Not Channels But Composite Signals: Speech, Gesture, Diagrams and Object Demonstrations Are Integrated in Multimodal Explanations

Randi A. Engle (randi@csli.stanford.edu)
Stanford University School of Education
Cognition and Learning Lab, CERAS 105
Stanford, CA 94305-3084 USA

Abstract

This paper provides empirical evidence that multimodal signals are produced and understood as integrated units of communication called composite signals, rather than being independently interpretable "channels" of communication. I propose that using composite signals relies on two communicative norms, co-expressivity and consistency:

- *co-expressivity:* each element of a composite signal refers to the same underlying referent
- *consistency:* elements of the same composite do not contradict each other.

This paper will show that these norms are consistent with data comprising a set of explanations of how locks work in which participants spoke while gesturing, drawing diagrams, and manipulating a sample lock. *Co-expressivity* is supported by the fact that co-expressive speech segments can be found in nearby speech for communicative nonverbal behaviors but not for non-communicative nonverbal behaviors. *Consistency* is evidenced in inferences that maintain number and modality consistency in cases of apparent contradiction.

Introduction

This paper is part of an effort to develop a general theory of multimodal communication. Recently there has been significant research on how gesture and speech are coordinated in face-to-face communication (e.g. Bavelas, 1994; Kendon, 1980; McNeill, 1992). I am extending this work on speech and gesture to a wider range of multimodal communication: cases where speech is also used in coordination with diagram drawing, object manipulations, and gestures that are superimposed over diagrams and objects. For example, I have found that synchronization patterns seen for gesture with speech generally apply to this wider range of multimodal signals (Engle, 1998).

This paper will develop the hypothesis that multimodal communication is achieved via integrated communicative units which Herbert Clark and I call composite signals (Clark, 1996; Engle & Clark, 1995). This hypothesis argues against the common assumption (as discussed in Sanders, 1987) that communicative modalities like speech, gesture, and diagrams are independently interpretable communicative "channels".[1] The composite signal hypothesis is an extension to multimodal communication of claims by Bavelas (1994), McNeill (1992), Sanders (1987), and others that gestures are "integrated" with speech. If composite signals that combine language with nonverbal signals do exist, understanding how composite signals are created and interpreted will be a key element in a more complete understanding of both language and communication.

Theory

I argue that composite signals are integrated in that communicators and their addressees jointly assume that all signals in a particular composite signal are intended to be treated from the start as contributing to a single, unified interpretation. This is a fundamental communicative norm that is in effect whenever more than one method or modality is used to communicate about some topic. It is a claim of the same sort as Grice's (1967/1989) cooperative principle. In most situations, we should expect people to try to be cooperative and to present multimodal signals that are easily integrated. However, in any particular case, a communicator may not be sufficiently skillful to produce a composite that can be integrated into a single interpretation. In such cases we should expect evidence that the composite did not successfully communicate (e.g., the communicator repairs his/her original version, an addressee asks for clarification, or addressees are shown to have misunderstood later). In addition, a communicative norm can be obtrusively not followed in order to lead addressees to infer something else (e.g., that simultaneously more than one conversation is occurring—one carried by conventional hand signals, the other by talk) (Grice, 1967/1989).

I will make two specific claims about composite signals:

1. Elements of Composite Signals Are Co-expressive

First, the elements of each composite are in McNeill's (1997) terms "co-expressive". Each element of a composite

[1] This alternative claim is usually assumed as self-evident rather than explicitly argued for. For example, in testing whether iconic gestures are communicative, Feyereisen, van de Weile & Dubois (1988) had raters watch videotaped gestures with the sound turned off, questioning whether they were communicative because raters were not consistent at attributing meanings to them. By an integrated view, it is assumed that when iconic gestures communicate they do so *in coordination* with accompanying speech. Since iconic gestures do not have conventional meanings, it is not surprising that they do not have independent, reliable interpretations separate from integration with conventional symbols provided in speech.

(whether it be a single gesture, an addition to a diagram, or a segment of speech) is assumed to refer to the same underlying referent (i.e., the same object, event, relation, etc.). The elements of a composite may do this by being completely redundant with each other, or they may provide complementary and only slightly redundant information about a larger referent that neither could establish alone[2].

In addition, the "integrated meaning" supported by a composite signal can vary. In my data, which are explanations of how locks work, most composites helped to update the construction of analogical models. The analogical models were spatially anchored nonverbally by diagrams, gestures or a sample lock, with speech primarily providing conceptual perspectives on the resulting models. A few composite signals in the beginning of the explanations seemed designed to help learners know what category of prior knowledge to access for understanding the explanation, with nonverbal signals used to exemplify the category[3]. In other work, gesture-speech pairs have been shown to combine to form indirect requests (Barr & Kelly, 1997).

2. Elements Are Assumed to be Consistent

Second, each element of a composite is assumed to provide consistent information about the underlying referent. For example, what is communicated in a diagram should not contradict what is said about the diagram. And generally it does not. However, if one interprets each element separately, it is sometimes possible to identify apparent contradictions between what was conveyed verbally and nonverbally. In most cases, these apparent contradictions would not arise if one assumed from the start that the elements are co-expressive and consistent, and guided interpretive inferences accordingly[4]. When contradictions remain despite assuming co-expressivity and consistency, we should expect communication failures. When a videotaped narrative was designed to artificially have gestural content not match coinciding speech, listener's retellings included both inferences maintaining consistency and errors (McNeill, Cassell & McCullough, 1994).

Method

Six undergraduates examined a sample lock and studied a written explanation of how locks work. Each undergraduate then individually explained locks to a series of three undergraduates who did not know how they work. For one explanation, the sample lock was available; for another, paper and pen were; for a third, no additional material resources were available. Thus, the data includes not only gestures with speech, but also the kinds of multimodal signals mentioned previously: diagram-drawing, manipulations of a sample lock and key, and gestures superimposed over diagrams or the lock system.

Order of material resource availability was counterbalanced. Each explainer-learner pair was told to stop only once both participants felt that the learner understood how locks work well enough to explain them to someone else. All 18 explanations were videotaped. One camera focused on the explainer, another on the learner; and a third was focused down on the table where participants gestured, drew diagrams, and manipulated the lock.

Analysis

Transcription

Explainers' and learners' speech was coded for pauses, false starts, emphatic stress, and intonational contours. All observable hand and other body movements–which included all gestures as well as any manipulations of pen, paper, or lock–were described.

Sampling

One episode from each explanation was selected for intensive analysis. Each episode was a coherent piece of the explanation that had been collaboratively grounded as a "conversational contribution" (Clark & Schaefer, 1989) and for which there was evidence later in the interaction about how well the learner understood the information presented. Beyond fulfilling these criteria, episodes were selected to maximize variation in communicative methods and information presented in order to capture as much of the variability in the larger dataset as possible. On average, each episode contained 7.1 spoken intonation units and 8.1 completed nonverbal behaviors, resulting in a total sample of 127 speech units and 146 nonverbal behaviors.

Ratings of Content Communicated

Pairs of independent raters watched the videotape up to the start and then up to the end of each selected episode. Each rater constructed representations of what had been communicated about the lock before the episode started and once it had finished. By comparing these start and end representations it was possible to measure what new information had been communicated in each episode.

One representation raters constructed was a diagram of the lock system. The second was a propositional style representation in which the rater specified what components of the lock system had been explained and what each of their so-far specified static properties (e.g., shape, location, etc.), kinematic properties (e.g., direction, trajectory, etc. of any part that moves, is able to move or is not able to move) and

---

[2] An example of complementarity from my data: in explaining how a lock works, an explainer wanted to establish where the cotter pins were located inside the cylinder. In speech, the explainer fixed 2 dimensions of their location (their location in a horizontal/vertical plane) while representing the 3rd dimension and one of the previous dimensions with a gesture over the lock (their location in the vertical/horizontal plane). In this case, though each element alone only represented two dimensions of the pins' positions, the composite as a whole represented their full 3D location with respect to the cylinder.

[3] For example, one explainer began his explanation by saying "it's a lock" while demonstrating himself putting an imaginary key into a door lock. This made clear what kind of lock would be explained, allowing the learner to access relevant knowledge.

[4] Examples of this will be discussed in the results.

causal relations (static or kinematic properties of objects in the lock system that cause other parts to move, be able to move or not be able to move in some way) were. Raters then constructed jointly agreed upon start and end diagrams and propositional descriptions of the lock system by comparing their individual diagrams and descriptions, and often watching the videotape again.

Communicative Status & Interpretability Codings

The investigator and one of the independent raters coded the nonverbal behaviors either as communicative about locks or not (i.e., meta-communicative, preparatory, or irrelevant). To be communicative, a behavior had to be more than simply informative about locks; its performance also had to be treated in the interaction at that moment as part of the explainer's intended presentation (see Grice, 1957). Of the 147 nonverbal behaviors, 108 were coded as communicating something about locks, 35 were coded as not communicating anything about locks and 3 were coded as unclear (89% agreement, with a third rater used to resolve difficult cases). Of the 108 nonverbal signals about the lock, 43 were hand gestures performed alone[5]). The remaining 65 involved the lock or diagram in some way:

- 7 involved drawing or writing something on paper;
- 10 were manipulations of the lock or key;
- 27 involved a diagram with a gesture superimposed over it;
- 20 involved the lock with a gesture superimposed over it;
- 1 combined drawing with gesturing over the diagram.

For each nonverbal behavior coded as communicative, an independent rater who had participated in the content ratings coded its meaning as either clear or unclear given its context.

Timing

Beginnings and endings of nonverbal behaviors were timed with respect to speech, to the nearest syllable. Behaviors were considered to begin as soon as a hand moved into diagram, lock or gesture space[6] and to end as soon as it began leaving diagram/lock/gesture space or initiated a new movement. For communicative acts, this criterion singles out when meaningful content is being conveyed by excluding preparatory movements. For gesture, this criterion has the "gesture proper" consist of the stroke and post-stroke hold phases, the two phases found to often coincide with co-expressive speech (Kendon, 1980; McNeill, 1992, 1997).

Results

Evidence for the Norm of Co-expressivity

1. Co-expressive Speech Is Identifiable for Communicative Nonverbal Behaviors. The intonation units just before and during each nonverbal signal rated as communicative about the lock were examined to see if they contained a speech segment co-expressive with the nonverbal signal. To count as co-expressive, both the speech segment and the nonverbal signal were interpretable as both referring to the same underlying referent. For all but one of the 108 nonverbal signals, a co-expressive speech segment could be found within this two intonation unit time window[7].

Synchronization of the nonverbal and verbal elements of composite signals was usually even more precise than that (see also Engle, 1998). Seventy-five percent of nonverbal signals[8] exactly coincided with a segment of co-expressive speech while another six percent[9] followed immediately after a co-expressive speech segment, results consistent with (and thus extending) the gesture-speech synchronization literature (Kendon, 1980, 1988; McNeill, 1992, 1997). In sum, communicative nonverbal behaviors are generally performed near a segment of co-expressive speech that they could be in the same composite signal with.

2. Co-expressive Speech Is Often Not Identifiable for Nonverbal Behaviors Rated as Not Communicative About Locks. Of the 35 nonverbal behaviors rated as not communicative, 17 were considered by the investigator to be behaviors that could have potentially communicated something about the lock in other communicative situations. For example, consider the action of putting a key into a lock. In some situations, explainers used this to demonstrate what happens when the key is put into the lock (e.g., the key's wedges can be heard pushing the cotter pins up). But in other situations, putting the key in the lock was used simply to get the lock into the right state for demonstrating something else (e.g., to prepare for demonstrating what happens when the key is turned). In such situations, putting a key in the lock *could* have played a communicative role, but in fact it did not.

How do people distinguish when nonverbal behaviors are playing a communicative role versus when they are only preparatory actions that support later communication? One factor appears to be whether or not a co-expressive speech segment can be found. If any of the 17 potentially communicative behaviors had in fact been communicative, it should have been possible to find a segment of speech co-expressive with its likely interpretation. So, for the example of a key being put in the lock, one could look for co-expressive speech that made some reference to something that happens when the key is put in the lock. Using this likely interpretation for each of the 17 potentially

[5] 31 of these 43 gestures were made in the no material resources condition when neither paper nor the lock were available to use.

[6] A hand is in diagram space once it or the pen is touching the portion of the paper's surface on which the diagram is being drawn. A hand is in lock space once it is touching the lock or key. A hand is in gesture space once it is in a location observable by both the gesturer and the addressee and from where the gesture proper can then be initiated.

[7] The one exception was drawing a line to represent a pin in a diagram after having twice earlier indicated that the "pin" was located along that line. To see the line as a "pin" was also confirmed in the next intonation unit when the line was traced over and extended slightly.

[8] 78% of nonverbal signals involving the lock or diagram while 70% of gestures not involving the lock or diagram.

[9] All of these involved the lock or diagram in some way.

communicative nonverbal behaviors, the current and immediately preceding intonation units were again examined for co-expressive speech. In stark contrast to communicative nonverbal signals, in 14 of these 17 non-communicative cases, no co-expressive speech was present[10].

Putting the last two sets of results together, we see that when a nonverbal behavior is likely to be part of a composite signal with speech (i.e., when the behavior is communicative about the topic of speech), it is easy to find a speech segment that can be interpreted as referring to the same referent. In contrast, when a nonverbal behavior cannot be part of a composite (i.e., when it is not communicative) a co-expressive speech segment is less likely to be available.

3. Noncommunicative Behaviors Are Desynchronized. Speech coinciding with each nonverbal behavior was coded as independently referring to something about the lock or as not doing so, either because no speech coincided or the speech segment was too incomplete to fix reference. The 35 nonverbal behaviors rated as not communicative about the lock were significantly more likely to occur without speech or be synchronized with such incomplete segments (overall: $\chi^2=61$, df=1, $p < .001$; for hand movements only: $\chi^2=30$, df=1, $p < .001$; for behaviors with diagram or lock: $\chi^2=32$, df=1, $p < .001$).

Evidence for the Norm of Consistency

1. Potential Contradictions Are Infrequent. Interpreting speech and nonverbal elements relatively independently, I looked for two kinds of potential contradictions between the 108 nonverbal signals and associated co-expressive speech segments in my data: number and modality contradictions. A number contradiction was coded if the same entity (object, event, etc.) was represented in both nonverbal signals and co-expressive speech and either: speech used a plural form while only one entity was shown nonverbally; or speech used a singular form but more than one entity was shown nonverbally. Modality contradictions were coded when a motion event was represented as being impossible (e.g. "can't move"), currently occurring (e.g., "moves"), or possible but not currently occurring (e.g., "can move") in speech but represented in a different modality nonverbally. Of the 108 nonverbal signal and co-expressive speech pairs in the sample, there were only 15 potential number inconsistencies and 2 potential modality inconsistencies.

2. Many Potential Contradictions Never Arise Because of Implicature-Style Inferences. In all but one of the 17 cases of apparent contradiction, it is easy to make an inference that re-interprets either the spoken or nonverbal element (or both) so that a consistent interpretation can be arrived as from the start.

Though not crucial for my argument, these inferences can be thought of as Gricean-style (1967/1989) implicatures by extending a subset of Grice's maxim of quality "do not say what you mean to be false" from purely linguistic contradictions like "the key is moving but it's not" to cases in which a proposition and its negation are distributed across modalities. Rather than being cases where a maxim is flouted, most of my cases are what Levinson (1983) calls "standard implicatures", cases where the communicator "implicates that which he must be assumed to believe in order to preserve the assumption that he is observing the maxim" (Grice, 1967/1989).

Prototype implicatures. Of the 15 apparent number inconsistencies between spoken and nonverbal elements, 13 do not arise if one assumes that the nonverbal signal was depicting a prototype that represents the characteristic features common to some set of entities described in speech.

Table 1 below is an example in which more than one pin is referred to in speech but only one pin is depicted nonverbally, resulting in a prototype implicature:

Table 1: Transcript of Paper-2B #11.2.[11]

nonverbal behaviors	speech
in one motion, explainer traces pen over an already drawn vertical line on the keyface diagram & draws in a new segment of the line so that it extends down into the cylinder	*11.2* these **pins**,

In this example, a diagram now includes a single vertical line that represents a pin and its location with respect to the rest of the objects visible on the keyface of the lock: an inner cylinder, a keyhole cut into it and an outer cylinder that encases the inner cylinder. Though only one pin is drawn, speech makes reference to more than one ("these pins"). This appears to be a contradiction, but is not if the single drawn pin is assumed to be depicting a prototypical pin. In that case, the single pin shown represents part of what the explainer means to be the case for all of the pins. Interpretation of the speech may also be affected. "These pins" may be assumed to refer to *this type of pin* rather than the more standard interpretation *these particular pins*, also helping to maintain consistency[12]. Based on these inferences, a model of the lock can be generated that has all of the pins include the common features depicted in the prototype (esp. their orientation & 2D location with respect to the keyface).

Prototype implicatures are not limited to drawing while speaking. The next page presents an example of a gesture superimposed over the lock that also supports a prototype implicature (Table 2):

[10] In the three cases where a co-expressive speech segment could be found, each speech segment appeared to be already linked into a composite with a nonverbal signal that coincided with it.

[11] Legend for transcript tables:
- actions in the same row occurred simultaneously
- actions performed later appear below earlier actions
- nonverbal behaviors rated as communicative about locks in plain text, those not communicative about locks in italics
- bold indicates emphatic stress; hyphen a self-cutoff

[12] Thanks to an anonymous reviewer for pointing this out.

Table 2: Transcript of Lock-3C #20.

nonverbal behaviors	speech
(none)	*20* and
explainer picks up lock & tilts the keyface up towards the learner	the pins are
traces right index finger down keyface of lock, from top of stator to keyhole	going down like that.

I am claiming that in these and eleven similar cases, it is reasonable to assume that the composite signal is referring to a prototype. Besides maintaining consistency of number and allowing participants to construct models, the existence of these prototype implicatures shows that when nonverbal representations like diagrams and gestures are used with language, they are not limited to depicting specific situations as is sometimes assumed. When used in composite signals with language, nonverbal signals can be used to depict types of situations, like prototypes.

Exemplar implicature. In one of the two remaining number inconsistencies, it is reasonable to suppose that the single entity that has been indicated nonverbally is an exemplar of the set of items described in speech (Table 3):

Table 3: Transcript of Lock-6A #64.

nonverbal behaviors	speech
with nail of left index finger, Explainer points to the top of the wedge of the key	*64* these things,
(none)	are just to push it into the groove.

In this case, the top of the frontmost wedge is supposed to be an example of one of the kinds of things that push each pin into each groove of the key. Given that the rest of the key's wedges are visible, the learner can use this single point to one wedge-top to find the rest of the wedge-tops being referred to linguistically.

Notice that in both prototype implicatures and exemplar implicatures, speech uses a plural while the nonverbal signal only depicts or indicates a single entity. An important question is how it becomes clear which inference to draw in a specific situation. One hypothesis consistent with my data is that exemplar implicatures will be drawn when the nonverbal signal draws attention to the actual entity being referred to (in this case, pointing to the actual key) while prototype inferences will be drawn when the actual referent is not present, forcing the nonverbal signal to only selectively represent the entity (as in iconic gestures and pictorial diagrams)[13]. Because of their potentially wide range of usefulness, it may also be true that both kinds of implicatures are at least partially conventionalized.

Implicatures resolving apparent modality contradictions. In one of the potential modality contradictions, the explainer said an object "can't move" while showing the object moving nonverbally. Since the non-existence of a motion cannot be shown nonverbally, the implicature is that the explainer is showing how the object would move if it had not been prevented from moving.

In the other apparent modality contradiction, the explainer talked about the key turning "to the left" while holding his fist stationary. Given that the fist's position is already twisted to the left, the implicature is that the gesture is showing the end state of the motion of turning the key to the left. In a larger sample, other kinds of implicatures would probably arise.

I hypothesize, consistent with other work on implicatures (see Gibbs, 1989), that participants assumed from the beginning that information represented in both modalities would be consistent and made any additional assumptions necessary to preserve consistency, instead of having to detect inconsistencies and then resolving them afterwards. In fact, in *none* of the cases of implicature posited above was there *any* evidence—from later in the interactions or from the ratings of what had been communicated—that a contradiction had been detected. Ratings were consistent with the hypothesized inferences, there were no repairs of understanding initiated by either explainers or learners, and there was no evidence from their written and oral re-explanations that learners misunderstood these composites.

3. An Example of an Unsuccessful Composite. In contrast to all other cases, in the one remaining contradiction (in number), the contradiction may have been detected as the speaker repaired it. As seen in Table 4 below, the original co-expressive speech segment (#8) referred to a single thing that would be heard in the lock demonstration, when in fact 3-4 separate sounds were easily distinguishable:

Table 4: Transcript of Lock-2A #8-9.

nonverbal behaviors	speech
(none)	*8* you can probably hear it,
explainer slowly moves the key into the keyhole: first 3 separate clicks can be heard	(long pause)
as key finishes sliding in, a 4th click can be heard	*9* you can
(none)	hear each pin go up?
(none)	Learner: mhm

The explainer repaired the original co-expressive speech segment (#8) by cutting into the middle of the demonstration to add speech segment #9 which modified the singular "it" to "each pin go up". Using the expression "each pin" resolves all potential number contradictions because this is a way of referring to a group of things while considering them individually. Because the explainer made

[13] For those familiar with Peirce's (1940) distinction between icons and indexes, exemplar inferences are drawn when an icon represents the referent while protoype inferences are drawn when there's an index that goes directly to the referent. See Clark (1996; Clark & Gerrig, 1990) for further discussion of how the icon/index distinction is being used here.

this and other clarifications in segment 9[14], apparently the explainer decided[15] that segment 8 was not sufficiently consistent with the lock demonstration to be a composite signal that the learner would understand and acknowledge (as the learner did soon after segment 9).

Discussion

The data patterns presented above do not support the common assumption that nonverbal communicative modalities always function as separate channels that convey meaning independently from speech. The existence of co-expressive speech near every nonverbal signal in the lock explanations is consistent with the idea that speech and nonverbal signals are to be integrated, though it could be argued from the separate channels view that the similarity in content only arises because explainers were thinking about a single topic which happened to be expressed in both verbal and nonverbal channels. What a separate channels view would find difficult to explain are the desynchronization and implicature results. If whether and what nonverbal behaviors will communicate is always independent of what is happening in speech, why did explainers desynchronize noncommunicative behaviors from speech segments referring to the lock? And if nonverbal signals are interpreted independently from speech, why would the nonverbal signal convey things on the same topic that on the surface appear to be contradicted by speech? Finally, how would prototype and exemplar implicatures be drawn if nonverbal signals were not to be integrated with speech? Until these questions are answered, it cannot be claimed that communication can always be divided into the linguistic and the nonlinguistic, which can each be studied and understood separately[16].

Acknowledgments

This research was supported by a dissertation fellowship and a research grant from the Spencer Foundation, an NSF graduate fellowship, and an APA dissertation research award. Of course, the claims made here are made by the author alone and are not necessarily the positions of the funders. Thanks to Gertraud Benke, Herb Clark, Karen Emmorey, Jim Greeno, Cathy Lachapelle, Muffie Wiebe, and two anonymous reviewers for comments on earlier drafts.

References

Barr, D., & Kelly, S. (1997). Nonverbal factors in understandings and remembering indirect requests. In M.G. Shafto & P. Langley (Eds.), *Proceedings of the Cognitive Science Society.* Mahwah, NJ: Erlbaum.

Bavelas, J.B. (1994). Gestures as part of speech: Methodological implications. *Research on Language and Social Interaction, 27,* 201-221.

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

Clark, H.H., & Gerrig, R.J. (1990). Quotations as demonstrations. *Language, 66,* 764-805.

Clark, H.H., & Schaefer, E.F. (1989). Contributing to discourse. *Cognitive Science, 13,* 259-294.

Engle, R.A. (1998, May). Gesture-speech synchronization patterns may extend to multimodal communication more generally. Poster presented at the *10th Annual American Psychological Society Convention*, Washington, DC.

Engle, R.A., & Clark, H.H. (1995, March). Using composites of speech, gestures, diagrams, and demonstrations in explanations of mechanical devices. Paper presented at the *American Association for Applied Linguistics* meeting, Long Beach, CA.

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

Gibbs, R.W. (1989). Understanding and literal meaning. *Cognitive Science, 13,* 243-251.

Grice, H.P. (1967/1989). Logic and conversation. *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Grice, H.P. (1957). Meaning. *Philosophical Review, 67,* 377-88.

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

Kendon, A. (1988). How gestures can become like words. In F. Poyatos (Ed.), *Cross-cultural perspectives in nonverbal communication.* Toronto: Hogrefe.

Levinson, S.C. (1983). *Pragmatics.* Cambridge, UK: Cambridge University Press.

McNeill, D. (1992). *Hand and mind.* Chicago and London: University of Chicago Press.

McNeill, D. (1997, February). *Imagery in motion event descriptions: Gestures as part of thinking-for-speaking in 3 languages.* Paper presented at the *Berkeley Linguistics Society*, Berkeley, CA.

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

Pierce, C.S. (1940). Logic as semiotic: The theory of signs. In J. Buchler (Ed.), *The philosophy of Pierce: Selected writings.* London: Routledge & Kegan Paul.

Sanders, R.E. (1987). The interconnection of utterances and nonverbal displays. *Research on Language and Social Interaction, 20,* 141-170.

[14] Segment #9 also clarified that "it" was to refer to a pin(s)-going-up event. But more than just that clarification was achieved because the explainer used the clause "each pin go up" rather than the simple singular in "the pin go up".

[15] Not necessarily with conscious awareness, of course.

[16] Because of space limits, I talk about composite signals here as if they primarily combine verbal and nonverbal elements. Instead, with Clark (1996) I think of them as combining conventional symbols with iconic and indexical ones (Pierce, 1940). In the data analyzed, all verbal signals were conventional and all nonverbal ones were iconic and/or indexical, but this is not always so. In a future article I hope to present a theory of composite signals made stronger and more precise by framing it in terms of these 3 ways of symbolizing.