, 189–196.
Feldman, H. M., Holland, A. L., Kemp, S. S., & Janosky, J. E. (1992). Language development after unilateral brain injury. *Brain and Language, 42*, 89–102.
Garber, P., & Goldin-Meadow, S. (2002). Gesture offers insight into problem-solving in children and adults. *Cognitive Science*.
Gershkoff-Stowe, L., & Smith, L. B. (1997). A curvilinear trend in naming errors as a function of early vocabulary growth. *Cognitive Psychology, 34*, 37–71.
Goldin-Meadow, S. (1997). The resilience of language in humans. In C. T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development* (pp. 293–311). New York: Cambridge University Press.
Goldin-Meadow, S. (2002a). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York: Psychology Press.
Goldin-Meadow, S. (2002b). *Hearing gestures: How our hands help us think*. Cambridge, MA: Harvard University Press.
Goldin-Meadow, S., & Alibali, M. W. (1994). Do you have to be right to redescribe? *Behavioral and Brain Sciences, 17*, 718–719.
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., & Morford, M. (1985). Gesture in early child language: Studies of deaf and hearing children. *Merrill-Palmer Quarterly, 31*, 145–176.
Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Sciences, 12*, 516–522.
Goodwyn, S. W., & Acredolo, L. P. (2000). Encouraging symbolic gestures: A new perspective on the relationship between gesture and speech. In J. M. Iverson & S. Goldin-Meadow (Eds.), *The nature and functions of gesture in children's communications* (pp. 61–73). San Francisco: Jossey-Bass.
Greenfield, P., & Smith, J. (1976). *The structure of communication in early language development*. New York: Academic Press.
Guillaume, P. (1927). Les debuts de la phrase dans le langage de l'enfant. *Journal de Psychologie, 24*, 1–25.
Iverson, J. M., Capirci, O., & Caselli, M. C. (1994). From communication to language in two modalities. *Cognitive Development, 9*, 23–43.
Karmiloff-Smith, A. (1979). *A functional approach to child language: A study of determiners and reference*. Cambridge: Cambridge University Press.
Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
Kendon, A. (1972). Some relationships between body motion and speech. In A. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 177–210). New York: Pergamon.
Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *Relationship of verbal and nonverbal communication* (pp. 207–228). The Hague: Mouton.
Kita, S. (1993). *Language and thought interface: A study of spontaneous gestures and Japanese mimetics*. Unpublished doctoral dissertation, University of Chicago.
Leopold, W. (1949). *Speech development of a bilingual child: A linguist's record, Volume 3*. Evanston, IL: Northwestern University Press.
Masur, E. F. (1982). Mothers' responses to infants' object-related gestures: Influences on lexical development. *Journal of Child Language, 9*, 23–30.
Masur, E. F. (1983). Gestural development, dual-directional signaling, and the transition to words. *Journal of Psycholinguistic Research, 12*, 93–109.
Mayberry, R. I., Jaques, J., & DeDe, G. (2000). What stuttering reveals about the development of the gesture-speech relationship. In J. M. Iverson & S. Goldin-Meadow (Eds.), *The nature and functions of gesture in children's communication* (pp. 77–87). San Francisco: Jossey-Bass.
McNeill, D. (1992). *Hand and mind*. University of Chicago Press.
McNeill, D., & Duncan, S. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), *Language*

and gesture (pp. 141–161). New York: Cambridge University Press.
Morford, M., & Goldin-Meadow, S. (1992). Comprehension and production of gesture in combination with speech in one-word speakers. *Journal of Child Language, 9*, 559–580.
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., & Meier, R. P. (1985). The acquisition of American Sign Language. In D. I. Slobin (Ed.), *The cross-linguistic study of language acquisition: Volume 1. The data* (pp. 881–938). Hillsdale, NJ: Lawrence Erlbaum Associates.
Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*, 359–400.
Perry, M., Church, R. B., & Goldin-Meadow, S. (1992). Is gesture–speech mismatch a general index of transitional knowledge? *Cognitive Development, 7*(1), 109–122.
Perry, M., & Elder, A. D. (1996). Knowledge in transition: Adults' developing understanding of a principle of physical causality. *Cognitive Development, 12*, 131–157.
Petitto, L. A. (1988). "Language" in the pre-linguistic child. In F. Kessel (Ed.), *The development of language and language researchers: Essays in honor of Roger Brown* (pp. 187–221). Hillsdale, NJ: Lawrence Erlbaum Associates.
Stare, E. M. (1996). *Congenital unilateral brain injury, gesture, and the transition to two-words*. Unpublished master's thesis, University of Chicago.
Stone, A., Webb, R., & Mahootian, S. (1991). The generality of gesture-speech mismatch as an index of transitional knowledge: Evidence from a control-of-variables task. *Cognitive Development, 6*, 301–313.
Thal, D., Tobias, S., & Morrison, D. (1991). Language and gesture in late talkers: A one year followup. *Journal of Speech and Hearing Research, 34*, 604–612.
Volterra, V., & Caselli, C. (1985). From gestures and vocalizations to signs and words. In W. Stokoe & V. Volterra (Eds.), *SLR '83, Proceedings of the Third International Symposium on Sign Language Research* (pp. 1–9). Silver Spring, MD: Linstok.
Volterra, V., & Iverson, J. M. (1995). When do modality factors affect the course of language acquisition? In K. Emmorey & J. S. Reilly (Eds.), *Language, gesture, and space* (pp. 371–390). Hillsdale, NJ: Lawrence Erlbaum Associates.
Zinober, B., & Martlew, M. (1985). Developmental changes in four types of gesture in relation to acts and vocalizations from 10 to 21 months. *British Journal of Developmental Psychology, 3*, 293–306.

[1]In this chapter, we consider gestures situated at one end of the continuum described by Kendon (1980)—spontaneous hand movements produced while talking (*gesticulation* in Kendon's terms). Thus, we leave aside more codified gestures such as emblems (e.g., OK, thumbs-up) that can be produced without speech.

[2]The high proportion of synchronous combinations found in each child's communications after gesture–speech integration (i.e., to the right of the vertical line in Fig. 5.1) was not unique to gesture + word combinations, but was also characteristic of the gesture + meaningless sound combinations that the child produced at this time.

[3]Once a child began to produce combinations in which the two modalities conveyed different information, the child tended to continue to do so. In the 26 sessions that followed the onset of combinations in which gesture and speech conveyed different information, there were only three sessions across the six children in which a child failed to produce such a combination (two in Joseph and one in Emily).

[4]We observed a few instances where the child pointed at an object (e.g., a pig) and labeled it with the wrong word (e.g., *cow*). Gershkoff-Stowe and Smith (1997) also observed combinations in which gesture indicates one object and speech another, apparently unrelated object in young one-word speakers naming pictures in a book. However, this type of combination might not actually be an instance of gesture conveying different information from speech if, in fact, the child really thinks that *cow* is the name for a pig. In their sample, Gershkoff-Stowe and Smith were able to tell that a given child had the word *pig* in his or her vocabulary, suggesting that these combinations were indeed an instance of gesture and speech conveying different information. We did not have enough data on each child's vocabulary to make comparable claims

for the children in our study. As a result, we did not count combinations of this type as instances where gesture and speech convey different information.

[5]In some cases of gesture–speech mismatch in older speakers, it is difficult to integrate the information conveyed across the two modalities into a single idea; for example, a child produces a *pour* gesture on the conservation task while talking about the container's width. Although these two pieces of information do not form a single conservation rationale, the child must have both in perspective in order to conserve—the child must understand that merely pouring the water is as irrelevant to its amount as the diameter of its new container. The pieces of information conveyed across the two modalities are very rarely in conflict even in older speakers; rather, they tend to reflect complementary aspects of the situation. Indeed, we suggest that mismatches reflect just those pieces of the situation that the speaker is currently working on integrating (Goldin-Meadow, 2002b).

[6]One argument against articulatory difficulties being responsible for the relatively late onset of two-word speech is that two-sign combinations are developed at the same, relatively late age in deaf children learning American Sign Language from their deaf parents (Newport & Meier, 1985; Volterra & Caselli, 1985). If articulation were playing a major role, one would expect some differences in the onset of two-unit combinations across the two modalities. Moreover, within this particular sample, all of the children except Beth produced a small number of meaningful words in combination with meaningless vocalizations several sessions before they produced two-word combinations (Beth produced both types of combinations for the first time during the same session.) Thus, the ability to produce two vocalizations in succession was not what prevented most of the children in this sample from combining two meaningful words into a single utterance.