In *Pragmatics & Cognition*, 6:2 (1998)

SPEECH-GESTURE MISMATCHES: EVIDENCE FOR ONE UNDERLYING REPRESENTATION OF LINGUISTIC & NONLINGUISTIC INFORMATION[1]

Justine Cassell
Media Laboratory
MIT

David McNeill
Psychology, Linguistics
University of Chicago

Karl-Erik McCullough
Linguistics
University of Chicago

Pantomime without discourse will leave you nearly tranquil, discourse without gestures will wring tears from you.

Essay on the Origin of Languages, Jean-Jacques Rousseau

Abstract

Adults and children spontaneously produce gestures while they speak, and such gestures appear to support and expand on the information communicated by the verbal channel. Little research, however, has been carried out to examine the role played by gesture in the listener's representation of accumulating information? Do listeners attend to the gestures that accompany narrative speech? In what kinds of relationships between gesture and speech do listeners attend to the gestural channel? If listeners do attend to information received in gesture, how is this information represented—is it 'tagged' as originating in the gestural channel? In this article research is described that addresses these questions. Results show that listeners do attend to information conveyed in gesture, when that information supplements or even contradicts the information conveyed by speech. And information received via gesture is available for retelling in speech. These results are taken to demonstrate that gesture is not taken by the listener to be epiphenomenal to the act of speaking, or a simple manual translation of speech. But they also suggest that the information conveyed in a discourse may be represented in a manner that is neither gesture nor language, although accessible to both channels.

Introduction

This article addresses the question of whether listeners attend to the kinds of gestures that people produce spontaneously while conversing; more specifically, whether information that is conveyed gesturally is integrated into the listener's understanding of what is communicated by a speaker. We show that listeners do attend to information presented in the gestural modality, and claim on the basis of this evidence that gesture and speech are integral parts of understanding—that they give rise to one unified representation in the listener, and can be drawn on as one unified representation.

A growing body of evidence shows that people unwittingly produce gestures along with speech in many different communicative situations. These gestures have been shown to elaborate upon and enhance the content of accompanying speech (Cassell & McNeill, 1991; McNeill, 1992; Bavelas, 1994, analyzing narrative or descriptions; Kendon, 1972; Streeck, 1993, analyzing conversation), often giving clues to the underlying thematic organization of the discourse or the speaker's perspective on events. Gestures have also been shown to identify underlying reasoning processes that the speaker did not or could not articulate (Church and Goldin-Meadow, 1986 examining Piagetian conservation; Crowder & Newman, 1993, examining science explanations).

Although there now seems to be substantial agreement on the natural occurrence of gesture, and its intimate relationship to some aspects of language, there is much less agreement on the *function* of gesture produced in the context of speech. Questions about "function" can be interpreted in several ways: researchers have asked at what stage in the production of language gesture functions, and they have asked whether the function of gesture is to help the listener or the speaker.

Thus, in answering the first question, some claim that gesture is a late occurring process in the production of language, either a translation of speech, or an outcome of the search for particular lexical items (Butterworth & Beattie, 1978; Butterworth & Hadar, 1989), hence the production of so-called word-finding gestures during pauses. These researchers believe that gesture is not integral to communication because, so they claim, the production of gesture is epiphenomenal to the production of speech. In answer to the second question, some claim that the primary function of gesture is not to communicate to the listener but to support the speaker's encoding of information (Freedman, 1972; Rime, 1982)—a kind of continuation of the sensorimotor stage of knowing that Piaget described for infants. This would be why gesture is produced when, for example, speaking on the telephone. Others argue the complementary position, that gesture is not communicative for the listener (Krauss, Morrel-Samuels, & Colasante, 1991).

Some researchers, however, reject the entire notion of a linear encoding and information processing model where a pre-verbal message is passed on to the language production module as input (Kendon, 1972; McNeill, 1992, to appear). In this case, the question for whom gesture functions loses its relevance. Since, according to McNeill, gesture and speech arise together from an underlying propositional representation that has both visual and linguistic aspects, the relationship between gesture and speech is essential to the production of meaning and to its comprehension. This explains why we find the strict temporal synchronization between the production of gesture and speech (Kendon, 1972), the parallel semantic and pragmatic content of gesture and speech (McNeill, 1992), the simultaneous acquisition in children (Riseborough, 1982) and tendency for the two systems to break down in parallel ways in aphasics (Feyereisen, 1983; Pedelty, 1987).

Surprisingly, the ongoing discussion of the function of gestures in the context of speaking has to a large extent neglected the question of what listeners get out of gesture. We know that gestures are still produced in situations where there is no listener, or the listener cannot see the speaker's hands (Rimé, 1982), although more gestures may be produced when an addressee is present (Cohen, 1977; Cohen & Harrison, 1973). Thus it appears that gestures must serve some function for the speaker, independent of any communicative intent. In addition, we know that information appears to be just about as effectively communicated in the absence of gesture—on the telephone, or from behind a screen (Short, Williams & Christie, 1976; Williams, 1977), and thus gesture is not **essential** to the interpretation of speech. But it has been shown that when speech is ambiguous (Thompson & Massaro, 1986) or in a speech situation with some noise (Rogers, 1978), listeners do rely on gestural cues (and, the higher the noise-to-signal ratio, the more facilitation by gesture). And, when adults are asked to assess a child's knowledge, they are able to use information that is conveyed in the child's gesture (and that is not the same as that conveyed by the child's speech) to make that assessment (Goldin-Meadow, Wein & Chang, 1992; Alibali, Flevares & Goldin-Meadow, 1994). So, the question that must be asked is, when gesture is available in a communicative situation, do listeners spontaneously heed the information it conveys. This article reports on a study where speech and gesture were deliberately **mismatched** in the telling of a story, that is, where the information conveyed by speech and that conveyed by gesture were different, as a way of examining whether there is uptake of gestural information by listeners, and if there is uptake whether gesturally-conveyed information is treated on a par with speech-conveyed information.

What is Gesture

What do we mean by 'gesture'? Our interest is in the communicative function of gesture in the context of speech; that is, in the **interaction** between speech and gesture. For this reason we consider only gestures that co-occur most often in conjunction with speech. Thus, we exclude the emblematic gestures that can replace speech (Ekman & Friesen, 1969). These gestures are culturally specified in the sense that a gesture may have a consistent interpretation that differs from culture to culture. Examples in American culture are the thumb-and-index-finger ring gesture that signals 'okay' or the 'thumbs up' gesture. Some cultures demonstrate much more consistent use of these emblems than others (Kendon, 1992).

This criterion also rules out gestures that only occur in the momentary absence of speech, for example 'word finding gestures' made when a speaker is engaging in a lexical search or when speech has failed (Butterworth & Beattie, 1978; Butterworth & Hadar, 1989).

Clearly, languages transmitted by way of a manual medium, such as ASL, are not the focus of the current research, likewise manual calques on languages—such as the Warlpiri Sign Language, studied by Kendon (1988), that is used when speech is prohibited—likewise the spontaneously emerging sign languages of deaf children who are not exposed to an existent language model at an early age (Goldin-Meadow, 1993).

Finally, our research does not address what might be called 'propositional gestures' (Hinrichs & Polanyi, 1986) such as the use of the hands to measure a particular space while the speaker says "it was this big". These gestures are not unwitting and in that sense not spontaneous, and their interaction with speech is more like the interaction of one grammatical constituent with another than the interaction of one communicative **channel** with another; in fact, the demonstrative "this" may be seen as a place holder for the syntactic role of the accompanying gesture.

Gesture Types

Our concern is, rather, with gestures that co-occur with speech and therefore may be serving a function as integral as that served by speech. Such gestures are the most frequent type in many communicative situations, including the narrative discourse situation on which we base this article. Four types of such gestures have been shown to occur with narrative discourse (McNeill, 1992).

1. *Iconics* depict by the form of the gesture some feature of the action or event being described; such as "he climbed up the pipe" accompanied by the hand rising upwards to show the path.

2. *Metaphoric gestures*, which are also representational, but where the concept being depicted has no physical form. An example is "the meeting went on and on" accompanied by a hand indicating rolling motion. There need not be a productive metaphor in the speech accompanying metaphoric gestures; sometimes the "metaphors" that are represented in gesture have become entirely conventionalized in the language. There does need to be a recognizable vehicle that mediates between the form of the gesture and the meaning of the speech it accompanies. Some common metaphoric gestures are the 'process metaphoric' just illustrated, and the 'conduit metaphoric' which objectifies the information being conveyed, representing it as a concrete object that can be held between the hands and given to the listener.

3. *Deictics*, which spatialize, or locate in the physical space in front of the narrator, aspects of the story being narrated; such as "Adam looked at Chuck, and he looked back" accompanied by a hand pointing first to the left and then to the right. Such gestures in narrations accurately index the left and right location of events from cartoon or event stimuli (McCullough, 1992).

4. *Beat gestures:* small baton like movements that do not change in form with the content of the accompanying speech. They serve a pragmatic function, occurring with comments on one's own linguistic contribution, speech repairs and reported speech. An example is "she talked first, I mean second" accompanied by a hand flicking down and then up on the word “second.”

Gesture Functions

It has been shown that speakers produce these four types of gesture during conversations and when narrating stories, and that these gestures play particular roles in narrative (Cassell & McNeill, 1991). In what follows, some of these narrative functions are described.

1. Iconic gestures may specify the manner in which an action is carried out, even if this information is not given in accompanying speech. For example, in the narration of a cartoon story, one narrator said "but he came right back out". Only in the narrator's gesture (right hand in a fist, sketching a very quick arc from in front of the body out to the right) is it clear that the character was **thrown** out. Another example comes from the same speaker who said "he went back and forth": once again, only the gesture (right hand in a fist, but first and second finger pointing down; fingers wiggle as if person walking) makes it clear that the character was **pacing** (as opposed, for example, to being tossed back and forth).

2. Iconic gestures may also specify the viewpoint from which the action is narrated. That is, gesture can demonstrate who narrators imagine themselves to be, and where they imagine themselves to stand at various points in the narration, when this is rarely conveyed in speech. For example, one narrator said "Granny whacked him and threw him out the window" while shaking her upright fist in front of her as if she were smacking someone with a bat, and then sweeping her hand away from her body. This narrator might equally well have shaken her head back and forth and rolled her eyes as if she were Sylvester being smacked by a bat, but she chose instead to take the point of view of the Granny.

3. Deictic gestures may locate characters in space, and make apparent the spatial relationships between them, even if this information is not conveyed verbally. This information may then serve the purpose of reference maintenance to the discourse entities. For example, one narrator said "Sylvester was in the Bird Watcher's Society building, and Tweetie was in the Broken Arms Hotel (. . .) he ran across the street". When referring to Sylvester, this narrator pointed to the right, when referring to Tweetie he pointed to the left. The referent of the pronoun was then disambiguated by a point to the right, meaning that it was Sylvester who ran across the street.

4. Beat gestures may signal that information conveyed in accompanying speech does not advance the plot of the story, but rather is an evaluative comment. For example, one narrator said "I can't remember the first thing he does" and accompanies this utterance with short formless beats. These gestures index the speech as a comment, in this case on the speaker's memory for the story, as opposed to a description of an event.

5. Metaphoric gestures may specify that a new narration or new segment of a narration is beginning. For example, one narrator begins her story by saying "okay, this is classic Sylvester and Tweetie". While saying "classic" she cups her hands and offers them to the listener in the conduit metaphoric shape. The 'conduit metaphoric' objectifies the information being conveyed, representing it as a concrete object that can be held between the hands and given to the listener. Conduit metaphorics commonly accompany new segments in communicative acts; the gesture described above indicates that accompanying speech refers to the classic status of the cartoon, which is about to be described (as opposed, for example, to referring to some behavior on the part of a character as being typical of the genre; in this latter case one would expect a beat gesture).

These five examples of the role played by gesture in narrative are interesting because they demonstrate narrative functions that appear to be universally important, and that linguists have attempted to tie to particular linguistic devices (Hopper, 1979; Hopper & Thompson, 1980), without much success (Cassell & McNeill, 1990). Because they exemplify the role played by gesture in the context of narrative speech, the narrative functions just described were the source for the creation of speech-gesture mismatches in the current study. Ultimately, only the first three were chosen because they illustrate the role of gesture in the communication of narrative information (as opposed to the communication of narrative structure), but the final two functions will be discussed in the conclusion to this article.

Method

Designing the Task

We investigate here the hypothesis that listeners take into account the information conveyed by gesture in their understanding of narrative discourse. Our study investigates whether gestures communicate information to listeners in the context of the speech with which they are produced. We therefore test the nature of comprehension when information differs between the two channels, rather than testing comprehension in the absence of one of the communicative channels. Although speakers do not appear to be able to recognize gestures that they have been exposed to previously, or to impute meanings to gestures when they are seen without their associated speech (Krauss et al., 1991), adults do seem to be able to "read" children's knowledge states off of the gestures that they produce in the context of speech (Goldin-Meadow, Alibali & Church, 1992; Alibali, Flevares & Goldin-Meadow, 1994). Two pilot studies (described below) led us to believe that we could determine the communicative import of gestures with respect to speech .

We therefore designed a task whereby listeners were exposed to a videotape of a narration in which the information conveyed by gesture can be distinguished from the information conveyed by speech. Our goal was to probe the uptake by listeners of information conveyed in the gestural channel, when it contrasted sharply with, or even opposed information conveyed in the verbal channel, as a way of determining the relationship between speech and gesture in the comprehension of narrative discourse.

Stimuli

An adult native speaker of English (the third author, depicted in Figure 1) was videotaped telling the story of a Sylvester and Tweetie Bird cartoon. He followed a script composed of phrases and gestures from actual subjects' narrations (see Appendix I for the script) of the Warner Bros. cartoon "Canary Row". This person was videotaped telling the story twice, using exactly the same words and intonation each time but modifying the gestures, in such a way that 14 target phrases accompanied by gestures were produced once with a normal gesture and once with a gesture mismatched to the content of accompanying speech. Speech-gesture mismatches such as these do not appear in the discourse of normal adults. Six additional phrases were accompanied by a gesture that gave supplementary information to that provided by speech (rather than contradicting information conveyed by speech); in these cases gestures gave supplementary information to speech in both videotapes, and the information provided by gesture differed between the two versions. Cases where gestures and speech provide complementary and additive information to one another are often found in normal discourse situations. Matched and mismatched gestures were compared to insure that the mismatched gestures did not call attention to themselves in some way. The fact that both speech-mismatched and some normal gestures were performed according to a script, and that it is the comparison between them that is being tested, minimizes the chances that some unintended salience of speech-mismatched gesture production accounts for any results. Nevertheless, gestures were compared according to standard criteria and, in fact, the speech-mismatched gestures and the scripted normal gestures corresponded to the narrator’s spontaneous gestures in terms of size of gesture, placement in space of gesture, and temporal association of gesture to accompanying speech.

Three types of gesture pairings were used in the stimulus videos: Two types of gesture-speech associations were designed where speech and gesture were contradictory and one type of gesture-speech association was designed where gesture gave additional non-overlapping but non-contradictory information. These three types were included and designed on the basis of the first three narrative functions of gesture described above, and building on lessons learned from the data used in an earlier pilot study (described in McNeill, 1992, and below). The three types of gesture-speech associations were 'anaphor', 'origo', and 'manner'. They are described below.

1. *Anaphor mismatches* set up two referents in space with deictic gestures and then violate the assumptions of referring back to the two referents by pointing to the wrong space. Thus, our knowledge of linguistic anaphors tells us that one discourse entity is being referred to, while gesture contradicts that by indicating another entity. An example from the stimuli is given below[2]:

> [Sylvester] is right near [Tweetie], watching him, and then suddenly [he lunges for him] and runs into the apartment after him.
> (a) *normal: right hand points to space representing Sylvester, left hand points to space representing Tweetie; then right hand lunges leftward*
> *(b) mismatched: right hand points to space representing Sylvester, left hand points to space representing Tweetie; then left hand lunges leftward*

In this example, our knowledge of the two characters and their behavior tells us that the cat does the chasing, and thus that the "he" in "he lunges" refers to Sylvester. In the mismatched version of the gesture, though, the hand that has been used to represent Tweetie does the lunging. Figure 1 shows Sylvester (right hand) lunging towards Tweetie (left hand) and Figure 2 shows the Tweetie (left hand) lunging away.


Figure 1: Normal Anaphor


Figure 2: Mismatch Anaphor

2. *Origo mismatches* are iconic gestures which represent the perspective from which the action is seen, (the term is borrowed from Buhler's theory of deixis). In particular, origo-mismatched-gestures provide a perspective on the action different from that assumed by the accompanying speech. An example of an origo mismatch from the stimuli is given below:

> Granny sees him and says "oh what a nice little monkey". And then she [offers him a penny].
> *(a) normal : left hand proffers penny in the direction of listener.*
> *(b) mismatched: left hand offers penny to self*

Figure 3: Normal Origo

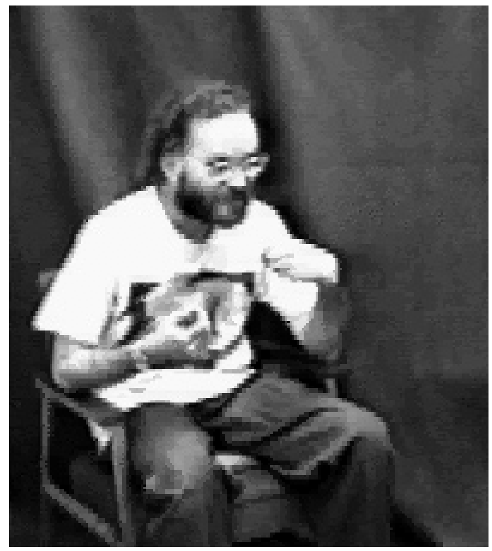

Figure 4: Mismatch Origo

Figure 3 shows the left hand offering something towards the listener. The mismatched version, in Figure 4, is especially striking because it shows a mixed perspective on the action. The speaker's arm represents Granny's arm, and his body represents Sylvester's body. This gesture, however, would seem perfectly normal if it accompanied the phrase "and then she *got a penny*".

3. *Manner mismatches* do not necessarily contradict speech, but they provide additional or different information about the manner in which an action is carried out, when this information is not given in accompanying speech. In particular, for verbs that describe movement from one point to another without specifying how that movement is carried out (e.g. "go", "leave", "enter"), these iconic gestures specify the manner of motion (e.g. by crawling, bouncing, being propelled). An example of manner gestures from the stimulus narrations is given below:

> So then Sylvester sees an organ grinder and a monkey on the street. He hides around the corner from them, and he gets the monkey [to come to him].
> *(a)* $manner_A$*: narrator makes 'beckoning' gesture with crooked forefinger*
> (b) $manner_{B:}$ *narrator makes 'grabbing' motion with left fist.*

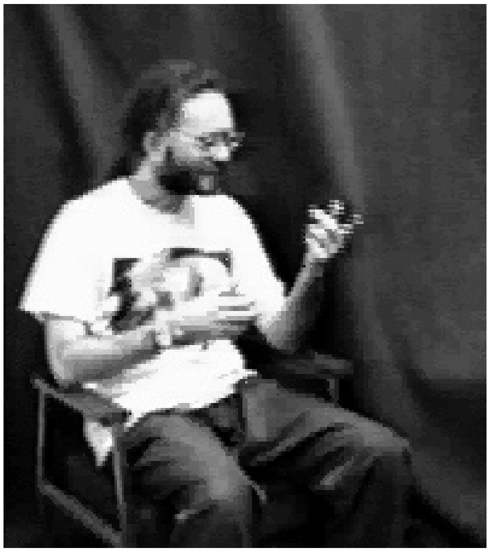

Figure 5: Manner (a) "beckon"

Figure 6: Manner (b) "grab"

In this example, the gesture provides information not given in speech, namely that the way in which Sylvester gets the monkey to come to him is either by beckoning the monkey to walk around the corner (Figure 5), or by grabbing the monkey and pulling him around the corner (Figure 6).

In the first pilot study that we carried out (described in McNeill, 1992), we videotaped one single version of a similar script, containing 20 gestures mismatched to speech, and showed it to four subjects. Our analyses of these data showed that in the actual study two speech-mismatched gestures must not be carried out one after the other, since it was then impossible to distinguish the effect of the two gestures in subjects' retellings. In a second pilot study (described in McNeill, Cassell & McCullough, 1994) we reorganized the order of the mismatched and matched gestures, and again showed the stimuli to subjects. This time we compared the performance of subjects exposed to the video to the performance of subjects exposed to audio only, asking whether we could distinguish gestures that had been influenced by what subjects had seen, from gestures that came from the subjects' own imagination. We believed that, once again, we could determine—on the basis of story retellings—cases in which subjects had been influenced by the gestures they had seen. However, it was clear from analyses of these data, that speech-mismatched gestures must be directly compared to matched gestures in the same context. We had until this point not compared like situations.

These two pilot studies also convinced us that, for manner gestures, the description "mismatched" was inappropriate. In the example given above, for instance, neither the 'beckoning' gesture nor the 'grabbing' gesture *contradicts* the phrase "he gets the monkey to come". Rather two gestures add information that leads the listener towards two different interpretations of the verb phrase "get to come".

Thus, as shown in Table 1, in the current study two counterbalanced versions of the narration were constructed in such a way that one version had a normal gesture where the other version had a speech-gesture mismatch, and in the case of manner gestures, one version had gesture (a) while the other version had gesture (b). Each version contained roughly equal numbers of normal and mismatched gestures, and equal numbers of the different kinds of mismatches. Speech-mismatched gestures were never carried out immediately after other speech-mismatched gestures. Each version was divided into three episodes of roughly 2 minutes each.

Gesture	Type	Description	Version A	Version B
1	Anaphor	Sylvester dashes across street	mismatch	match
2	Manner	Sylvester exits hotel	“walking”	“flying”
3	Manner	Sylvester goes up pipe	“climbs”	“goes up”
4	Anaphor	Sylvester grabs Tweetie	match	mismatch
5	Manner	Grannie whacks Sylvester	“punches”	“slaps”
6	Origo	Tweetie drops bowling ball	match	mismatch
7	Anaphor	Sylvester falls down pipe	mismatch	match
8	Manner	Sylvester goes down the street	“goes down”	“bounces”
9	Manner	Sylvester lures monkey	“grabs”	“beckons”
10	Origo	Sylvester steals clothing	mismatch	match
11	Origo	Granny offers penny	match	mismatch
12	Anaphor	Granny whams Sylvester	match	mismatch
13	Origo	Sylvester knocks	match	mismatch
14	Anaphor	Sylvester throws suitcase	mismatch	match
15	Origo	Sylvester removes cover	match	mismatch
16	Anaphor	Granny strikes Sylvester	mismatch	match
17	Anaphor	Sylvester grabs Tweetie	match	mismatch
18	Origo	Weight lands on Sylvester	mismatch	match
19	Manner	Sylvester makes plans	“taps head”	“draws”
20	Anaphor	Sylvester slides down wall	mismatch	match
21	Anaphor	Sylvester escapes trolley	match	mismatch
Total	**9 Anaphor** **6 Manner** **6 Origo**		**7 mismatch** **8 match** **6 manner**	**8 mismatch** **7 match** **6 manner**

Table 1: Distribution of Gesture Types

Subjects

Sixteen college undergraduates served as subjects, of whom eight observed the videotape and narrated what they saw and eight were listeners (role was chosen by coin toss). Only the narrators are of interest here, and so data from the eight narrators will be discussed. Subjects were randomly assigned to one of the counterbalanced video conditions. Narrator-listener pairs were tested in sessions that lasted about 3/4 hour. Subjects' participation fulfilled a course requirement. All subjects were native English speakers, and data from subjects who had seen this Sylvester and Tweetie Bird cartoon before were discarded before analysis.

Procedure

Narrators watched the first narration segment of the stimulus video (showing our narrator telling the story of a Sylvester and Tweetie Bird cartoon) on a 24-inch color monitor (the entire video was not shown at one time as, in pilot testing, subjects were unable to recall the entire story). After having watched the narration segment, the narrators were requested to tell the listener (who had been sitting in the next room) the whole story thus far, everything they had seen and heard. The procedure was repeated for the subsequent two narration segments. The instructions and procedure were identical to those of previous experiments from this laboratory, so that data would be comparable[3]. In order to determine whether subjects were conscious of the speech-gesture mismatches, after they had finished narrating all three segments, and before debriefing, subjects were asked if anything stood out, or if they found anything unusual about the storytelling. No subject reported noticing the narrator’s gestures, or relationship between speech and gesture. The comment that came closest to being a remark about gestures was one subject’s report that the narrator was “very animated”

Coding

Two coders independently transcribed speech and gesture for each subject's narration of the stimulus video (using the transcription conventions from McNeill, 1992). The two coders then compared their transcriptions, and the few disagreements were resolved by discussion. This provided a complete written transcript of each subject's narration.

The critical question to be answered by analysis of these transcripts is whether gesture contributed to subjects' understanding of the narration. This question may be addressed by looking at whether there was a difference in understanding of the narration between subjects in the two video version conditions, i.e. between subjects who heard the same speech but saw two different gestures.

Analyses of pilot data (McNeill, 1992 and McNeill, Cassell & McCullough, 1994), where only one video version was shown, led us to believe that effects of speech mismatched gestures could be traced in subjects’ subsequent retelling of the story. Analyzing those data, we looked for places in the subjects’ retellings where a choice of word or gesture could be traced back to a gesture or a word that the subject had been exposed to in the stimulus narration. Despite the roughness of this technique, we were led to believe that effects of the speech-mismatched gestures were recognizable. In the current study, then, we refined this technique to blind comparisons of subjects retellings of the two video versions.

Thus, those clauses in subjects' narrations that narrated the 21 target segments in the stimulus video (that is, segments that in one version of the stimulus video were narrated with one gesture and in the other version were narrated with another gesture) were highlighted and subjects' transcripts were given to three coders blind to the version of the stimulus narration seen by each subject. These coders were also given the speech transcript followed by the narrator in the stimulus video and asked to judge, for each highlighted section, whether the subject's retelling (including both speech and gesture) was accurate or inaccurate with respect to the speech that the stimulus narrator had produced. If the coders judged that the re-telling was inaccurate with respect to the original, they then compared the inaccuracy with the gestures produced in the two versions of the stimulus narration, and judged if the inaccuracy was provoked by seeing one of those gestures[4]. Each transcript was looked at independently by three coders; in rare cases of disagreement, the decision of two out of three coders was followed.

If gesture has no effect on understanding of a narration, then subjects who saw the two versions of the video (which contained exactly equal speech) should produce equally accurate-to-speech re-tellings. However, if gesture has an effect on the understanding of a narration, those subjects who saw gestures that differed in content from accompanying speech should be likely to produce a re-telling reflecting what they saw in the stimulus narrator's gestures (and consequently inaccurate with respect to his speech).

If the information conveyed by gesture does contribute to one's understanding of a narration, then a second question arises: in a subsequent retelling, is information that was taken in via gesture subsequently re-told only in subjects' gestures, or may it cross channels to appear in their speech? In order to answer this question, if coders judged that a subject's retelling was not accurate with respect to the speech of the original narration, then they coded the locus of the inaccuracy (that is, what in the subject's transcription was inaccurate with respect to the speech heard) as speech, gesture, or speech and gesture.

For example, subject X saw the following manner mismatch in the stimulus video: the narrator on the video said,

> Luckily Granny is there, Tweetie's friend, and she [whacks] him one, and throws him back out the window

and simultaneously made a punching gesture towards the front. In her subsequent narration of the video, Subject X said,

> "but Tweetie's granny is there and she punches Sylvester out"

while with her hand she made a vague waving motion of the hand.

Referring to the transcript of the stimulus video, the coder saw that the original verb used was “whacks” while the subject says “punches out”. Judging that there is an inaccuracy with respect to the original, the coder went on to look at the two possible gestures seen by this subject. One version of the stimulus materials was accompanied by a punching gesture, and one by a slapping gesture. Not knowing which of the gestures this subject had actually seen, the coder judged that this retelling was due to the gesture “punch” that was present in one version of the stimulus materials. The coder therefore counted this as an instance of inaccuracy (in essence meaning that the coder believes that the subject's production of the verb "punch" arose from the subject having seen a punching gesture), with locus of effect in speech (meaning that the effect of the mismatch was found in the subject's speech) . If the subject's gesture had represented punching (rather than simply representing some kind of hand movement), then the coder would have counted it as an instance of inaccuracy with locus of effect in speech and gesture, because the mismatch would have had an effect on speech **and** gesture.

Analysis

The general question asked in this study concerns whether gestures are taken into account by listeners in their representations of the information conveyed during spoken events. Our methodology used gestures that differed from speech so that it was possible to tell through subjects’ retellings of a story whether they took into account the content of gestures as well as the content of speech. In particular, the gestures differed in two ways: origo and anaphor gestures contradicted information conveyed in accompanying speech, and manner gestures added non-redundant information. For gestures that contradicted the speech they accompanied, we question the data in the following way: for each potential gesture-speech mismatch point (target phrase) do the retellings of subjects differ between the version of the narration where that gesture is mismatched to speech from the version where it is paired to speech?

The question is asked slightly differently for manner “mismatches” than for anaphor and origo mismatches, since there is no version in which the gesture is redundant with speech. Thus, for manner gestures, we ask if it is possible to tell which gesture has been viewed by a subject from that subject’s retelling—assuming that if coders *guessed* which gesture had been produced, they would be correct 50% of the time. Thus we ask whether coders judgments as to which gesture was viewed were significantly better than 50% correct.

We also ask, of all of the data, whether some kinds of relationships between speech and gesture are more likely to be taken up by listeners than others: that is, do manner mismatches, where the relationship between speech and gesture is additive, get taken up more often than anaphor mismatches, where information conveyed in gesture directly contradicts that conveyed in speech? Finally, we ask whether information conveyed via gesture is represented subsequently only in the re-teller's gesture, or does that information also cross channels to be subsequently represented in speech?

Results

Effect or Not

If gesture is not attended to by listeners, if it is not communicative, then we expect that listeners will not notice the mismatched gestures or the gestures that add additional non-redundant information. That is, if there is no uptake of gestural information, then listeners will tell correct stories, and make correct gestures ('correct' with respect to the speech that they heard) since the story is interpretable if the gestures are ignored.

Subjects produced a mean of 6.63 retelling inaccuracies (where retellings did not match the speech of the original stimulus narration). A mean of 3.25 of these retelling inaccuracies occurred in the context of manner gestures. A mean of 3.38 of these retelling inaccuracies occurred in the context of the other two types of speech-gesture associations (anaphor and origo gestures). Of this latter number, only .38 retelling inaccuracies occurred in the context of speech-gesture matches. The other 3 mean retelling inaccuracies occurred in the context of speech-gesture mimatches. Figure 7 represents what percentage of each type of speech-gesture association was reflected in a retelling inaccuracy (that is, what percentage of the 7.5 mean mismatched gestures, 7.5 mean speech-gesture matches, and 6 manner gestures seen by subjects).

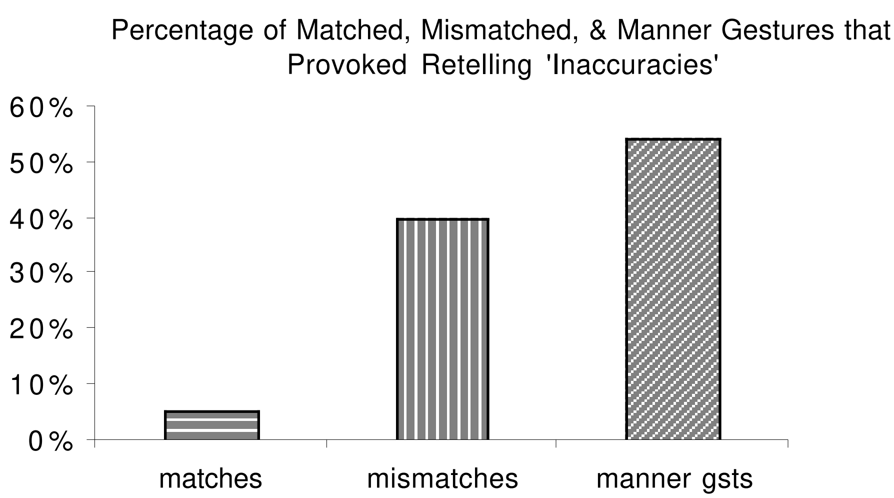

Figure 7

Figure 7 shows that only 5% (SD = 1.67) of speech-gesture matches provoked retelling inaccuracies, while 40% (SD=1.69) of speech-gesture mismatches provoked a retelling inaccuracy and 54% (SD=.71) of manner (non-redundant) gestures provoked a retelling inaccuracy.

Of course, as explained above, manner gestures cannot be compared directly to speech-gesture matches or mismatches. For that reason, two different analyses were carried out on the data. Data were analyzed using analysis of variance with repeated measures, and the Wilcoxon Signed Rank test.[1] The ANOVA was carried out on those gestures with a speech-matched and speech-mismatched version (origo and anaphor gestures), with matchedness (gesture matched to

[1] The Wilcoxon Signed Rank test is a nonparametric procedure used with two related variables to test the hypothesis that the two variables have the same distribution. It makes no assumptions about the shapes of the distributions of the two variables. This test takes into account information about the magnitude of differences within pairs and gives more weight to pairs that show large differences than to pairs that show small differences. The test statistic is based on the ranks of the absolute values of the differences between the two variables.

speech vs. speech mismatched to gesture) and mismatch type (anaphor vs. origo) as within-subjects factors, and version of the stimulus video as the between-subjects factor. A preliminary ANOVA showed that the version of the stimulus video that subjects watched had no significant effect and so subsequent analyses collapsed the data across the two groups. The difference between retelling ‘inaccuracies’ in the context of speech-gesture matches and speech-gesture mismatches is significant ($F(1,7) = 10.14$, $p<.02$). No differences were found due to mismatch type (anaphor vs. origo). Thus, when gesture conveyed information that contradicted the information conveyed by speech, listeners took that information into account when retelling the story they heard.

The Wilcoxon Signed Rank test was carried out on those gestures where both stimulus versions contained gestures that differed from the speech. In this case, both gestures convey information that differs from the information conveyed by speech, although it does not contradict it. We are interested here in the question of whether the information conveyed by gesture is taken up by the listener and conveyed in her subsequent retelling of the story. We know that the information conveyed by gesture was taken up if we can recognize the information conveyed only by the stimulus film’s gestures in the subject’s retelling. The question that was asked here, therefore, concerned whether it was possible to recognize which gestures subjects had seen from their retellings with significant frequency. And, in fact, the Wilcoxon shows that to be the case (Sgn Rank 18 $Pr>=|S|$, $p<.008$).

Type of Mismatch

If gesture is communicative, but not an equal partner in the building of a representation of information, then one might expect manner 'mismatches' to be regularly attended to by listeners, while the two other kinds of mismatches would not. This is because manner mismatches expand on and support speech, as opposed to contradicting it, and this is the most unmarked (and least controversial) function for gestures to have. Origo and anaphor gesture mismatches convey information that contradicts that conveyed by accompanying speech. If listeners incorporate this information, they must have over-ridden the information they received via speech, or reconciled the information conveyed by the two channels.

Figure 8 shows the comparison between the percentage of each gesture type that provoked retelling inaccuracies (that is, what percentage of origo mismatches resulted in retelling mismatches, what percentage of anaphor mismatches resulted in retelling mismatches, and what percentage of manner gestures resulted in retelling effects). In fact, a glance at the data seems to show that no difference exists between the percentages of manner mismatches, anaphor mismatches, and origo mismatches that were incorporated into listeners' retellings. As reported above, no significant differences were found between the effect of origo and of anaphor gestures ($F(1,7) = 3.91$, $p<.09$).

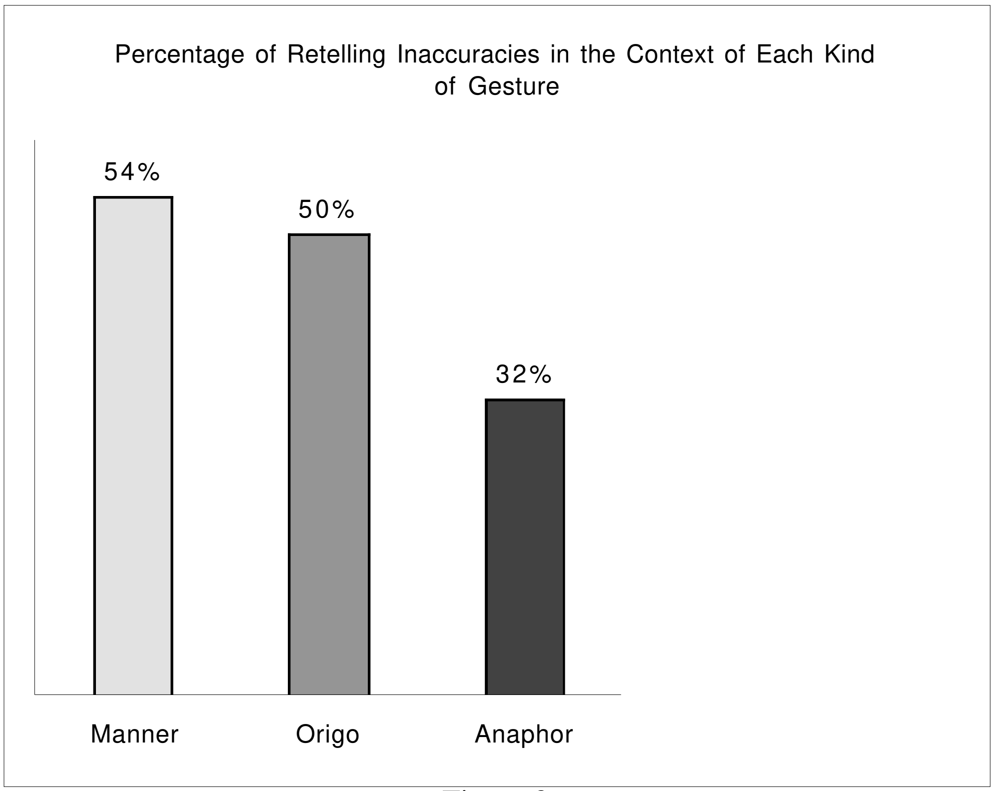

Figure 8

Manner

As reported above, and shown in Figure 8, 54% of manner mismatch stimuli had an effect on listeners' retellings. An example of a retelling "inaccuracy" in the context of a manner gesture comes from a subject who heard the narrator say "and Granny whacked him one" but saw the narrator making a punching gesture (narrator stimulus is shown in Figure 9). The subject said,

> "And Granny like punches him or something and you know he whacks him."

and, as shown in Figure 10, with his right hand in a fist, the subject punches forward.

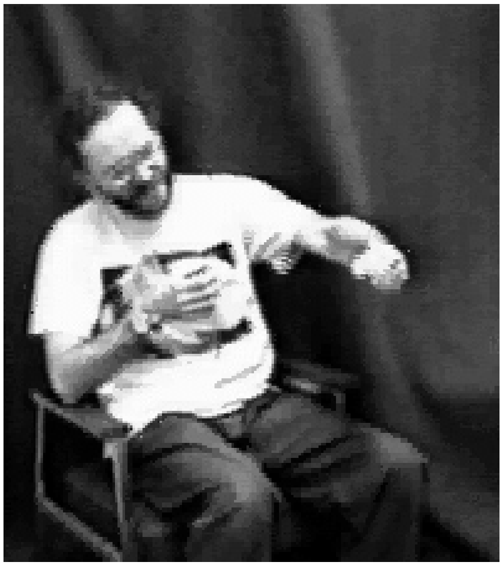

Figure 9: Manner stimulus "punch"

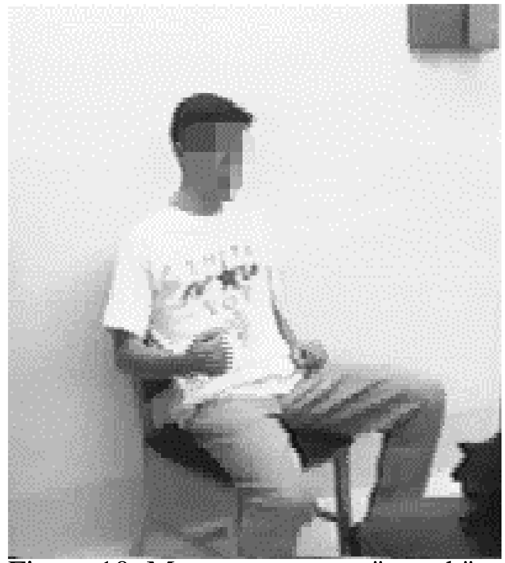

Figure 10: Manner response "punch"

On the other hand, in the following two figures, we show the effect of the same speech with a different gesture on a subject's retelling. This subject also heard the narrator say "and Granny whacked him one" but saw the narrator making a slapping gesture (narrator stimulus is shown in Figure 11). The subject said,

> "And Granny gets him and and whacks him and clobbers him."

and with his right hand entirely open, the subject makes a slapping movement (see Figure 12).

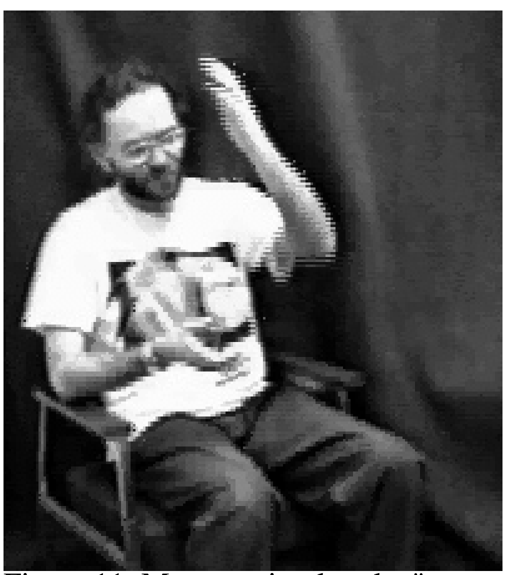

Figure 11: Manner stimulus slap"

Figure 12: Manner response "slap"

Origo

50% of origo mismatch stimuli had an effect on listeners' retellings. An example of a retelling inaccuracy in the context of an origo mismatch is a subject who heard "and then Granny gives him a penny" but saw the narrator make a proffering gesture towards himself. She said,

> "Granny sees him and says "oh what a nice little organ grinder" [and she]
> gets—[goes to give hi]m [a penny]—a little monkey excuse me"

and with her hand first makes a giving gesture towards her listener, and then towards herself and then towards her listener once again (the repetitive nature of the movement in this and the next example make it impossible to include images along with the text).

Anaphor

32% of anaphor mismatch stimuli had an effect on listeners' retellings. An example of a retelling mistake in the context of an anaphor mismatch comes from a subject who saw Sylvester and Tweetie set up in space and, as a description of Sylvester chasing Tweetie, heard "suddenly he lunges for him", but saw the Tweetie hand doing the lunging, and who said:

> Sylvester is in the window of an apt building
> and he's looking at Tweetie
> and Tweetie sees him
> and (pause) Tweetie go- follows him into the
> follows- goes into the apartment
> and (pause) Sylvester lunges at him

while his hands rest on his lap. Note that this person has reconciled information from the two modalities by adding an additional event to the story: Tweetie follows Sylvester into the apartment (derived from the Tweetie hand doing the lunging) before Sylvester lunges (derived from the statement that "he (Sylvester) lunges for him").

Thus, even when the information that gesture conveyed contradicted the information conveyed by accompanying speech (origo and anaphor mismatches), subjects still took gestural information into account. The fact that retelling inaccuracies resulted from manner, origo **and** anaphor mismatches also tells us that the fact of mismatching gesture to speech is not the cause of subjects' attention to gestural information; even when, as is usual, gestural information enhances speech, subjects demonstrate that they have taken gesture into account when representing the narration that they viewed.

Locus of Effect

Although it is clear from the results presented thus far that people do attend to information conveyed by gesture, it is possible that they isolate that information into a gestural representation that does not interact with the linguistic representation (as was assumed by Krauss *et al.*, 1991 when they asked subjects to recognize gestures that they had previously viewed in the context of speech, or alone).

Thus, if gestural information is attended to by listeners, there are still several possibilities for how information conveyed by gesture and that conveyed by speech will interact. If listeners do attend to information conveyed in gesture, but keep it separate from information conveyed verbally, then we might expect to see the subjects' subsequent retellings show the gestures copied and the speech copied. That is, the listener would tell the story as heard, and produce the gesture mismatches as seen. If we found uptake of gesture only by gesture, then this would be evidence for uptake of gestural information, but a strict separation between the two channels; for two distinct representations—one motoric and the other linguistic—of the information conveyed. This would argue against the position that gestural information is used as an equal partner with speech to form a representation of the information conveyed, rather, suggesting that gesture is taken up but not processed as relevant to the content of speech.

If however, the gesture-speech mismatches provoke a new representation—accessible to both speech and gesture—of a particular event of the stimulus story, then information conveyed solely by gesture may find its way into speech. This would arise because gesture and speech are both understood and produced as expressions of different aspects of one single kind of representation.

Of course, we are depending here on retellings as an index of representation, as we don't have direct access to the outcome of the comprehension process. Thus, in order to examine this question, we look at the number of times in which information crosses channels; that is, we compare the number of times in which information received via gesture alone is retold in speech with the number of times in which information received via gesture is re-enacted by gesture.

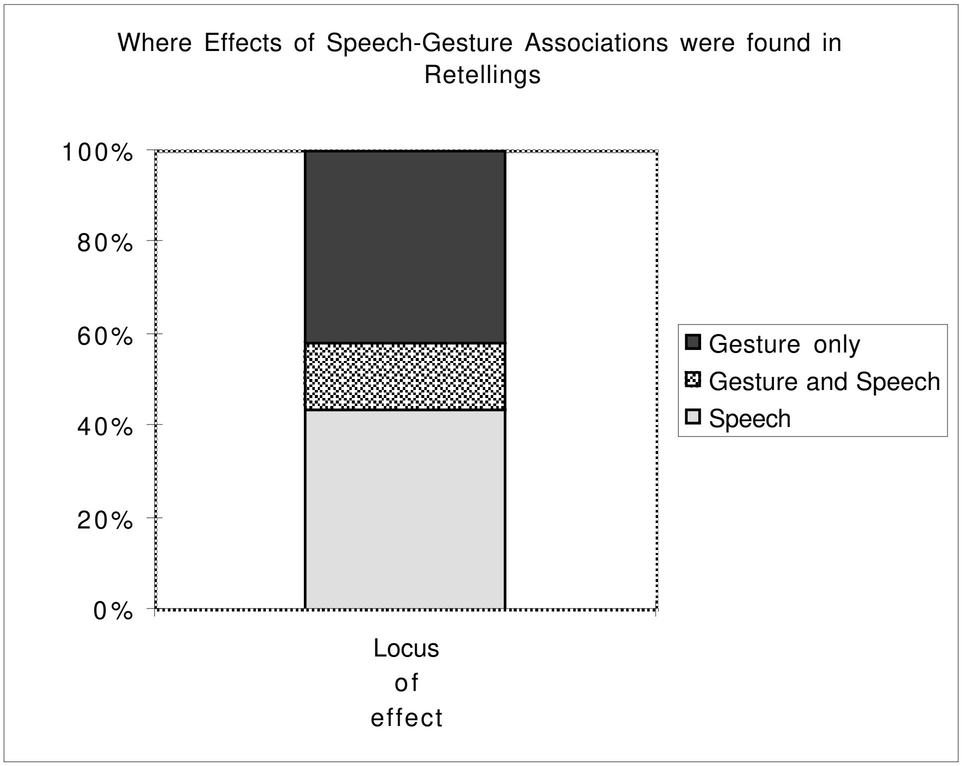

Figure 13

As shown in Figure 13, in fact, of the mean of 6.63 inaccuracies in retelling produced by each subject, 56% were displayed in speech and 63% were displayed in gesture (of which 19% were also displayed in speech)[5]. Thus, it is not the case that information conveyed by gesture is retold in gesture, while information conveyed by speech is re-represented in speech. Instead, in more than half of the cases where a subject re-told information that s/he received only via gesture, that information was retold in a modality other than the one it was received in (i.e. speech). Examples of information received in gesture and retold in speech were given above—all three examples of mismatch type described in the last section demonstrated a lexical choice response to information received only in the gestural modality. The second example, where a subject hesitates between "gets a penny" and "gives a penny" in speech, and movements of giving towards the other and of giving towards the self in gesture, is particularly interesting because it is possible to observe two competing representations arise and then become reconciled. Note that the first gesture is accurate with respect to the speech heard in the stimulus tape, while the first speech is only accurate with respect to the gesture seen; the subject then repairs her speech to what she knows must be the case (given her knowledge of grannies and organ grinders' monkeys) and—perhaps automatically because she is repairing her speech—also changes her gesture to reflect the gesture she saw, but which no longer reflects her own speech. Finally, she produces a third gesture which fits the speech that she herself has at this point produced.

Gesture and speech, in this instance, at first display different representations of narrative information; representations which cannot co-exist and which must be reconciled for the story to continue. In this sense, we have evoked an experimental situation much like the naturally occurring one examined by Goldin-Meadow et al. (1993) where children may describe one representation of math knowledge in speech and another, competing, representation in gesture. Like Goldin-Meadow et al., we conclude that mental representations are equally accessible to both modalities, and that differences between the modalities may trigger new representations. That is, in the case of narrative retelling, it is not the case that motoric and linguistic modalities are displaying information from motoric and linguistic input, respectively.

Of course, the current paradigm only allows us to discern cases where information transmitted through gesture is re-transmitted through speech, and not the inverse. Cases of the latter—information received in speech and retransmitted in gesture—are abundant in the literature. In fact, in our own work, when we have had subjects *listen* to a narration without viewing the video image, we see just as much gesturing in the re-telling as when we have subjects both listen to and watch the narration on the video screen (McNeill, Cassell & McCullough, 1994). It is, however, more critical to the current argument to show that gestural information can give rise to linguistic. This is because only in cases where gestural information gives rise to linguistic retellings can one dismiss the conclusion that gesture is understood as a motoric translation of speech and therefore information conveyed in gesture is discarded or maintained in comprehension separately from information conveyed by speech. Our results provide evidence for the conclusion that gesture coexists with speech at the first level in which represented information is received from communication.

Omission of information

Finally, we asked whether a gesture-speech mismatch might provoke an *omission* of information. "Asynchrony" of gestures and speech has been found to impair the recall of a spoken message (Woodall & Burgoon, 1981) and thus, while no subject reproduces the entirety of the story, we can compare the omission of information for particular events between subjects who saw a gesture-speech fit, and those who saw a gesture-speech mismatch. If gesture and speech are equally likely to contribute to the representation of narrative information then, in cases where gestural and speech information conflict one might expect that listeners, unable to resolve the conflict, would simply omit the description of the event. To look at this issue, we compared how many events were omitted from subjects' retellings when speech and gesture matched, and when they mismatched. These data are shown in
Figure 14.

Figure 14

A mean of 1.9 events was omitted per subject (out of the 21 potential mismatch points examined in the story). Of these omissions, a mean of .9 (.045%) occurred in the context of a gesture-speech mismatch, .5 (.025%) occurred in the context of manner gestures, and a mean of .5 (.025%) occurred in places where speech and gesture matched. There is no reason to think that manner gestures would provoke an omission of information in the retelling—these gestures are common in everyday conversation, and in other experiments in our laboratory are remembered without difficulty. We might, therefore wish to group manner gestures and speech-gesture matches together, which would indicate that information is roughly equally likely to be omitted in cases where gesture and speech are matched and in cases where they are mismatched. The percentages are small and in any case these data are inconclusive concerning the effect of speech-gesture mismatches on memory for narrative events.

Note, however, that if omission of information is an expected consequence of two conflicting representations of narrative information, then we have already seen that there are other ways of dealing with such conflicts. The example of a response to an origo mismatch described above shows that the reteller is able to, after several repairs, recount a plausible scenario for the story. Another subject, who also saw the mismatched 'offer-to-self' gesture along with the phrase "and then Granny gives him a penny", said

> and she threw him a penny
> so he picked up the penny

while gesturing towards herself. This listener was able to continue the story, but she too had to insert additional information that was not in the original to reconcile the information received in gesture with the information received in speech. Prior, schematic, knowledge about stories and in particular about cartoon stories, and most probably even about Sylvester and Tweetie Bird, that most listeners bring to this task, undoubtedly helps in the task of reconciling gestural- with speech-conveyed information, perhaps diminishing the number of events omitted.

Discussion

The present study suggests that in a situation where, as is natural, gesture is produced in conjunction with speech in narrative discourse, listeners attend to the information conveyed via the gestural modality. McNeill (1992), Kendon (1972) and others have shown that there is an intimate temporal, semantic, and pragmatic relationship between gesture and speech in the production of discourse. The data from our study suggest that the semantic relationship between gesture and speech is taken into account by listeners when constructing a mental representation of what has been communicated to them. In particular, in their retellings of a narrative, listeners narrate a version of events that takes into account information conveyed only in gesture (in the case of manner gestures) or that attempts to reconcile conflicting information from speech and gesture (in the case of origo and anaphor mismatches).

The fact that our results hold for gestures that convey contradictory information to the speech they accompany, and for gestures that supplement the information conveyed by speech, supports the claim that these findings are not due to the somewhat artificial situation of having a narrator mismatch his gestures to his speech. Additional support for this claim comes from work by Goldin-Meadow and her colleagues (Church & Goldin-Meadow, 1986; Goldin-Meadow, Alibali & Church, 1993; Goldin-Meadow *et al*, 1992; Alibali *et al*, 1994) that shows that children may naturally produce speech-gesture mismatches when they are at transitional states of knowledge, and untrained adult speakers have the capacity to interpret information that is only conveyed in gesture by these children. Such an ability is important in the context of the classroom since, as these researchers have shown, children may demonstrate in gesture knowledge of concepts that they cannot yet articulate in speech. Our subjects' ability then, to reconcile conflicting information in speech and gesture may not be unexpected.

Using a mismatch paradigm allowed us to maintain the temporal and pragmatic relationships between speech and gesture while varying the semantic relationship between the two media. It also allowed us to look at uptake of gesture in a natural context: the understanding of information conveyed by a narrator whom the listener can see, who gestures and speaks naturally.

It is undoubtedly because we manipulated the semantic relationship between gesture and speech, but maintained the temporal and pragmatic relationship between them, that our results and our conclusions differ from those of researchers such as Krauss and associates (1991) who asked listeners to *identify* gestures they had seen before (a task in which performance diminished when gestures were presented along with their linguistic context). If the meaning conveyed by gesture is used along with speech to build a single representation of what has been said, then we would predict that recognition of (the surface form of) gestures would be adversely affected by appearing in their communicative context because the surface form of the gestures would have been lost while their meaning was retained[6].

Gestures are less codified than speech and, unlike speech, do **not** engage in consistent, socially ratified form-meaning pairings. The index finger pointing up may sometimes convey that one is referring to the sky, and sometimes convey that one is describing someone climbing upwards. Therefore, we do not expect gestures to have decontextualized transparent meanings (cf. Tuite, 1993). Their interaction with speech is more like the interaction of a schematic diagram (arrows and boxes, for example) with the instructional text it accompanies in the automatic generation of multi-media text, than like the interaction of one word with another[7]. The two media display different aspects of one single idea.

Finally, our study, as Alibali et al. (1994), has shown that the representation that listeners form of gestural information is accessible to subsequent communication in both words and gestures. Information received via gesture was not re-told only in gesture but also crossed channels and found its way into the re-teller's speech. This is evidence for the integral role that gesture plays in comprehending language: we believe that the gesture and speech of the narrator are decoded by the listener and the information gleaned from both used to build a unified representation of the story. In terms of ‘growth points’ (McNeill, in press) the listener, in understanding what was being said, matched or nearly matched the speaker’s growth points. The retelling, as with all of the narrative and cartoon retellings that we have examined over the years, contains both speech and gesture. What information is conveyed in speech and what in gesture, however, obeys the constraints of narrating in general, and not any constraints of what medium the information was received in. Thus, it appears that gesture is not kept modularly distinct from speech in the comprehension of narrative since a subsequent retelling of events may communicate information in speech that was received in gesture, and vice-versa.

Of course, as described above, only the narrative functions of gestures that concern plot-line information were addressed here. It remains to be seen whether listeners attend to the discourse structure conveyed by gesture; that is, information about speaker's perspective and stance on events. In order to address these issues one must mismatch the pragmatic relationship between gestures and speech; for example accompanying narrative level speech about plot-line events with beats. We are attempting to carry this out by building an animated cartoon narrator in whom the temporal relationship between speech and gesture can be manipulated (Cassell et al., 1994).

Likewise, here we have concentrated on the role of gesture in comprehension and not on the relationship between gesture and speech in production. This latter is more difficult to examine experimentally. However, the work by Goldin-Meadow and her colleagues on naturally occurring gesture - speech mismatches, along with examples such as that described in the results section on origo mismatches, where a subject clearly attempts to reconcile two competing representations of events, one generated in gesture and one generated in speech, does lend support to the position that gesture plays a primary role in the production of language.

In sum, in demonstrating that speakers' gestures are integrated into hearers' representations of the information conveyed, this work supports the position that gesture and speech are two integrated systems, allied from the very moment they are viewed to the last stage of their comprehension and use in constructing a model of what was said. The present study suggests that when people listen to a story, they are also watching.

References

Alibali, M.W., Flevares, L. & Goldin-Meadow, S. 1994. *Going Beyond What Children Say to Assess their Knowledge*. Manuscript submitted for publication.

Bavelas, J. (1994). “Gestures as Part of Speech: Methodological Implications”. *Research on Language and Social Interaction* 27, 3: 201-221.

Butterworth, B. & Beattie, G. 1978. “Gesture and silence as indicators of planning in speech”. In R.N. Campbell & P.T. Smith (eds), *Recent Advances in the Psychology of Language: Formal and Experimental Approaches*. New York: Olenum Press.

Butterworth, B. & Hadar, U. 1989. “Gesture, speech, and computational stages: A reply to McNeill”. *Psychological Review* 96, 168-174.

Cassell, J. & McNeill, D. 1990. “Gesture and ground”. In K. Hall, J.-P. Koenig, M.Meacham, S.Reinman, & L.A. Sutton (eds), *Proceedings of the 16th Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society, 57-68.

Cassell, J. & McNeill, D. 1991. “Gesture and the poetics of prose”. *Poetics Today*, 12(3): 375-404.

Cassell, J., Steedman, M., Badler, N., Pelachaud, C., Stone, M., Douville, B., Prevost, S., & Achorn, B. 1994. “Modeling the interaction between gesture and speech”. In *Proceedings of the Cognitive Science Society Annual Conference*.

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

Cohen, A.A. 1977. “The communicative functions of hand illustrators”. *Journal of Communication* 27(4): 54-63.

Cohen, A.A. & Harrison, R.P. 1973. “Intentionality in the use of hand illustrators in face-to-face communication situations”. *Journal of Personality and Social Psychology* 28: 276-279.

Crowder, E. & Newman, D. 1993. “Telling what they know: the role of gesture and language in children's science explanations”. *Pragmatics and Cognition* 1(2): 341-376.

Ekman, P. & W.V. Friesen. 1969. “The repertoire of nonverbal behavioral categories—origins, usage, and coding”. *Semiotica* 1: 49-98.

Feyereisen, P. 1983. “Manual activity during speaking in aphasic subjects”. *International Journal of Psychology* 18: 545-556.

Freedman, N. 1972. “The analysis of movement behaviour during the clinical interview”. In A. Siegman & B. Pope (eds), *Studies in Dyadic Communication*. New York: Pergamon 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*. Cambridge, England: Cambridge University Press.

Goldin-Meadow, S., Alibali, M., Church, R.B. 1993. “Transitions in concept acquisition: Using the hands to read the mind”. *Psychological Review* 100 (2): 279-297.

Goldin-Meadow, S., Wein, D. & Chang, C. 1992. “Assessing knowledge through gesture: Using children's hands to read their minds”. *Cognition and Instruction* 9 (3): 201-219.

Greenbaum, S. & Quirk, R. 1970. *Elicitation Experiments in English: Linguistics Studies in Use and Attitude*. London: Longman.

Hinrichs, E. & Polanyi, L. 1986. “Pointing the way: A unified treatment of referential gesture in interactive contexts”. In A. Farley, P. Farley & K.-E. McCullough (eds), *Proceedings of the Parasession of the Chicago Linguistics Society Annual Meetings (Pragmatics and Grammatical Theory)*. Chicago: Chicago Linguistics Society.

Hopper, P. 1979. “Aspect and foregrounding”. In T. Givón (ed), *Syntax and Semantics, Volume 12: Discourse and Syntax*. New York: Academic Press.

Hopper, P. & Thompson, S.A. 1980. “Transitivity in grammar and discourse”. *Language* 56: 251-299.

Kendon, A. 1972. “Some relationships between body motion and speech”. In A. Siegman & B. Pope (eds), *Studies in Dyadic Communication*. New York: Pergamon Press.

Kendon, A. 1988. “How gestures can become like words”. In F. Poyatos (ed), *Cross-cultural Perspectives in Nonverbal Communication*. New York: C.J. Hogrefe.

Kendon, A. 1992. “Some recent work from Italy on quotable gestures ('emblems')”. *Journal of Linguistic Anthropology* 2(1): 72-93.

Krauss, R., Morrel-Samuels, P. & Colasante, C. 1991. “Do conversational hand gestures communicate?”. *Journal of*

Personality and Social Psychology 61(5): 743-754.

McCullough, K.-E. 1992. “Visual imagery in language and gesture”. *Annual Meeting of the Belgian Linguistic Society*. Brussels (November).

McNeill, D. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.

McNeill, D. (in press). “Growth points cross-linguistically”. In J. Nuyts & E. Pederson (eds), *Language and Conceptualization*. Cambridge: Cambridge University Press.

McNeill, D. to appear. “Models of Speaking (to the Amazement) Meet Speech-Synchronized Gestures”. In McNeill (ed) *Language and Gesture: Window into Thought and Action*.. 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*..

Pedelty, L. L. 1987. “Gesture in Aphasia”. Ph.D. dissertation, Department of Behavioral Sciences, University of Chicago.

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

Riseborough, M.G. 1982. “Meaning in movement: An investigation into the interrelationship of physiographic gestures and speech in seven-year-olds”. *British Journal of Psychology* 73: 497-503.

Rogers, W.T. 1978. “The contribution of kinesic illustrators toward the comprehension of verbal behavior within utterances”. *Human Communication Research* 5: 54-62.

Sachs, J. 1967. “Recognition Memory for Syntactic and Semantic Aspects of Connected Discourse”. Ph.D. dissertation, University of California.

Short, J., Williams, E., & Christie, B. 1976. *The Social Psychology of Telecommunications*. New York: Wiley.

Streeck, J. 1993. “Gesture as communication I: Its coordination with gaze and speech”. *Communication Monographs* 60 (4): 275-299.

Thompson, L.A. & Massaro, D.W. 1986. “Evaluation and integration of speech and pointing gestures during referential understanding”. *Journal of Experimental Child Psychology* 42: 144-168.

Tuite, K. 1993. “The production of gesture”. *Semiotica* 93 (1/2): 83-105.

Williams, E. 1977. “Experimental comparisons of face-to-face and mediated communication: A review”. *Psychological Bulletin* 84: 963-976.

Woodall, W.G. & Burgoon, J.K. 1981. “The effects of nonverbal synchrony on message synchrony and persuasiveness”. *Journal of Nonverbal Behavior* 5: 207-223.

Notes

[1] Preparation of this article was supported in part by National Science Foundation Visiting Professorship for Women GER 9350179 to Justine Cassell; and NSF BNS 8518324 and National Institute for Deafness and Other Communication Disorders DC 01561 to David McNeill.

Many thanks to Brecky Church and Kevin Tuite for participating in the design of the experiments, to Laura Stachnik and Julie Amato for collecting and transcribing data, to Mike Kelly and Sue Marcus for help with statistical analyses, and to Nicola Bessell, Susan Goldin-Meadow, and Mark Steedman for insightful comments.

Correspondance should be sent to Justine Cassell, at the Media Lab, Massachusetts Institute of Technology, E15-315, 20 Ames Street, Cambridge MA 02139-4307.

[2] The narrator's speech is represented above the accompanying gestures; square brackets in the speech indicate the extent of the gesture (which words were accompanied by gesture).

[3] In particular, these instructions are identical to those used for other previous experiments in our laboratory where subjects watched a video of a cartoon; we therefore do not believe that the instructions particularly drew subjects' attention to the videotaped narrator's hands.

[4] Why do the coders need to be aware of the form of the two gesture stimuli, one of which the subject has seen? For two reasons: First,because subjects' retellings **never** reproduce exactly the speech that they heard, the coders must be given a direction in which to look for inaccuracies. A subject may interject that Tweety is a canary, even though she did not hear this in the stimulus narration. This addition is not counted as an instance of gesture influencing the understanding of the narration because it is not in the direction of either of the gestures. If gesture has no effect on understanding, then, because the coders are blind to the version of the stimulus narration seen by a given subject, they are likely to ascribe inaccuracies equally to the normal and mismatched version of a given segment. Only if gesture has influenced the subject's understanding of the narration will more gesture-provoked inaccuracies be found in the mismatched version. The second reason stems from the fact that, for the manner mismatches, in both versions of the stimulus narration the gesture gave information that supplemented that provided by speech. Thus, retellings that take into account the information conveyed by manner mismatches will be inaccurate for both versions; in order for the coders to correctly judge whether there was uptake of gestural information, they must know which gesture was seen.

[5] Thus, percentages do not sum to 100% because of a mean of 1.25 retelling inaccuracies, or 19% of inaccuracies, were apparent in both speech and gesture.

[6] In fact, the situation is no different than has been found for speech. Six unrelated words are easily recalled, whereas people may have difficulty recalling the exact form of the words in a sentence of equal length (Greenbaum & Quirk, 1970, Sachs, 1967). The reason is that in the second case, people have extracted semantic interpretations, and have not retained the form of the words.

[7] Thanks to Matthew Stone for this insight.