

*To appear in Linguistics and Education, 10 (3)*

## **MODES OF MEANING IN A SCIENCE ACTIVITY**

Gordon Wells

Ontario Institute for Studies in Education

University of Toronto

When I selected the videotape of the shoebox experiment as the basis for practical work in my graduate class, I did so for two reasons. The first was because of its completeness as the record of the various stages of an activity, from initial selection by the two children to the writing of the texts in which they report what they have done and learned; the second was that I thought that, as I had been the 'visiting teacher', there might be advantages in my being able to offer a participant's perspective. As it turned out, both of these expectations were fulfilled - though the second not exactly as I had anticipated. For as we viewed, reviewed and discussed the tape analytically, it was the nature of my own participation that became problematic for me; in particular, I was disturbed by the differential manner in which I interacted with the two children. What was it, I wondered, that had led me to give so much more of my attention to Jessica than to Alan?

At first I thought it was just that Jessica was more interested than Alan in my attempt to get them to see their efforts to replicate the phenomenon of 'bending light' - an 'experiment' that they had found in an illustrated book on light - as an opportunity to engage in the more scientific activity of formulating and testing an explanatory hypothesis. Certainly, on a purely verbal level, Jessica was much more willing than Alan to engage in my reconstruing of the activity, and she seemed to be the main beneficiary of my efforts. In fact, from the transcript made of the recording, Alan hardly seemed to participate in the discussion at all. However, as became

apparent when we paid close attention to the videotape, Alan was just as interested as Jessica in the phenomenon itself. It was he who manipulated the flashlight in such a way that the crossing beams of light could be clearly observed to 'criss-cross' on the bottom of the box. Was it, then, that his command of English was more limited and that, for this reason, he left the speaking to Jessica?<sup>1</sup>

Comparison of the two written reports that they subsequently produced certainly suggests that Jessica was more fluent than Alan and that, to a much greater extent than he, she was writing to communicate rather than simply to complete a required task. However, her assumption that she should do most of the speaking seemed to have a different basis, for it was not only when interacting with adults that Jessica monopolized the interaction. As I discovered later, when I viewed the episodes earlier in the morning in which they looked through the book to select an experiment, and again when they carried it out in the dark under the teacher's desk, it was clearly Jessica who was in charge. She made all the decisions, told Alan what to do, and monitored his actions - to the extent of taking over from him when she was not satisfied with his performance. However, when I first interacted with the two children I was not aware of this power imbalance; if I had been, I might have made a greater effort to ensure that Alan had an equal share of the floor.

On still further viewing of the tape, a different and perhaps more plausible - though less pedagogically acceptable - explanation for my concentration on Jessica occurred to me, namely that she was socially more adept than Alan. Enjoying adult attention, she was more skillful in securing and responding to an unknown adult's interest and, without being aware of it, I responded by giving most of my attention to her. There certainly seemed to be grounds for accepting this interpretation. For example, where Alan was apparently absorbed in his own goal of repeating the phenomenon of the 'criss-cross' beams of light, Jessica was willing to play up to

my agenda and engage in the attempt to convince the sceptical adult that their claim to have seen the beams criss-cross could be justified by empirically supported explanation.

It was thus in order to check out these hypotheses that I first began to look more closely at, and listen to, the extra-linguistic dimensions of the interaction. Up until that point, although having viewed the videotape many times, I had been satisfied, for the purposes of analysis, to work from the written transcript that had been prepared from the audio track of the recording. When I began to pay closer attention to all the information available in the videotape, however, I came to see how inadequate the transcript was as a record of what had been going on.

### **Attending to Non-Verbal Meaning in Discourse**

Face-to-face interaction involves many dimensions of meaning, only some of which are represented in a written transcription of "what was said". In particular, a written transcript fails in general to capture meanings that are conveyed by such non-verbal means as intonation, facial expression, gesture, and participants' spatial orientation to each other and to the material artifacts involved in their activity.<sup>2</sup> However, this is hardly surprising, since for those who grow up in highly literate cultures, meaning is often equated almost exclusively with linguistic meaning, that is to say with the semantic distinctions that are realized lexico-grammatically through words and the syntactic relationships between them.

To understand the probable reasons for this "logocentrism" it is necessary to consider the different ways in which meaning is related to observable behavior in the different semiotic modalities and how, in the historical development of writing, only some modes of meaning making came to be represented directly in the visuographic display.

A central feature of language is its tendency towards categoriality. For example, despite the continuous variation of the sound that is produced in speaking, the stream of speech is treated as consisting of a succession of tokens of discrete phonemic categories, and particular sequential patterns of phonemes are construed as particular morphemes or lexemes. So, although in fluent speech sounds may be produced that are intermediate between, for example, "p" or "b" or "r" and "w", they are nevertheless heard, in context, as instances of one or the other phoneme, and the larger segments of which they are a part are heard as discrete lexemes - for example, either "pie" or "by". There are no intermediate possibilities.<sup>3</sup> And the same is true of the meanings that words and syntactic structures encode; for the most part, they are treated as equally categorical.

This is not the case, by contrast, for intonation or gesture, where parameters such as amplitude, rate or directionality can vary continuously, with corresponding continuous gradation of the meanings expressed. For example, within the tone group (i.e. a segment of speech containing a tonic syllable on which there is a terminal pitch movement), the greater the amplitude on the tonic or any other stressed syllable, the more the meaning of the word or phrase in which that syllable occurs will be perceived as being emphasized and therefore in contrast with alternatives that potentially might have occurred in that position (including the unstressed, and therefore unmarked, occurrence of the same item). Similarly, with gesture, the orientation of the palms towards or away from the speaker, for example, can signify varying degrees of commitment, on the part of the speaker, to the point of view being expressed.

The strong tendency for linguistic forms and meanings to be treated as categorical may also explain why non-verbal dimensions of meaning are largely ignored in writing. From recent work on the development of writing it seems that, in the earliest stages of the evolution of most writing systems, the primary aim was not to represent speech directly, but to find a way of giving permanent visual representation to the meanings expressed through speech (Harris, 1986). That

is to say, the visual symbols that were used were iconic, and directly represented the categories of objects and events referred to in speech rather than indirectly through the representation of the linguistic units through which they were encoded. So, by the time that writing systems came to represent the formal organization of speech more directly through the incorporation of syntax (e.g. 'two horses' instead of 'horse, horse') and, in some cases, through the use of a syllabary or alphabet to represent the phonology of speech, the basic organizing principles had already been established. Writing represented, first and foremost, the semantic categories encoded in lexical items and in the syntactic relationships between them (Harris, 1986). Distinctions of meaning of a pragmatic kind, that are often signalled non-verbally, such as how the sentence should be taken - as assertion or conjecture, description or declaration - were left to the reader to infer from the context. Similarly, attitudinal meanings, such as degree of certainty about or agreement with the proposition - to the extent that they were encoded non-lexically and also non-categorically - were also left to the reader's imaginative reconstruction as he or she read the text aloud.

However, this almost exclusive concern with the lexico-grammatical dimension is not confined to the representation of meaning in writing; it also extends to the way in which we think about meaning in speech. For, whatever the earliest relationship between visual representation and speech, as written scripts came to be linguistically based and thus to represent the sequential organization of speech, they also began to function as models of speech. As a result, as Olson (1994) argues, folk theories of the way in which meaning is represented in speech are dependent on which features happen to be captured in the orthographic system used to represent it in writing. Furthermore, since the "non-verbal" dimensions of speech are not captured in standard written texts, they are more difficult to bring to the level of conscious awareness. The result is that modes of meaning that are not directly represented in writing are not readily available for subsequent inspection and reflection - at least not with any specificity.

Evidence for this bias for attention to be given to meaning that is realized lexico-grammatically can be found in the nature of the glosses and formulations that participants use to characterize previous contributions when a breakdown in communication occurs (Jefferson, 1972; Wardhaugh, 1985). Choice of particular lexical items or of grammatical features, such as tense, may be questioned or disputed, but rarely is mention made of intonation, gaze or gesture, even though one of these dimensions may be the cause of the difficulty. The reason, it seems, is that we simply do not have "meta" terms, in everyday language, with which to refer to the distinctions that are perceptible on these dimensions of interactive behaviour or to the differences in meaning that they communicate.

However, that is not to say that, on these non-verbal dimensions, meaning is always indeterminate. On the contrary, some distinctions are almost as conventional and categorical in their meaning as lexical items. For example, rising and falling pitch movement in English are conventionally used to signal the contrast between information that is treated as shared and that which is offered as new (Brazil, 1975) and this contrast, taken together with the syntactic arrangement of the subject and first element of the verb in the main clause, is systematically used to distinguish between the major classes of speech act function. Other intonational features, such as the range of pitch movement, variation in pace, and marked stress, are also regularly used to communicate the speaker's orientation both to the content of the message and to the intended recipient (Chafe, 1985; Lemke, 1994). and for those aspects of meaning that are signalled by these means the speaker may also be held accountable. Nevertheless, as already suggested, it is usually more difficult for a listener, if challenged, to explain the grounds for his or her interpretation.

Similarly, direction of gaze and facial expression can also be conventionally interpreted, at least to some degree. Looking away when asked a question, for example, is usually interpreted as unwillingness or inability to give an answer in the terms proposed by the questioner, and smiling

or frowning are normally interpreted as signalling, respectively, a positive or negative orientation. Some gestures also have a conventional interpretation, ranging from nods and shakes of the head to thumbs up and a whole range of scatological gestures employed by irate drivers.

Nevertheless, with the exception of the latter, it still remains the case that, when compared with meanings that are encoded lexico-grammatically, distinctions of meaning that are expressed non-verbally are much more difficult to pin down and identify. Lacking means for representing or referring to them, participants rarely appeal to them in attempting to resolve disagreements about what was 'said' and what was 'meant'. For the same reason, they are rarely recognized as making a significant contribution to the meanings that are co-constructed in face-to-face interaction.<sup>4</sup>

However, it cannot be assumed that, because these dimensions of meaning normally operate outside conscious awareness, they do not influence participants' interpretations of each other's meanings (Schiffrin, 1994). In fact it was because I suspected that they might have played an important part in the shoebox episodes that I decided to investigate them more systematically.

### **Analyzing Non-Verbal Meanings**

As already pointed out, one of the key differences between verbal and non-verbal dimensions of meaning is that the latter tend towards continuous variation. There is, thus, an inherent contradiction in proposing to code non-verbal meanings by assigning tokens to category types. First, it is often difficult to decide when two instances are sufficiently different to be treated as tokens of different types and, second, since, unlike lexico-grammatical categories, they do not wear their 'names' on their sleeves, it is difficult to find appropriate ways of describing the categories that are distinguished. Nevertheless, until at least tentative decisions are made with respect to these problems, it is impossible to proceed towards a systematic investigation. A next

part of this paper will be devoted, therefore, to a description of the ways in which we are addressing these issues.<sup>5</sup>

A further problem that has to be solved is that of accurately representing the video-recorded information that is treated as significant in such a way that it can be typed into a computer and subsequently displayed in a form that is easily interpreted by a reader who has no more than the printed text. Since, in practice, the solution to this latter problem interacts with the former, I shall start by describing the procedure that, after a considerable amount of experimentation, we are currently using. It goes without saying that we continue to try to improve it.

### Transcription

In the past, it was generally our practice to produce the first draft of the transcript from an audio copy of the sound-track taken from the recorded video-tape. There are two advantages to this procedure. First, in copying the sound onto an audio tape, the signal can be passed through a graphic equalizer and adjusted to maximize interpretability and, second, an audio machine with foot-pedal control that incorporates a variable rewind function makes the process of transcribing much easier than when using a standard VCR. Then, following the production of the first draft of the transcript, the video-tape would be reviewed in order to check the identification of speakers and to add information about the accompanying non-verbal activity. Up until this point, however, we had been sparing in the amount of information included, simply placing brief glosses or descriptions (in parentheses) at appropriate points in the verbal text. In other words, the transcripts we have worked with up until now have been largely restricted to lexico-grammatically realized meaning.

In the present case, a transcript of this kind already existed, so selected episodes from this were used as the basis for the current work. These were first reformatted so that each 'utterance'<sup>6</sup>

appeared as a separate text paragraph and then the whole episode was entered into a table, using the "text to table" command in the WordPerfect (Mac 3.0) word processing program. When the necessary tidying up had been carried out, additional columns were inserted to the right of the text in which to add codes for the non-verbal dimensions. Next, tone unit boundaries within utterances were marked (|), where appropriate, and an indication of pitch movement added immediately preceding the word containing the tonic syllable in each tone unit. Five options were recognized: / rise; \ fall; ∨ fall-rise; ∧ rise-fall; and -- level.<sup>7</sup>

Finally, on the basis of repeated viewing of each utterance, a plain language description of the significant non-verbal behavior accompanying each utterance was added in a last, wider column prior to coding gesture, gaze and spatial orientation.<sup>8</sup>

## **Coding**

Although in interaction the many dimensions of meaning making are seamlessly woven together in a continuous flow, analysis requires that they be separated and coded independently. Only then can the relationships between the different dimensions be described in a systematic manner. As a first step in this ambitious program, we decided to focus on a limited number of dimensions, selecting those that we expected to be relatively easy to code, at least at a gross level of delicacy, and those that we hypothesized might be particularly significant for the attempt to explain the unequal manner in which the teacher interacted with the two children. Our approach to the coding of intonation, or at least one aspect of it (the management of the 'given-new' contract (Brazil, 1975)) was briefly presented above. The three remaining dimensions are described below.

## Spatial Orientation

Participants' orientation to each other and to any artifacts that play a significant role in their joint activity can be expected to affect the nature of their participation. For example, when students sit in rows facing the teacher at the front of the class, the possibilities of peer interaction are substantially diminished compared with a situation in which a small group is arranged in a circle with the teacher absent. However, even in the latter case, if one participant has greater access than the others to a critical artifact such as a book or computer keyboard, that participant is likely to play a more central role in determining how the activity proceeds (Wegerif, in press).

Conventions of varying degrees of delicacy might be devised to represent information about spatial orientation, although any scheme would quickly become overly complex if the number of participants were large or if there were frequent or continual changes of orientation. Fortunately, in the episodes analyzed here there were only three participants and a single artifact cluster - the experimental apparatus - as the focus of their shared attention. The solution adopted here, therefore, was to use initials to identify the human participants and X to indicate the apparatus. A new coding was entered each time there was a change in their relative positions. Thus, a typical configuration might be represented as:

J  
T   X A

which is to be interpreted as, viewed from the position of the camera 'below', Jessica and Alan close to the apparatus on adjacent sides of the desk, and Teacher facing Alan, but further away than they from the apparatus. (Significantly, the end from which the beams were projected through the two slits in the box was the one facing T; however, this information could only be included in the prose description, as no concise way of coding it could be devised.)

## Gaze

In informal conversation between peers, participants frequently look at each other and engage in periods of mutual gaze. The speaker typically looks at one or more of the addressees at the beginning of a turn and again at the end; during short turns, gaze is often maintained throughout, while in long turns the speaker may look away for part(s) of the turn. Listeners behave in a similar manner, although they may spend a larger proportion of the time looking away. Thus, the points at which mutual gaze most frequently occurs is around points of potential turn change. This pattern is modified when an object, person or event is the current focus of joint attention. In these circumstances, all participants' gaze tends to be directed to this 'third party', although mutual gaze still frequently occurs at potential turn changes.

Mutual gaze seems to perform a number of functions. First, as its most frequent temporal location suggests, it plays an important role (together with other cues in the auditory channel) in managing the smooth transfer of the floor from one speaker to another, including the selection of next speaker. This aspect of gaze has been described in considerable detail in the work of conversation analysts (Sacks, Schegloff and Jefferson, 1974). A second function is that of monitoring the gestural and postural behavior that accompanies oral communication. Speakers monitor listeners' uptake of their messages, while listeners monitor the additional 'information' that speakers provide, particularly through gesture and facial expression (Schiffrin, 1994). Without mutual gaze, this important means of negotiating intersubjectivity would be unavailable, as is the case in telephone conversations. A third function is related to the relative status of the participants; here, avoidance of gaze may be just as important as looking one's interlocutor 'in the eyes' in signalling in a culturally appropriate manner the recognition of one's inferior status.

Determining co-participants' direction of gaze is a skill acquired in very early infancy (Bower, 1974) and, from early childhood, humans are able to make a similar determination for people with whom they are not actually interacting, provided they can see the full face. This skill even extends to the observation of naturalistic representations of people in paintings and in the images captured by still or video cameras. It is therefore relatively easy to determine the gaze of participants from a videorecording of their interaction, provided one can see their faces. The difficulty, however, is in deciding how much of the resulting information to code, and how to do so.

In the present case, we decided to focus on speaker's gaze, and simply to code whether or not s/he looked at the addressee(s) for at least some part of his or her speaking turn and, if not, whether his or her gaze was directed to the current focus of joint attention or elsewhere. We also coded whether mutual gaze between speaker and addressee(s) occurred during a speaker's turn. This information was recorded using initials, as with spatial orientation, with > to indicate directed to, and <> for mutual gaze. Obviously, a much more fine-grained analysis could be carried out, including a coding of the periods during which each participant was looking at one or other of the co-participants.

### Gesture

Surprisingly, rather little attention has been given to gesture, and almost none to its contribution to interaction in educational settings. One of the most comprehensive general accounts is provided by McNeill (1992), based largely on videorecorded observations of subjects retelling a Sylvester and Tweety Bird cartoon. In introducing this work, he makes a clear distinction between gesture (classified as 'gesticulation' by Kendon (1988), on whom McNeill draws) and other forms of manual communication. Unlike 'emblems' (and even more, the signs in a sign language such as SLA), which have a conventional meaning and a standard form, gestures are

"idiosyncratic spontaneous movements of the hands and arms accompanying speech" (McNeill 1992, p.37). He then distinguishes four different kinds of gestures: iconics, metaphoric, deictics and beats.

We found the continuum from gesticulation, through emblems, to full sign languages a helpful starting point. Like McNeill, we were mainly interested in the gesticulation end of this continuum. However, on viewing our data in the light of the criteria discussed, we decided that the continuum needed to be extended further to the 'left' to include bodily movements that, while not intended to communicate, are interpreted as meaningful by other members of the culture and oriented to, where necessary, in order to coordinate joint activity. What we are referring to here are the movements that make up a recognizable action or succession of actions, such as making a bed or sawing and nailing pieces of wood to construct a box. In our data, the second episode contained a considerable amount of such interpretable action, as the participants together invented and applied the test for determining whether the beams of light bent as they passed through the jar of water.

However, what finally persuaded us that such action should be included at one end of the continuum of semiotic behavior was a category intermediate between action and gesture as defined above. We noticed that when the two children had completed their experiment, they were anxious to demonstrate it to others. Now, however, although they repeated many of the same actions, they did so not to achieve the experimental result but rather to demonstrate to others how the action was performed.<sup>9</sup> On these grounds, we are proposing an extension of 'Kendon's continuum' as follows:

action -- demonstrative action -- gesture -- emblem -- sign language

Equally fruitful as an initial framework was McNeill's categorization of the different kinds of gestures. However, here again, we found that our data suggested the need for some further distinctions. Currently, our scheme distinguishes the following categories and sub-categories:

<u>Category</u>	<u>Code</u>
Action	A
Demonstrative Action	Da
Deictic: a) Point to concrete object or location	Po
b) Trace the path of an object or phenomenon	Pt
c) Abstract Point to a 'concept' referred to in either concurrent or preceding speech	Pe
Iconic: Gesture that iconically represents an object	Go
or action referred to in the concurrent utterance	Ga/t
Metaphoric: a) Gesture that iconically represents a 'concept' in the concurrent utterance (e.g. closing the fist to represent 'contraction', or a horizontal movement of one hand away from the body, palm down, to represent 'it was straightforward'.)	Gc
b) Gesture that represents an abstract relationship, such as 'equivalent to' or 'contrasted with'	Gr
c) Gesture representing a mental state (e.g. touching one's temple to represent 'thinking')	Gm
d) Gesture representing communicative action, such as 'telling' or 'writing'	Gc
Attitudinal: a) Gesture that represents an affective state, such as pleasure	Gf
b) Gesture that represents an epistemic attitude with respect to the concurrent utterance, such	Gz

as uncertainty

Emphatic: Referred to as 'beat' by McNeill (1992), this E  
typically up-down gesture marks emphasis or  
"indexes the word or phrase it accompanies as  
being significant, not for its own semantic content,  
but for its discourse-pragmatic content" (p.15)

In addition, gestures were coded as either underlining (u) the meaning of the accompanying speech or substituting (s) for speech.

As with the development of most coding schemes, the categories above were arrived at after a lengthy period of to-and-fro between different versions and their application to increasingly large segments of the data. Finally, all three episodes were coded using the above categories and an \* placed before the word in the utterance on which the 'stroke', or point at which the gesture comes to fruition, occurred.<sup>10</sup> For the most part, little difficulty was experienced in assigning gesture tokens to types.<sup>11</sup>

Table 1 presents a short coded extract from episode one in order to illustrate the way in which the coding was carried out.

Table 1. Sample of Coded Transcript

44:25	T	Now . can you \tell me .. what question have you \answered?	T<>J TJA<x			[A is off picture] T sits on ch A comes to stand opposite T All look at shoebox
		or what question . have you <u>found</u> <u>\out</u> about by doing this?				A continues to look down at t
	J	<u>Well</u> . <u>we</u> x-	T<J			
		Well we had- . well first our question \was .	J<X			J holds box, lifting it up and d from table, nodding head from side to side
		can light- does light always shine */straight .	J<>T	Gcu		turns to face T on 'can' and makes strong rh horizontal ge from r to l across chest on 'straight'
		/straight*forward?	J<>T	Gcu		repeats gest. into lh
	T	And /does it?	T<>J			
	J	^No .	J<>T	Gzu		J wobbles head=qualified
		cos it sort of *\bends		Pe		rh to inside of box
		cos before when we put this *RIGHT \there .	J<X	Dau		picks up jar on 'cos' and place in box on 'right'
	T	Yeh				
	J	and this part right through there it was very *\dark .		Pe		said as an addition looks to right, r.h. point
		but we had trouble getting it /dark and it was just a ^torch	J<>T			laughs after 'dark'
		and we put it *right /there .		Dau		J demonstrates where they hel the f-light, as if holding it, clc to the slits
	A	<u>xxx</u>		Dau	JA G X	A moves close to J's side and : shows where they held the lig
	J	but it couldn't be*right /close to it	J<T	Dau		(marked rising inton.) J dem. light close to box
	T	uhhuh				
	J	because it would be *too <u>bright and</u> <u>straight</u>		Gzu		J openhand gest. both hands, indicating excess
	A	<u>It was just</u> too straight <u>xxxx</u>	A<T	Gtu		A rh horizontal gest along pa of light through box x2

## Results of the Analysis

At this early stage in our investigation of the contribution of non-verbal dimensions to the meanings negotiated and co-constructed in face-to-face interaction, it would be inappropriate to attempt to draw conclusions of any generality. Instead, the results of the analyses so far carried out will be used to illuminate the particular task-related interaction from which they are derived. In particular, I shall use them to compare the interactional behavior of the two children and of the teacher as he interacted with each of them. For this purpose, I shall focus chiefly on the results of the analysis of gesture. First, however, I will comment briefly on the other dimensions examined.

The three episodes discussed are those in which I interacted with the two children in the role of 'teacher'. Summary details are shown in table 2.

Table 2. Summary of the episodes analyzed

Episode	Start	Stop	Activity
One	44:25		Jessica and Alan describe their experiment to the teacher
Two	47:51	52:12	Jessica and Alan, in collaboration with the teacher, develop a method of demonstrating that the beams cross
Three	1:14:24		In discussion with the teacher, Jessica and Alan plan the writing of their report

## **Spatial Orientation**

In all three episodes, the three participants tended to situate themselves on three sides of the desk with the shoebox in the centre. In the first two, while the experiment was being described and then extended, the shoebox was placed with the slits on the left (from the camera's perspective). Initially, Jessica was positioned on one side of the box, facing the camera, with the teacher at the end with the slits and Alan facing him at the other end. As they demonstrated what they had done, Alan moved to stand by the side of Jessica in order to share in the demonstration.

Episode two started with the same arrangement as did episode one but, early on, the teacher changed places with Jessica, thus putting her in the privileged position of controlling the flashlight and directing its beam through the two slits. On one occasion, Alan leant across the box in an effort to direct the light, but Jessica refused to cede control. Alan did, however, take the initiative in finding an exercise book to cover his end of the box so that the beams emerging from the jar could be more easily seen. More or less excluded from the action, Alan wandered away for a few seconds but, on returning, he again leant over the box to test the effect of covering one of the slits with his finger.

In episode three, as they began to plan the report, Alan and the teacher were sitting on two adjacent sides of the table, Alan facing the camera; Jessica was standing opposite the teacher. Initially, the box was oriented as previously. Part way through this episode, two boys came to show the periscopes they had made and both Alan and Jessica engaged in conversation with them. Jessica turned the box so that the slits were facing her and she demonstrated what they had been doing to one of the boys. A little later, Alan stirred the water in the jar with his finger and briefly continued the conversation with one of the boys, who looked at Alan through his

periscope. Then, on Jessica's insistence, he turned his attention back to the planning of the report.

In reviewing the results of this analysis, three things are apparent. First, whether by design or by chance, Jessica is always in a more dominant position than Alan; throughout, it is she who controls the flashlight and, in the third episode, she stands facing the teacher while Alan sits to one side of him. Second, although Alan makes several attempts to play a more central role in the action, Jessica does not allow him to do so. Third, despite this, Alan remains involved and, apart from a few brief interludes, appears to give his full attention to what is going on.

### **Gaze**

There is a strong contrast between the gaze behavior of the two children across all three episodes. Except when all three participants are looking at the apparatus, Jessica looks at the teacher when she speaks, often with a smile or some other facial expression showing interest. On almost all occasions, he looks at her while she speaks and there is a period of mutual gaze. By contrast, Alan rarely looks up when he speaks and, when he does, it is more often at Jessica than at the teacher. He rarely makes eye contact with the teacher. The difference between them in this respect can be captured in the finding that whereas Jessica looks at her addressee on more than 90% of her speaking turns, the proportion for Alan is only about 25%. When to this is added that Jessica speaks twice as often as Alan, and typically at much greater length, it is apparent that there are many occasions of mutual gaze between the teacher and Jessica and hardly any with Alan.

This finding goes a considerable way towards explaining the differential distribution of the teacher's attention: Jessica, to a much greater degree than Alan, kept up her end in the interaction; through frequent exchanges of mutual gaze, she signalled that she was on the same wavelength as the teacher and so elicited further interaction. By contrast, Alan failed to give such signals and so

was interpreted as not wishing to engage in the interaction. Whether this was an appropriate interpretation will be further considered below. (See also Buzzelli & Johnston, this issue.)

## **Gesture**

As previously described, all gestures were assigned to one of the five categories and, where appropriate, to one of the subcategories within them. In addition, where a nonverbal action that accompanied or took the place of a speaking turn could be considered meaningful within the framework provided by the activity, it was coded as either 'action' or 'demonstrative action' according to the criteria described above. Then, as frequencies in some subcategories were very low, a regrouping was carried out to create the following seven clusters: *Action*; *Pointing* to material objects and locations; *Demonstrative Action*; *Iconic* gestures; *Metaphorical* gestures (including metaphorical points); *Attitudinal* gestures; and *Emphatic* gestures. Table 3 shows the distribution of types of gesture by participant.

As can be seen, overall, Jessica produced the greatest proportion of gestures, and Alan the least. However, what is more interesting is the different rankings of gesture types within participants. Although for both Jessica and Alan, action was the most frequent form of non-verbal communicative behavior (most tokens of which occurred, as we shall see, in episode 2) the two children differed in their next most highly ranking categories. For Alan, it was iconic gestures, and, in particular, gestures tracing the path of the beams of light; for Jessica, on the other hand, it was attitudinal gestures (mainly her affective attitude to the things she was talking about) that ranked second, closely followed by pointing and demonstrative action. Meanwhile the teacher made most use of pointing gestures and gestures that emphasized the ideas he wanted the children to grasp. It would not be unreasonable, I believe, to take these relative distributions of gesture types as indicative of the overall stance of the three participants to the activity in which they were jointly engaged.

Table 3. Distribution of Gestures by Category and Participant

Category	Jessica	Alan	Teacher	Total
Action	16	14	2	32 (19.8%)
Point (to object or aspect of situation)	11	6	23	40 (24.7%)
Demonstration	10	2	0	12 (7.4%)
Gesture: Iconic	5	9	6	20 (12.3%)
Gesture: Metaphoric (incl. abstract point)	6	3	5	14 (8.6%)
Gesture: Attitudinal	12	1	2	15 (9.3%)
Emphasis	8	2	19	29 (17.9%)
Total	68 (42.0%)	37 (22.8%)	57 (35.2%)	162 (100%)

Equally illuminating is to look at the distribution of gesture types across the three episodes.<sup>12</sup>

From table 4, it can be seen that the explanation-oriented activity in episode two evoked the largest overall number of gestures and that, of these, more than 75% took the form of goal-

oriented action and deictic points intended to make explicit the specific material phenomena between which a causal relationship was being proposed, e.g. "T: and the light that goes in \*Vhere (pointing to left slit) . is showing up on \*/^that side (pointing to bar of light on right)". Not surprisingly, most of the action was performed by the two children, whereas most of the pointing gestures were made by the teacher, in his attempt to focus the children's attention on the phenomena that he considered particularly relevant for the construction of an explanation.

Table 4. Distribution of Gestures by Category and Episode

Category	Episode 1	Episode 2	Episode 3
Action	1 (3%)	22 (34%)	9 (14%)
Point (to object or aspect of situation)	6 (18%)	28 (42%)	6 (10%)
Demonstration	8 (24%)	1 (1%)	3 (5%)
Gesture: Iconic	8 (24%)	6 (9%)	6 (10%)
Gesture: Metaphoric (incl. abstract point)	4 (12%)	1 (1%)	9 (14%)
Gesture: Attitudinal	6 (18%)	4 (6%)	5 (8%)
Emphasis	0	4 (6%)	25 (40%)
N =	33 (100%)	66 (100%)	63 (100%)

In the much shorter first episode, by contrast, there is a more even distribution. Here, a variety of gesture types support the children's attempts to describe and demonstrate to the teacher the actual procedures they had used in producing the 'criss-cross' effect that the book had instructed them to look for. It is in this episode too that, through her attitudinal gestures, Jessica expresses her strong involvement in solving the problems posed by the 'experiment' that they have chosen to carry out.

Episode three shows yet another distributional pattern, with a high preponderance of emphatic gestures - largely contributed by the teacher - as the three participants consider the content and sequence of the different elements that should be included in the report that the children are required to write about the activity in which they have engaged. It is significant that it is in this episode, too, that the majority of metaphoric gestures occur, as Jessica considers the (absent) audience and the sort of information that they would want to know. The action and demonstrative action gestures, by contrast, occur chiefly when Alan and Jessica are showing what they have been doing to two boys who pay a visit.

Taken together, the results of the comparison across episodes suggest that participants produce different types of gesture according to the nature of the task in which they perceive themselves to be involved. That is to say, as with speech, participants make choices in terms of the genre appropriate to the task in hand (Miller, 1984; Kamberelis, 1995). Episodes one and two are clearly distinguishable from each other in this respect, with the first being concerned with demonstration and the second with making an explanation. In these two episodes, all three participants seem to share a similar orientation to the task and use a similar range of gestures (including action and demonstrative action as categories of gesture for this purpose). In episode

three, by contrast, the participants appear to have adopted somewhat different orientations; while the two children were exploring the alternatives of content and organization for their written reports, the teacher was providing guidance and support from the perspective of one who had already considered the options.

In work with fifth and sixth grade science students, Crowder (1996) found evidence of a somewhat similar matching of gesture type to task undertaken. However, for the pairs of students she compared, the task was relatively constant - to explain a phenomenon - and what differed was the orientation to the task adopted by the different students. Where students were describing a model or explanation with which they were already familiar, their gestures tended to be iconic or points which added no information that was not already present in the accompanying utterances ("gesture-redundant talk"). In other words, their gestures tended to add emphasis for the benefit of the audience. The gestures of students who constructed explanations in the moment, on the other hand, often added meaning to what was verbally expressed or even substituted for a word that was unavailable ("interwoven gesture talk"). These gestures often foreshadowed the meaning that was verbally expressed, and the stroke was much less likely to coincide with the intonationally marked tonic syllable in the accompanying utterance than was the case where students described an already familiar model.

In the data from the shoebox experiment, the children provide some evidence of both types of orientation. In the first episode, both are describing the experiment they have just carried out and their gestures serve to reinforce the meanings expressed in speech, while at the same time giving directional or locational precision to words or phrases such as "very far", "too straight" or "like that". In the second episode where, in response to the teacher's request, they are attempting to prove that the beams of light bend, it is their actions that are the primary mode of meaning making and their speech provides an interpretive commentary on what they are doing and observing. In episode three, where the task involves both envisaging the likely informational

needs of their audience and reviewing the procedures that might need to be described, both children produce metaphorical gestures that seem to help them to clarify what they will need to report. However, the ratio of gestures to utterances is lower than in the previous episodes and they rarely substitute for meanings unspoken.

The analysis reported by Crowder (1996) is considerably more detailed than the one attempted here and the students to be compared were chosen specifically because they were judged to demonstrate contrasting orientations to the same task. In the present case, with a lesser degree of analytic delicacy, it is genre differences between tasks that are most salient, rather than individual differences in task orientation. But for that reason, the two studies can be seen to complement each other in demonstrating genre-related differences in choice of gesture, both within and between tasks.

Table 5. Percentage Distribution of Utterances by Participant Within Episode  
(Distribution of Gestures in Parentheses for Comparison)

	Jessica	Alan	Teacher
Episode 1	56.6 (63.6)	18.9 (21.2)	24.5 (15.2)
Episode 2	41.7 (40.9)	12.6 (21.2)	45.6 (36.3)
Episode 3	36.4 (31.7)	28.8 (25.4)	34.7 (42.9)
Total	42.3 (42.0)	20.8 (22.8)	36.9 (35.2)

However, Crowder's distinction between "gesture-redundant" and "gesture interwoven" talk prompted me to look more closely at possible differences between Jessica and Alan in the way their gestures and talk were related. Most apparent is that Jessica both speaks and gestures more than Alan (see table 5). She also produces many more attitudinal gestures than he does and, leaving aside metaphorical points, more metaphorical gestures as well. However, what table 4 reveals is that Alan's ratio of gesture to utterance is slightly higher than Jessica's, and that this is most apparent in episode two, where making an explanation is the task in hand. On more than one occasion in this episode, Alan silently performs task-related actions and makes gestures that trace the path of the beams of light while Jessica does the talking. The same occurs in episode three. It seems, therefore, that Alan is constructing an explanation in the moment through action and gesture, even though, from the record of his speech alone, one might judge that he was less fully engaged than Jessica.

## **Discussion**

I started this paper with a number of possible explanations for my unequal treatment of the two children in the interaction I had with them around the shoebox experiment. And because the original transcript of the verbal dimension of our interaction did not provide an adequate basis for testing all my hypotheses, I decided to investigate certain non-verbal dimensions as well. By examining participants' spatial orientation, gaze and use of gestures, in addition to their verbal contributions, I gained a much more complete understanding of the factors that influenced the way in which the interaction unfolded.

In general, my hypothesized explanations were confirmed. Jessica was indeed more socially adept at interacting socially with an unknown adult than Alan - at least on the terms that the adult took for granted. Alan's habit of looking down when speaking instead of at his addressee(s)

led me to undervalue his contributions and to give more of my attention to Jessica, who behaved in a way that I found more responsive and engaging. However, having now noticed Alan's preference for avoiding mutual gaze -- a preference of which I was not consciously aware at the time - I am left wondering whether this was a cultural difference to which I was insensitive or simply a form of shyness. In either case, his gaze behavior should not have triggered such an inappropriate response on my part and I shall endeavor to be more sensitive to patterns of behavior that differ from my culturally-based expectations in my future encounters with students, whatever their cultural background.

What I was not aware of until I viewed the whole videotape, including the episodes in which Jessica and Alan chose, set up, and carried out their experiment, was the extent to which they had established a relationship of unequal status, in which Alan allowed Jessica to make most of the decisions and to direct their joint activity. When I joined them, therefore, I was unwittingly drawn into their role relationship and, for the reason given above, I was also too ready to collude in it (cf. Buzzelli & Johnston, this issue). Close examination of the two children's spatial orientation to the apparatus and to me provides ample evidence of Jessica's assumption of a controlling role, both in the manipulation of the flashlight in relation to the slits cut in the box - to which she had greater access than Alan - and in her tendency to speak for both of them, frequently speaking over Alan so that his contributions went unnoticed. Again, had I been more sensitive to Alan's relative disenfranchisement, I would have made a greater effort to ensure his equal participation.

But just as important as the discovery of my own inappropriate way of privileging Jessica at Alan's expense was the enhanced understanding I gained, from attending to the fine detail of extra-linguistic behavior, of the complex, multidimensional, and mutually constructed nature of face-to-face interaction in any situation. From this more dispassionate perspective, I came to see that Alan's apparently less intellectually engaged involvement in the 'scientific' activity of explaining

the phenomenon of bending light was not simply the result of his individual attributes or of my ineptitude, but was jointly constructed by all three participants through the totality of their communicative actions and their responses to each other's actions in ways of which they were probably to a large extent unaware.

However, the greatest surprise came from what I discovered by closely attending to the participants' gestures. My own preference for deictic points and emphatics made me realize that I had more of an 'instructional' agenda than I had thought. Jessica's fluent metaphorical and attitudinal gestures confirmed my impression that she really enjoyed being challenged to move beyond simply carrying out the instructions found in the book to developing and testing an experimental proof and then sharing her discovery with others. (This impression was amply confirmed by her reader-oriented and copiously illustrated report; see Haneda, this issue.)

But the real surprise was to discover how fully Alan was involved in the whole process. As our interaction proceeded in real time, I did not notice how he was shadowing the talk between Jessica and myself with gestures that showed he fully grasped the significance of the 'finger test' and that later, in episode three, he used a tracing gesture to explain what he could not put into words. Clearly, he was not as uninvolved as his lack of speech, his willingness to be distracted by the interventions of his peers, and his apparently aimless stirring of the water in episode three had led me to suppose.

## **Conclusion**

Although I had no intention of investigating my own 'teaching' when I selected this videotaped interaction as a 'text' for practical work in my class on discourse analysis, the prompts that my colleagues' interpretations of it gave me to engage in some post-hoc action research has been doubly beneficial. First, I have learned just how 'logocentric' I have tended to be, both in my

interactions with others and in my studies of classroom discourse. From now on, I shall try to be more attentive to all dimensions of meaning making in face-to-face interaction.

Second, I have become fascinated by gesture itself and by the role that it plays in interaction. Since I was apparently not aware of the significance of Alan's gestures - nor indeed of my own - until I studied them on the recorded videotape, it appears that gestures cannot function on the same level as the verbal meanings that are co-constructed in interaction, at least not in the normal course of events. On the other hand, where informal speech is not accompanied by spontaneous gestures, it is typically experienced by listeners as lacking in conviction; similarly, for many speakers, to be debarred from gesturing is to be deprived of an important means of making meaning clear. However, if neither speakers nor listeners are regularly conscious of the gestures that accompany the utterances that they produce or hear, what exactly is the nature of their contribution?

## References

- Bower, T. (1974) Development in infancy. San Francisco: W.H. Freeman.
- Brazil, D.C. (1975) Discourse intonation. Discourse Analysis Monographs 1, University of Birmingham, English Language Research.
- Chafe, W.L. (1985) Linguistic differences produced by differences between speaking and writing. In D.R. Olson, N. Torrance and A. Hildyard (Eds.) Literacy, Language, and Learning. Cambridge: Cambridge University Press.
- Crowder, E.M. (1996) Gestures at work in sense-making science talk. Journal of the Learning Sciences, 5 (3): 173-208.
- Deacon, T.W. (1997) The symbolic species: The co-evolution of language and the brain. New York: Norton.
- Halliday, M.A.K. (1970) A course in spoken English: Intonation. London: Oxford University Press.
- Harris, R (1986) The origin of writing. London: Duckworth.
- Jefferson, G. (1972) Side sequences. In D. Sudnow (Ed.) Studies in social interaction. New York: The Free Press.
- Kamberelis, G. (1995) Genre as institutionally informed social practice. Journal of Contemporary Legal Issues, 6: 115-171.
- Kendon, A. (1988) How gestures can become like words. In F. Poyatos (Ed.) Cross-cultural perspectives in nonverbal communication. Toronto: Hogrefe.
- Lemke, J.L. (1994) Intertextuality and educational research. Linguistics and Education, 4 (3-4): 257-268.
- McNeill, D. (1992) Hand and mind: What gestures reveal about thought. Chicago: University of Chicago Press.
- Miller, C. R. (1984) Genre as social action. Quarterly Journal of Speech, 70: 151-167.
- Olson, D.R. (1994) The world on paper. Cambridge: Cambridge University Press.
- Sacks, H., Schegloff, E.A. and Jefferson, G. (1974) A simplest systematics for the organization for turn-taking for conversation. Language, 50 (4): 696-735.
- Singleton, J.L., Goldin-Meadow, S. and McNeill, D. (1995) The cataclysmic break between gesticulation and sign: Evidence against a unified continuum of gestural communication. In K. Emmerey and J.S. Reilly (Eds.) Language, gesture and space. Hillsdale, NJ: Erlbaum.
- Schiffrin, D. (1994) Approaches to discourse. Cambridge, MA: Blackwell.
- Wardhaugh, R. (1985) How conversation works. Oxford: Blackwell.

Wegerif, R. (in press) Factors affecting children's talk at the computer. In R. Wegerif and P. Scrimshaw (Eds.) Computers and talk in the primary classroom. Milton Keynes, UK: Open University Press.

Wells. G. (1981) Learning through interaction: the study of language development. Cambridge: Cambridge University Press.

## Notes:

---

1. For both Jessica and Alan, English was a second language. Both are Chinese-Canadian students, born in Canada, whose first language is Cantonese. Both have been speaking English in school since kindergarten, but Jessica uses English more often at home; she has a native-like fluency, whereas Alan seems to have a more limited command, particularly with respect to vocabulary. However, this may be as much a lack of confidence as a lack of command of English.
2. By "non-verbal", I mean those distinctions of meaning that are not encoded through choices from the lexical or morphosyntactic systems of a language. Some aspects of intonation in English, I concede, are grammatical (cf. Halliday, 1970; Brazil, 1975), but there are also dimensions that are not, and it was these in which I was particularly interested. I should also make it clear that my observations are based on my own dialect of (standard British) English, although I believe they apply to a large extent to other dialects of English and to other European languages.
3. See Deacon (1997) for a discussion of the possible evolutionary antecedents of this distinction.
4. This is not to say that non-verbal dimensions of interaction have gone unstudied; there is, in fact, a considerable body of research on each of the channels I have mentioned. My point is, rather, that they do not play a significant part in the lay-person's 'theory of speech'; nor has their significance been widely recognized in work on discourse in the classroom.
5. I should like to acknowledge the significant part played by Mari Haneda in the development of the scheme presented in this paper.
6. The term 'utterance' , as used here, refers to the smallest unit of spoken text that, in context, makes a potentially complete contribution to the ongoing discourse; it corresponds, we believe, to what Chafe (1985) calls an 'idea unit'. The segmentation of spoken text into analytic units is notoriously problematic, and usually makes appeal to several criteria. In earlier work on language development (Wells, 1981) we defined an utterance syntactically as "one independent clause and

---

any elements dependent upon it"; however, intonation and pausing were also taken into consideration. In our current work, we would give greater weight to intonation, defining an utterance as one potentially free-standing tone unit and any tone units dependent on it. In the following examples, adapted from transcripts of mother-child interaction (Wells, 1981, pp.279-280) the text on each line is treated as a separate utterance (for conventions of intonation coding, see p.#).

M: Put the /top | back on the \washing basket

C: On \there | /Mummy?

M: \Yes

No not the towel in /there

Put the /lid | .. on top of the \basket

C: Where \man gone?

M: I expect he's gone \inside | because it's \snowing

7. As we are well aware, a system with only five options falls far short of representing the full range of distinctions in pitch movement found in most dialects of English. However, it was considered adequate for the present purpose. More delicate descriptive schemes can be found in Halliday (1970) and Brazil (1975). (For an alternative and more comprehensive notation, see Wells, 1981, p.278.)

8. Many aspects of non-verbal behavior are potentially significant in accounting for the way in which an interaction develops. In the present case, only those aspects relating to the dimensions selected for coding were systematically included in the accompanying description.

9. It has been pointed out that this distinction between action and demonstrative action is somewhat similar to that between use and citation in linguistic description (Johnston, personal communication, Nov. 1997).

10. McNeill's scheme of analysis is considerably more detailed than the one employed here. Not only does it describe the shape of the gesture and its location in relation to the gesturing space,

---

but in mapping it onto the accompanying utterance, three phases are marked: preparation, stroke, and retraction.

11. Because the results are based on the same data-set that was used to develop the coding scheme, no attempt was made to carry out a formal evaluation of inter-coder reliability.

12. In interpreting these frequencies, it must be borne in mind that the three episodes were not of equal duration. Episode one was substantially shorter than the other two, and episode three was the longest (in part due to the interruption caused by the two boys with periscopes).