The Body in Scientific Discourse

Wolff-Michael Roth

University of Victoria

Paper presented at the annual meeting of the Society for the Social Studies of Science, San Diego, October 1999. All correspondence concerning this paper should be addressed to Wolff-Michael Roth, Lansdowne Professor, Applied Cognitive Science, MacLaurin Building A548, University of Victoria, Victoria, BC, Canada V8W 3N4. E-mail: mroth@uvic.ca. Tel: 1-250-721-7885. FAX: 1-250-721-7767.

Abstract

Past research provided evidence that gestures and talk are deeply integrated with cognition. Evidence from developmental psychology suggests a shift in the function and frequency of gesture use from early ages (~4 yrs) to early adolescence. To date, however, little evidence exists about the function of gestures as students from 10-18 years learn new, formal discourses. Based on a large database containing videotapes of students during inquiry science lessons, the relationship between gestures and talk is described and theorized in three types of settings: students make scientific arguments over and about (a) drawn (still) graphical models, (b) computer-based “runable” graphical models, and (c) three-dimensional models of architectural structures. Based on the analyses, evidence for three major claims is provided. First, in the absence of scientifically appropriate discourse, students’ gestures already pick out, describe, and explain scientific phenomena. Second, during the initial appearance of scientific discourse, deictic and iconic gestures precede the associated utterances. Third, as students’ familiarity with a domain increases, scientific talk takes on greater importance and gestures begin to coincide with the talk.
1. Introduction

Pragmatic approaches generally agree about deixis as a grammatical formalization of situation-bound gesticulative action to physical objects (Matras, 1998) and about deixis as the most obvious way in which language and context are related (Levinson, 1983). In its primary function, deixis is related to pointing and imaging gestures in actual space. In practical situations, people use both words (deixis, indexical expressions) and gestures (which can involve any part of the body) to construct their discourse contributions. Here, the term ‘gesture’ is used to mean any distinct bodily action which is directly involved in the process of communication (Kendon, 1985). However, the expressive possibilities of spoken language and gesture are different. In spoken language, distinct (typological) elements are concatenated into a temporal sequence which itself is structured by rules of syntax; these rules constrain both the form of discourse elements and their place in the sequence. Spoken utterances are therefore only indirectly linked to any aspect of the structure of what is being referred to. In contrast, many gestures have a topological character so that they express (temporal) action sequences, spatial relationships, and metaphors.

Words and gestures are deep features of cognition. Research in linguistic anthropology, for example, has provided evidence for the interdependence of apparently different cognitive capacities: deixis, gesture, and spatial orientation (Haviland, 1993; Levinson, 1997; Widlok, 1997). Gestures are deeply integrated into systems of directional reference which is of fundamental importance to understanding the cognitive background before which actions take their significance. For example, indexical reference plays an important role in accumulating spatial knowledge among the Hai||om bushpeople and is embodied in the form of topographical gossip involving directionals (Widlok, 1997). This interdependence clearly demonstrates that gestures and language are not simply aspects of language but of a broader communicative system (Levinson, 1997). Talk and gesture are almost always copresent during communicative acts (e.g., McNeill, 1992), especially during science activities (Crowder, 1996; Roth, in press). Although early research suggested that talk and gesture are separate components of communication (e.g., Kendon, 1983), recent research suggests that it is more advantageous to study these features as aspects of the same event (Goodwin, 1996). We are therefore held to study the visible structure in the stream of speech as it interacts with grammatical knowledge to accomplish social action.

Gestures have been classified into different types including those of deictic (pointing), iconic, metaphorical, and symbolic nature (Freedman, 1977; Kendon, 1985; McNeill, 1992). Deictic gestures “point” out some aspect of the context and therefore make it salient against everything else which becomes background. I do not claim that pointing unequivocally picks out referents. Rather, as shown below, pointing, indexical ground, and utterance together provide the resources for the listener’s interpretation. Whether the
entities they stand for are characterized by mapping relationships such that both be understood in terms of the same topological (i.e., conceptual) features. 2 For example, a scientist might outline a graph by following it using her finger, and thereby foregrounds the shape of its line. In a similar way, metaphorical gestures provide a visual expression of a metaphor. Finally, symbolic gestures function as independent signs and obtain their meaning through shared social conventions (raising of middle finger to signal an obscenity).

In this paper, featuring analyses of students’ conversations that occur over and about visual representations of scientific phenomena, I focus on deictic and iconic gestures. Deictic gestures are generally intended to make salient some object, the topic of the speaker’s communication. Iconic gestures, more or less transparently depict objects and events (or aspects thereof) in ways that are not spoken and often are difficult to render in words. Gestures are particularly well suited, in a four-dimensional analogue channel, to portray shape, form, space, and position that are not often verbally (and thus digitally) encoded. Gestures also can depict actions and motion that unfold in space and time in ways that bypass the essential linearity and sequentiality of speech (Haviland, 1993). That is, the body remains a locus for meaning even with its most trivial actions; the human body therefore maintains an essential rationality and provides others with the interpretive resources they need for building common ground and mutual intelligibility (Goodwin, 1986).

Developmental studies of communication show that there is a shift in the relative importance and in the relative placement of gesture and talk (Freedman, 1977). Around the age of 4 years, representational movements occur prior to verbal responses; at the 10-year old level, gestures start with the onset of verbal responses; finally, for speakers 14 years and older, gestures are generally subordinated to speech and do neither precede nor occur throughout the verbal response. In the adult, gestures and speech units of semantic expression are precisely coordinated (Kendon, 1983). Timing studies showed that there is virtually no delay between the gesture and talk, with 64% of the verb gesture coordination occurring within ± 200 milliseconds of each other, and 92% of the agreements occurring within ±400 milliseconds (McNeill & Levy, 1982).

While these developmental changes in the relation between gesture and talk concern those that occur ‘naturally’ in children’s appropriation of their mother tongue, one may question, Do the same or similar shifts occur when individuals learn a second language or new forms of discourse? The purpose of the present article is to provide empirical case

---

2 Because iconic similarity is learned (Eco, 1976; Quine, 1995), communicating individuals and analysts of communication have to ascertain for the situation that an iconic relation exists between the gesture and the phenomenon thereby being picked out.

---

speaker who points in the general direction of a rose and utters, “Look at this one” actually wants to foreground “rose,” “flower,” “blossom,” “petal,” “stamen,” or any other possible aspect very much depends on the specific context.
evidence for the use of talk and gesture while science students learn to make scientific arguments over and about drawn (still) graphical models, computer-based ‘runable’ graphical models, and three-dimensional models of architectural design. The main argument of this article is that in situations where students learn new (scientific) discourses and where the context supports the use of gestural resources, we observe changes in the relative importance of gesture and talk similar to those in the developmental studies. Initially, there is a considerable delay (sometimes weeks) between a gesture and the corresponding scientifically-correct words. The gap closes with experience until they coincide, that is, gesture and corresponding talk occur simultaneously. Before describing the study and its results, I turn to the nature of science laboratory talk and the analytic framework for the present research.

1.1. Laboratory talk in professional science and schools

Whereas deictic phenomena link language and the contexts of its use, science and scientific language are generally praised for their context-independence. That is, there seems to be no place for deixis in science. Yet a considerable number of ethnographic studies in scientific laboratories (also in engineering design studies) document the dependence of laboratory talk on the context; that is, in the absence of certain physical representations (drawings, photographs, graphs, models) to which scientists and engineers point to and relative to which they gesture, the interactions between them come to a grinding halt (Amann & Knorr-Cetina, 1990; Henderson, 1991). Talk continues as soon as the physical representations are available, associated with a lot of gesturing and pointing. If no representation can be fetched or drawn as facsimile, scientists and engineers often resort to gestures that render an ephemeral facsimile in the air especially when the collaborators are familiar with each other, and the topic of their collective work. This research therefore suggests that scientific laboratory talk is highly indexical, occurs in the presence of the object of talk, and involves a lot of gesturing (Suchman & Trigg, 1993): scientific thinking, literally, is handwork.

We can say that scientific talk is both over and about the representations. Talk is about the representations because these are the current conversational topic. Talk is also over the representation because the latter serves as background to which the talk is anchored and against which gestures become significant. The actual presence of objects and representations therefore provide important resources to scientists’ interactions in the laboratory. It has been noted that physicists may come to their understandings and interpretations of the subject matter in part through sensori-motor and symbolic reenactments of events (Ochs, Gonzales, & Jacoby, 1996). Furthermore, the availability of gesture as expressive medium provides resources to collaborative thinking-through processes in the sense that it affords the production of public and therefore witnessable and accountable events.
Recent research related to discourse and cognition in science education intimates that gestures are a common phenomenon and may play a role as expressive medium (Crowder, 1996; Crowder & Newman, 1993; Roth, in press). In Crowder’s work, gestures have been described as playing ancillary roles in the development of discourse. When gestures are used to describe scientific models, they are typically timed with speech and provide ‘redundant’ information. Gestures used when constructing explanations in-the-moment more frequently precede related verbal content and remedy the shortcomings of gesture-carried talk. In this paper, I argue that gestures are more deeply integrated to cognition and play an important role during the genesis of scientific discourse in school-aged individuals. Furthermore, the fact that gestures can be used to identify children’s ideas even before they are able to express the same content in the verbal modality (e.g., Cassell, McNeill, & McCulloch, 1999; Goldin-Meadow, Alibali, & Church, 1993) and that untrained adults correctly infer cognitive content (Goldin-Meadow, Wein, & Chang, 1992) is further evidence that there is more to gestures than paralleling talk.

Over the past decade, I have conducted intensive research concerning students’ science laboratory talk over and about scientific representations and artifacts. The research documents students’ appropriation of scientific language from Grade 4 to university and in four countries (Australia, Canada, Germany, USA), and contains several hundred hours of videotaped conversations in small groups and whole-class assemblies (e.g., Roth, 1995b, 1998a). The data support the following three broad generalizations. First, laboratory talk is highly elliptic and indexical, and very different from written language. Second, laboratory talk is accompanied and therefore rendered comprehensible by a large amount of gestures that point to, or stand in iconic relation to, the representations and phenomena at hand. Third, laboratory talk evolves, sometimes slowly, at other times more rapidly from stammered ‘muddled talk’ (Rorty, 1989) to scientifically-appropriate language. My initial intuition that gestures play an important role in the appropriation of scientific language (at least under the conditions described in this article) arose from the analysis of this extensive body of data (often conducted in collaboration with colleagues and graduate students).

1.2. Scientific representations as figure and ground

Much of the work on gesturing and verbal deixis has been conducted during some form of telling stories or providing information. Those settings are different from the present research on gesturing which occurs in the presence of graphical representations and three-dimensional objects. It can be expected that the presence of the object talked about changes to role and function of gesture and object. Relevant to the present issues, Hanks (1992) introduced the notion of ‘indexical ground’ with its figure/ground distinction. Deictic reference in general, but gesture in particular organizes the field of interaction into a salient foreground and an (in gesture and speech) unrepresented, more diffuse background. The discreteness, individuation, definiteness and singularity that come with deictic reference are typical figure characteristics. The referential figure is defined against a
ground characterized by diffuseness and variability. For the purpose of the present research, I only focus on those gestures of the participants that can be considered in the main-line or story-line track (Kendon, 1992; Levinson, 1983) of the interaction.

To exemplify the analytic framework used here, consider the following excerpt (Figure 1) from the conversation of three high school students trying to make sense of a computer-animated Newtonian microworld that layers phenomenal (moving objects) and scientific representations (arrows standing for, unbeknownst at this time to students, velocity and force). Figure 1.a shows that the three students G, E, and R sit in an arc-formed arrangement that provides all three of them visual and gestural access (but not equal access to keyboard and mouse); because of their specific task, this is a focused encounter in what has been called an F-formation system (Kendon, 1990: 209ff). Figure
1.b illustrates the beginning of the analysis which transcribes the utterances and associates it with particular video images (A), and time markers (C), reference numbers, and time difference between subsequent markers (E).

In this episode, G begins his attempt to describe and explain the events “So like” [1] but it takes 0.80 seconds before he points to the outline arrow [2], and another 1.03 seconds, before his utterance makes reference to the arrow [3] further made salient through a beat gesture [3-4] (Figure 1.b). His hand then slowly moves across the field of vision (0.47 seconds) until his utterance provides a description of the arrow’s effect on the circular object.

Figure 1.c illustrates the different levels of actor, action, and context. The computer monitor is the medium for a complex graphical display originating from the software product. The speakers are oriented to this display, talking about relevant features. Talk and gesture are designed by the speaker to individuate some objects or events, pushing other objects and events into the unattended background. (For example, students never talked about the visual aspect of the windows, most icons on the screen, etc.) The listener has to interpret gesture, graphical aspect, and talk together to identify just what the named objects and events are. What is actually significant has to be worked out by the listener who has to find a convergent interpretation of utterance, gesture, and the phenomena that constitute the indexical ground; the gesture individuates an object from the ground to the same extent that the ground lends significance to the gesture. This interpretation, from a phenomenological perspective, is a function of the individual’s current orientation, discourse practices, past experiences, and so forth. In this episode, G used deictic gesture [2], a beat gesture [3–>4], and an iconic gesture [4–>5]. The deictic gesture individuates the outline arrow and the iconic gesture hypothesizes the effect (“forces”) of this arrow on the circular object. Despite the delays between the gestures and their corresponding verbal expressions, the context allows a coordination of the triplet [arrow, “this arrow,” INDEX]. In the present situation, the figure itself splits into salient and nonsalient aspect (Figure 1.2). G’s utterance is about the arrow only, not about the second arrow and the circular object. Both of these also slipped into background and were not attended to.

The analysis begins by noting the computer monitor as the medium upon which something is inscribed. As Figure 1.c shows, this medium is the primary ground and that which is inscribed “floats” on top of it. Talk occurs in front of both figure and ground. The gestures change the configuration by introducing additional layers. As G points to and taps on the outline arrow ([2], [3–>4]), this object is picked out and made salient and becomes the primary aspect about which his utterance makes (or will make) a statement. The other part of the figure recedes to the background. Similarly, the gesture iconically enacts the motion of the object across the screen; although “trajectory” is an abstract notion, evidence for it remains visibly accessible in the background in the form of “multi-
exposure photographic still.” In this case we have a visual similarity of a gesture with the event described, so that we are dealing with an iconic gesture.

As a matter of foreshadowing the analyses to come, the present episode shows that deictic and iconic gestures sometimes considerably precede the utterance which has the same referent. Thus, at a micro level, gestures precede the representation of events in verbal form. Furthermore, gestures often anticipate verbal forms at a macro level when initial gestures correctly describe situations where people lack the appropriate verbal (scientific) descriptions. For example, although G uses the verb “forces” he does not yet describe the outline arrow as force. That is, from a schooling perspective, he does not know Newtonian physics because he does not use the required words force and motion in appropriate sentences. But he provides an appropriate description of it before describing and theorizing the arrow as a force, or representation of force. This description is developmental precursor of his subsequent verbally-coded knowledge of Newtonian physics. However, it was only two weeks later after several activities that he evolved a scientifically correct way of describing and explaining the microworld events that occurred on the monitor surface.

2. Data

The excerpts presented here derive from a large database of videotapes accumulated over a 9-year period in physics and biology classrooms from Grade 4 through college level and in four countries. In all of these studies, I documented the interdependence of language, context, and cognition (e.g., Roth, 1996a; Roth, McRobbie, Lucas, & Boutonné, 1997). However, I only recently realized that gestures may be an important aspect of the genesis of scientific cognition (Roth, 1996b, 1998b, in press; Roth & Duit, 1997; Roth & Welzel, 1999). I therefore reanalyzed data from different classrooms with students at different age levels studying a variety of scientific phenomena. The present analyses show examples of patterns which I have found consistently within and across these diverse settings. Because of space limitation, however, only a small number of samples are presented here. The video-based episodes presented in this article were all collected in physics lessons in Canada but at three different Grade levels (grade 4-5, 6-7, and 11) where the curricular topics were architectural structures, simple machines, and motion, respectively. In all studies, two or three cameras were used to follow as many individual groups during student-centered activities. During whole-class discussions and lectures, two cameras were used to triangulate the speakers to record all utterances with maximum reliability. Details of each episode are described as they are presented.

In all studies, the videotapes were transcribed within hours to a few days on a word by word basis, but without pause length or overlaps. The transcriptions of episodes with apparent theoretical appeal were then enhanced to include those features common to conversational analysis. That is, the enhanced transcriptions included the extent of
pauses, overlaps, stresses, and so forth. In addition, representations of the focal situations (e.g., artifacts, drawings, etc.) over and about which conversations took place were included in the transcripts. Here, these representations are video stills from the actual presentation. Because the videotapes were recorded at a rate of 30 frames per second, timing of gestures and speech and the coordination between the two channels is accurate to within one frame or 1/30th of a second.\textsuperscript{3}

During the analysis of videotaped laboratory activities, I assumed that the way students view curricular artifacts is likely to be different than when viewed by a physicist or physics teacher. I therefore reconstructed as faithfully as possible students’ ontologies from their actions and utterances. Here, an ontology is understood as the sum total of the elements which are salient to students in their task-related world. Past research showed that these ontologies are likely to differ between students and teachers, and often among students (Roth, 1996a; Roth, Duit, Komorek, & Wilbers, 1997; Roth, McRobbie, Lucas, & Boutonné, 1997). Careful analysis of activities suggests that students’ responses are very often plausible within the experiential worlds salient to and as individuated by the student.

3. Representations, gestures, and laboratory talk

The conversations considered and analyzed here are distinguished from many other conversational settings in that they occur in the presence of scientific representations (drawings, computer-animated objects, 3D architectural models). This presence comes with particular affordances to the speakers and their efforts of appropriating a new, scientific discourse. In all situations, the speakers featured are within reach of the representation that is the focal point of the conversation. The conversations therefore evolve in the context of copresent interlocutors who share a place and moment in time. All of the episodes analyzed here feature discourse contributions in which one participant attempts to elaborate an understanding of the focal situation. In this sense, though the interlocutor is absent in these episodes (because the analysis focuses on different issues), all are examples of \textit{participatory acts} (Clark & Schaefer, 1989) performed by the respective agent to contribute to the collective act of constructing a public understanding of the issues at hand. The three episodes presented are characterized by the different types of representations over and about which the conversation occurred: a drawing of a pulley system on a chalkboard, a computer-animated diagram (including representation of phenomenon and scientific concepts), and a three-dimensional model of a bridge design. The three situations are organized to represent increasing complexity in the kind of feedback speakers have to the kinesthetic experience of their iconic gestures: no feedback,

\textsuperscript{3} The following transcription conventions are used:
(3.1) -- Time in seconds; (.) pause less than 0.2 seconds; \textit{Italics} -- to indicate stress of an utterance; why:: -- colon to indicate lengthening of phoneme; | -- marker for time reference; \textsuperscript{\textuparrow} -- coordination of utterance and video frame displayed.
visual feedback, and proprioceptive feedback, respectively. This increasing complexity also allows for an increased accountability (veracity) for what has been said. Furthermore, the order of the episode represents the temporal unfolding of learning: the first two episodes represent speech-gesture relations at the beginning of students’ explorations in a new domain whereas the third episode represents speech-gesture relations at the point when students are said to understand the science involved.

3.1. Gestures, Talk, and Drawings

In this section, I analyze an episode from a Grade 6-7 science class where students argued over and about scientific models (drawn on the chalkboard) of a tug of war mediated by a pulley system. I show that in the early stages of understanding physics, deictic and iconic gestures play an important role in the construction and communication of ideas. An example from a posttest is provided which illustrates that by that time, the verbal means have now taken on much greater importance in the communication of students.

3.1.1. Setting and ethnographic context

This course was designed so that students could develop discursive and material practices in authentic situations. In this class, students therefore learned about the physics of simple machines by designing and testing artifacts of their own design. To develop students’ discursive competencies, their designs were presented and critiqued in whole-class conversations. Periodically, the teacher (Roth) prepared activities to focus the whole-class conversations on specific scientific issues. In the present case, to help students develop scientific descriptions and explanations of pulley system, the teacher had invited the entire class to a tug of war. However, he had literally rigged the event by using a block and tackle that changed the force ratio by a factor of 5 in his favor. After the class (20 students) had lost the competition, the teacher invited students to describe and explain what had happened. In the initial part of the conversation, students and teacher drew different pulley systems on the chalkboard to describe the situation. But then, students began to construct models for situations in which the teacher would not have benefited despite the use of a pulley system. At the end of the 17-minute conversation, there were 8 different designs students and teacher had drawn on the chalkboard; these designs served as topic and background to the ongoing talk. The present excerpt is from the second part of the conversation featuring one student (Shaun) who, in the course of his presentation, constructed one alternative configuration. He had initially tried to explain his model verbally from the seat; a classmate then attempted to construct a drawing according to Shaun’s instructions but failed. Because neither teacher nor classmates understood what Shaun attempted to say, he then went to the chalkboard facing the teacher, who also stood next to the chalkboard. Both looked at the drawings on the chalkboard between them while the other students watched and listened to the interaction.
3.1.2. The episode

The episode begins as the teacher asked Shaun where in the diagram the different parties would pull (line 01) because, up to this point, Shaun’s drawing and talk had not provided this information (Figure 2). The teacher’s question, “Where do I pull? Which end do I pull?” (line 02) indicated that he did not understand where the string was attached or who was pulling at that end. Until then, Shaun had only attributed one rope, the one for the teacher. If the class was to be at the other end, it would leave one end of the rope loose. In a sweeping gesture (line 02), Shaun then attributed to the teacher the rope that attached the pulley to the banister.

01 Where do I pull? Which end do I pull?

02 You can pull on here.

03 With the, OK (3.8 s) this is a banister.

04 But we have the banister here.
05 Oh God

06 (5.2 s)
07 ((Erases diagram, begins new one))
08 Banister

09 OK

10 Long string

11 OK, pulley.
12 Roth

13 pull here

14 OK.

15 Then *there* is a pulley.

16 And where do you pull?

17 (1.7 s)
18 And then *there* is another banister

19 and then we pull (.) here

20 (Several students applaud)

21 You can’t, with this.
Visibly disconcerted that the teacher did not understand, Shaun started a new diagram (line 03), but the teacher stopped him short by pointing out that there already was a banister (line 04). That is, the teacher held Shaun accountable to the earlier drawing which already included a banister. By means of theatrical emphases, Shaun aligned the teacher and his peers to the key elements of his design and their placement (lines 06-19). Through his emphases in the uttered words and the accompanying deictic gestures, he rendered unequivocal the placements of exactly those elements that had been ambiguous during his earlier verbal presentations (though some of his peers apparently understood him even then). Shaun completed his drawing.

Unlike his earlier verbal presentation from the back of the classroom (which was difficult and even incomprehensible to his peers and teacher alike), his description in the form of a drawing was much less ambiguous. Furthermore, his design now existed in material form and was available for critique which had made him accountable. This critique was immediately instantiated. It came both in the form of Alain’s observation that the teacher did not have to pull at all and in the teacher’s subsequent applause which celebrated Alain’s comment. Alain’s comment had furthermore underscored a concern that had been voiced earlier in the conversation. Sharon had pointed out that the pulley was attached in such a way that the class was pulling on the banister rather than competing against the teacher.

3.1.3. Analysis

In this episode constructing a scientific model in the form of a drawing allows the student to use gestures in different forms to construct an argument even before he understands and can correctly talk about pulley systems. Here, pointing to an element of the drawing and drawing one while uttering the associated label both have deictic functions. Iconic gestures both indicate where the different parties in the tug of war pull (and therefore also have a deictic function) and enact the pulling (and therefore forces acting in the system).

At the heart of the ongoing debate was the exact placement that the different elements in the pulley system (i.e., support/banister, rope, pulley) should take. Deictic gestures therefore played the important role of identifying elements and their location in the system. Here then, deictic gestures isolated and individuated individual elements. Drawing an element did double duty in (a) individuating the element itself and (b) indicating its placement in the equipmental whole. For example, the teacher uttered “But we have a banister here” while touching the top part of the drawing (line 04). Both verbal and gestural action contributed to bringing this element (and the fact that it already existed) to

---

4 Again, the utterance as situated action not only indicated that there already was a drawing with a banister, but also suggested that the speaker should make his argument by referring to it rather than producing a new one.
the foreground. Whereas the teacher pointed to an existing element, Shaun made the element he talked about salient by drawing them. For example, he drew some lines on the chalkboard and uttered “banister” (lines 03, 07-08). Here, because he drew the element, it and its location with respect to the other features on the chalkboard stood out. This is even more evident in the case of the elements following the “banister,” for each new element was placed in respect to the already existing ones.

Following line 08, Shaun therefore established the domain ontology by drawing each element, attributing the actors to pull on particular ropes, and uttering the verbal expression associated with each element. In this case, the nature of the elements was especially salient because Shaun drew them on the chalkboard, which served as ground against which the diagram stood out. Furthermore, for each new element, not only the blackboard functioned as ground, but also the already existing parts of the emerging diagram. Individuating an element by pointing to it or by drawing it are therefore slightly different. In the first instance (e.g., line 04), the deictic gesture points to part of the drawing. However, much of the work of delineating the extension of the element and thereby isolating it from the remainder of the drawing has to be done by the listener (observer). When Shaun drew an element such as “long string,” the gesture actually pointed to and thereby individuated the element in its entirety.

In addition to the deictic gestures, Shaun used gestures where his hands moved along a particular string. These gestures did double duty as they were both deictic and iconic. While Shaun indicated that someone (teacher, students) “pull[s] (on) here” (lines 02, 13, 19, [21]), the gesture both made salient where the agent was to hold on (deixis) and in which direction to pull (iconic gesture). These gestures, which enacted pulling in iconic ways are further elaborated in the video off-prints (Figure 3) of line 02. As Shaun uttered “you,” his hand was already in upward motion reached the rope above the pulley while pronouncing “pull on,” and then continued the motion along the rope until the completion of “here” (Figure 3). Thus, in this case, “here” is not a simple deictic device to point out the element on which someone pulls (e.g., “long string”). But “pulling here” is enacted together such that the gesture makes salient both the element pulled at as well as the direction of the pull. Thus, he was not just indicating some element where there is some force, but the forces appear in his gesture which iconically enacts them. At this point, the sweeping gesture not only pointed out where the teacher would be pulling, but also the direction of the forces. These gestures therefore “animated” aspects of the drawings and exhibited the dynamic of the system part in addition to declarative aspects of the drawings at hand. Thus, talking in the presence of the diagram allowed the participants to express important typological features of the situation, including forces and their direction, which the topological nature of talk cannot easily express. It is noted that in the absence of any feedback from the drawing, Shaun cannot know whether the system actually operates and mediates forces in the way he argued. Thus, whereas he enacted the individual forces accurately, his presentation failed to portray the system in a scientifically accurate way.
You can pull on here

Figure 3. Video-off print from conversation where Shaun explains and enacts the pulling forces in the simulated tug of war competition.

The episode is also interesting in terms of the changing origo of deixis. When Shaun enumerated the elements, he always placed them “there,” in front of himself, the teacher (opponent in argument), and the class as audience. But when he suggested where the different parties pulled, he always indicated them with the discursive deictic “here,” taking the perspective of the agent pulling at the place indicated. In addition, the swiping movement of the hand also indicated the direction, and in fact, enacted the pulling and therefore took a first person perspective of the event. With a shift occurring according to who he presently talked about. The inside observer perspective added flexibility to distance by bringing the listener into the space of the narrative, which heightens both the narrative and the rhetorical effects (McNeill, 1992). There is the opposition in “you,” and the shift in perspective to “here.” Phenomenologically, “you” designates the Other opposite to the Self; simultaneously, objects and others are placed “there,” whereas the Self is placed “here” (Ricœur, 1990). In this, we have a blending of the animate (student) and the abstract (inanimate) physical event; Shaun makes interpretive journeys from the intertwined constructed realms of (a) the world of physical events and (b) the world of the visual representation of the objects/events (Ochs, Gonzales, & Jacoby, 1996).

3.1.4. Discourse appropriation and learning

In mechanical systems, the direction of the forces are quite important, for the physics of these systems is given in terms of the vectorial properties of the concept of force. These vectors add, in the equilibrium case, to a net force of 0 (Newton's) in the system. Should the forces remain unbalanced, the system would move into the direction of the net remaining force. The cognitive outcomes of these lessons, particularly the effect of talking over and about scientific representations, has been analyzed in great detail elsewhere (Roth, 1996b). Here, I provide but one piece of evidence for the competence Shaun developed for analyzing such diagrams in terms of the forces acting in the system they represent is documented below.
Pretests and recordings of conversations early in the unit showed students’ limited competencies for describing and explaining simple machines, including pulley systems. Their utterances were brief and demanded considerable time and support before they were actually audible. In contrast, the posttests at the end of the unit (including writing and talking about and manipulating mechanical systems) provided evidence of the remarkable discursive competence students developed. After they had completed their posttests individually, Shaun and Alain were invited to discuss their respective answers to a question in which they had to identify the forces in a pulley system. (The system involved three forces related in the following way: $F_1 = F_2$, $F_1 + F_2 = F_3$.) The following transcript provides evidence that the students still referred to the drawing in front of them, using deictic and iconic gestures, but their discourse now has shifted much more into the verbal modality, identifying the magnitude of forces, analyzing their relationship, and producing different scenarios that would change the amount of weight the person (hand) would have to hold. By this time (at the end of the unit), the students had evolved ways of talking about the simple machines and mechanical devices more generally that stood in stark contrast to their earlier stumbling talk.

01 A: This [ceiling, $F_1$] takes half and this [$F_2$] takes half, so I got 55 grams, I didn’t measure the scale, I forgot about the scale, I put ahm, 10 grams for the pulley and a 100 grams for the weight, so I got a 100 and 10 and like, so there is 2 scales, so I divided by 2 and I got 55.

07 S: And I got a 120, I just add them [$F_1$, $F_2$] all ‘cause it would take the same. He is not asking what they both hold at the same time pulley, he wasn’t saying like if you pull them both at once, what, ahm, what they would both come to, like, if you were pulling one of them and just holding the other like that, then that’s what the scale come to.

12 A: OK, what I have to say is, see how it’s tipped to the ceiling?

13 S: The ceiling is taking half of the weight.

14 A: Exactly and it shows on Scale 1 [$F_1$], so Scale 1 [$F_1$], gets half of it and this gets half of it, it doesn’t, I understand what’s you’re saying but I think that if this wasn’t connected to the ceiling, then I will agree with you that it would be 100 and whatever you said but.
Here, Alain and Shaun argued whether the scales in a diagram should read 55 or 60 grams (F₁ and F₂). Both accounted for their written answer, and in arguing for their own solutions, revised their previous answers. Alain and Shaun maintained this conversation with turns of considerable length and without further prompting which attests to their discursive competence. In the end, they agreed on the scientifically correct answer and added that, if the pulley had some weight, this too would add to the magnitudes of F₁ and F₂. Here, Shaun not only identified the nature of the forces, but correctly indicated their magnitudes and directions and therefore, their relationship.

While they explained their answers, both students used deictic and iconic gestures tightly coordinated with their utterances to make their points. For example, while Alain uttered “so there is 2 scales [F₁, F₂],” both of his hands simultaneously moved above the drawn pulley system downward along the two scales and strings. The finger tips of each hand touched as if he was holding a string, much as he had done during the laboratory investigations with the pulleys. In the same way, as Shaun uttered “you were pulling one of them,” his right hand moved upward along the corresponding scale in the drawing and string thereby enacting the pull [F₂] required to keep the system in balance. As he continued, “and just holding the other like that” the fingers of his hand closed as if he was holding one of the strings, the hand placed above the point where the string was fixed in the drawing.

3.2. Gestures, Talk, and Computer-Animated Drawings

When the scientific models are animated, the relation between gesture and ground (diagram) is different than in the previous case with still diagrams. Because the computer-animated diagram itself changes through time, iconic gestures can now be visually compared to and evaluated against the changes in the ground. Traces of these changes may still be visible on the monitor, or changes may be simultaneous with the communication, or changes are invisible because they have occurred prior to or will occur after the communicative act has been completed. In this episode, gestures are not merely incidental but play a central role in the communicative efforts prior to the emergence of suitable vernacular descriptions on the order of seconds, but precede correct scientific discourse on the order of weeks.

3.2.1. Setting and ethnographic context

The present episode was recorded in a junior year physics course. This course was designed so that students could develop competencies in the discursive and material practices of physics. In addition to designing and conducting their own experiments, students were provided with activities that highly encouraged theory talk including the construction of semantic networks and explorations of microworlds. The present episode
derives from one such activity where students explored a computer-animated Newtonian microworld (Interactive Physics™) for the purpose of constructing their own theories of the events. The computer-animated microworld interface provides users with a range of tools to set up experiments and to measure, display, and record a range of motion events and physical quantities. The microworld really layers phenomenal and a theoretical-conceptual worlds by displaying both phenomenal objects and conceptual quantities as force and velocity vectors. In the present case (see Figures 1, 4), there was one circular object characterized by its velocity (bold single arrow); there was a second arrow force (outline arrow). The students’ task was to explore the behavior of the objects and to find descriptions and explanations for the arrows and their relation to the motion of the object on the screen.

3.2.2. The episode

In this episode, three students (Gaalen, Linda, and Randy) sit in front of a computer monitor (seating arrangement displayed in Figure 1). At the moment of the episode, they still had no grasp of what the arrows stand for and how they related to the moving object. The students had previously associated them with time, energy, time step, and many other ideas. Here, then, Gaalen attempted another description and explanation of what they just observed (still visible in the top left slide). While uttering the sentence “Wouldn’t the length of the arrows (1.60) Since that arrow’s longer, the velocity is higher. That’s how it is pushing that’a way” both of his hands enacted the arrows and their behavior as he had seen it previously. Gaalen had held his right hand with fingers parallel to the outline [force] arrow—in the form more clearly seen in the second frame—for 3.47 seconds prior to specifying its referent in Frame 2 ([1]-[3] in Figure 4). He then made another, 0.10-second gesture which marked the right hand while uttering “that arrow” which immediately preceded the causal meaning unit “that’s why it is pushing it . . .” Before he uttered “the velocity” ([4]), his left hand can already be seen at the top left of the image paralleling the velocity (single bold) arrow on the screen. Frame 3 shows the two hands set up, the right parallel to the force, the left parallel to the velocity. In the next frame, the right hand was already “pushing” against the left hand which is moving off the frame to the left a movement which continues to the end of the sentence (off screen). This pushing motion of the right hand began 0.83 seconds before its verbal equivalent “pushing.” Here, the shape of the object’s trajectory (visible in Frame 1) which he attempted to explain was already completely described by the iconic gesture (and trajectory) of his left hand. The episode is complex because there is one term “arrow” but two arrows on the monitor, and the same words “that” and “it” occur repeatedly but have different referents and function.
Wouldn’t the length of the arrows (1.60) Since that ‘s longer the velocity is higher
the arrow (1.47) ↑ (0.33) [(0.10)

that’s why:: it’s pushing it that’a way.
↑ (0.20) ↑ (0.53) | (0.83) ↑

**Figure 4.** Video-off print from conversation where Gaalen attempts to describe and explain the motion of an object in a computer simulation and the relationship of two arrows to the moving object.

### 3.2.3. Analysis

Among students, there was considerable variation in the words and phrases for the different elements in the microworld, and in understanding how these elements (object, arrows) interact. In this group, students referred to the outline arrow (force) using eight different verbal terms {little arrow, big arrow, time-set, time, direction, time-and-direction, velocity, redirection, gravity, force, reverse-gravity, gravity}. Similarly, they referred to the other arrow in eleven different ways {little arrow, big arrow, initial speed, velocity, force, effort, strength, speed, direction, speed-and-direction, velocity}. In this seeming chaos, deictic and iconic gestures were crucial to establishing common ground, finding appropriate descriptions, and ultimately, arriving at a theoretical discourse that was consistent with Newtonian physics. Thus, in the present episode, Gaalen’s gestures already provide descriptions and explanations which are more appropriate than his verbal descriptions. Here then, his gestures considerably preceded his verbal representation of Newtonian physics. In this situation, gestural (and verbal) deixis was crucial in coordinating utterances, gestures, and the phenomena in the microworld.

“That” appears three times, and each time not only the referent but also the function is different. In the first instance, “that” ([3]) has a deictic function designating a particular arrow standing in opposition to the speaker (distal use). Coinciding with the utterance, the right hand which had moved to the right, came to a sudden stop. As can be seen from
Frame 2, the fingers of the right hand stand parallel to the outline arrow (force). This finger position, the noticeable (abrupt) stop of motion, and the coincident utterance “that arrow” lead to a convergent interpretation that the referent of the right hand is the force arrow. The listener can draw further confirmation for this interpretation from the causal connection between “that arrow” and the force arrow because it is the one that they had previously manipulated, whereas the other arrow only changed as a function of their action.

In the second instance, “that” introduces the causal consequence (“that’s why”) of the hand arrangement he had set up and described in the previous part of the utterance; “that” falls at the beginning of gestural trajectory which iconically re-represents the earlier visible trajectory (Frame 1). Finally, in the third instance, “that” is linked to “way,” the immediately preceding trajectory (”way”) enacted by the gesture. In vernacular, “that way” most frequently expresses a specific direction. Here, however, “that a way” together with the curved motion of the hand, when read against the ground of the earlier curvi-linear motion of the object and the corresponding positioning of the arrows, highlights not only the existence of the trajectory but in particular its curvi-linear shape.

The indexical locution “it” was used twice, but gesture and ground help disambiguate the two referents (“It’s pushing it...”). The utterance occurred while the right hand followed, fingers pointing to, the left; heard together with “It’s pushing it,” the right can be understood as literally pushing the left hand (Frames 3-6). Here, the first use has as referent the hand/ arrow which is pushing (enacted by the hand) and the second occurrence has as referent something that is being pushed which, in this case, could be the second arrow/left hand, or the object.

At the time of this episode, Gaalen (as his two peers) did not yet describe the arrows in scientific terms, that is, as force and velocity. He used the appropriate scientific (verbal) discourse only two weeks later during the subsequent lesson with the microworld. However, in the present episode, his gesture correctly describes the relationship between the concepts of velocity and force as the arrows represented them. He characterized the action of the outline arrow as “pushing,” which is a vernacular form of describing forces. Finally, he associated the longer pushing arrow with a resulting higher velocity. Here, the referent of “velocity” is not completely clear and two readings are possible. Because the utterance coincided with the positioning of the left hand, “velocity” can be heard as the referent to the left hand: therefore, the longer force (right) arrow pushes more and therefore leads to a longer velocity (left) arrow. But the fragment “Since that arrow’s longer the velocity is higher” could also be interpreted such that the longer right arrow is equivalent to a higher velocity in which case “velocity” would have been anchored in the right arrow (incorrectly so from a scientific perspective). However, the nature of the referent for each of the two hands was disambiguated by their position in space in the course of the motion. The fingers of the right hand kept a constant direction just as the outline (force) arrow, whereas the left hand changed direction, though
less rapid so, in the way the single-line arrow did previously. Thus, Gaalen’s gestural
description and explanation of the events was scientifically correct long before his verbal
explanations.

McNeill and Levy (1982) suggested that gestures were generally coordinated with the
verb of the accompanying utterance and that there was little delay between the verbal and
gestural meaning units. In the present situation, this means that the gesture of the pushing
action began 0.73 seconds before the utterance of the corresponding verb. Furthermore,
the deixis in “that’a way” is at the very end of the 1.56-second iconic gesture of the
trajectory in which it was grounded. As such, this situation can be read as one in which
the gesture primed the associated utterances (“pushes,” “that’a way”). On the other hand,
the entire utterance can also be heard as a nested logical argument of the type ([longer
force arrow] IMPLIES [higher velocity]) IMPLIES [being pushed through curvi-linear
trajectory]. In this case, the gestures coincide precisely with each part of the argument.
As part of the premise, deictic gestures identifying force ([3]) and velocity ([4]) arrows
(see Figure 4) fell together with their verbal equivalents. However, the consequence of this
premise “longer arrow —> higher velocity” is an event extended in time and space so that
we might search for a temporal coordination of gesture and the utterance of the entire
verbal meaning unit. In this case, we observe a precise coordination of gesture and
utterance.

At the moment Gaalen reenacted the motion, the actual “flash photograph” of the
event was no longer available. Thus, the participants could not directly compare the
iconic gestures (with a curvi-linear trajectory) with the event they were trying to explain.
Nevertheless, there was from an analyst’s perspective a high congruence between the
gesture, the position of the arrows, the trajectory, and the circular object. This is
significant, for what students see and remember having seen is strongly related to what
they expect to see (this is so strong that we were able to document strong disagreements
between students whether or not a teacher demonstration showed motion or not [Roth,
McRobbie, Lucas, & Boutonné, 1997b]). But students too were able to evaluate their
gestures either retroactively against the trace on the screen they remembered, or
prospectively with a trace that would show with their next investigation. Thus, compared
to the chalkboard diagram in the previous case study, the computer-animated diagram
afforded to the conversation additional features in the indexical ground for using iconic
gesture and comparing these to events. Furthermore, this system allows users to record
the events like a single-exposure multiple-flash photography: students can see in one
image the the different positions the object had. This then makes ephemeral aspects of the
motion, such as the “trajectory,” available in visual form which, in turn, provides new
resources to the communicative effort, particularly to the iconic gestural channel.

3.2.4. Discourse appropriation and learning

Despite these uncertainties, and despite initial use of the same term for different
objects, or different terms for the same objects, and despite the potential uncertainties
associated with the shifting referents of the same deictic terms, students ended up using appropriate scientific language in consistent way. Thus, out of the conceptually “muddled” talk in the beginning, students developed consistent and scientifically correct forms of communicating about the microworld events. I presume that the presence of the computer-animated diagrams which afford deictic and iconic gestures (that are the transitional elements) facilitates students’ trajectory from one to the other form of talk.

The three students’ competence to describe and explain the computer-microworld phenomena in verbal form exists in two forms. (Verbal expressions of scientific concepts and theories are predominant forms in the scientific literature and even more so in school science.) First, posttests showed that the students used appropriate scientific language. Furthermore, an episode recorded 2 weeks later provided evidence that the students’ understanding was consistent with physics. Linda described the situation on the display as a world governed by “anti-gravity”, that is, a world in which objects fell towards the top of the screen. She consistently used the word “gravity” to describe the force arrow, and in two instances used the prefix “anti-” to indicate that the direction of gravity in the microworld was opposite to that in her physical world. Similarly, Randy and Gaalen used “gravity” and “velocity” correctly to label the respective vectors. The relative position and length of the force and velocity arrows before their attempt to achieve the desired trajectory indicated that the students not only talked about the vectors appropriately, but that they also understood the relative effects of the two on the motion of the circle in their microworld. Later in the transcript they described the shape of the trajectory as “parabola” caused by a gradual change of velocity over time.

We’ll design the experiment with Newton’s law including opposing forces, one has two equal forces, there’s no velocity, and the next one has unequal opposite forces they will accelerate in the direction of the greater force.

Their actions in the microworld also coincided with their verbal counterpart. Thus, students hypothesized that the object had to “go faster” which was followed by a lengthening of the velocity vector. An utterance that they needed “greater gravity” was followed by the lengthening of the corresponding arrow. Students therefore do not just manipulate words but their experimental practices deeply integrated the scientific meaning of these vector elements. “It needs a tiny less speed” was followed by a shortening of the arrow but not of the direction. So from the “it’s pushing it” enacted by the hand that picked out the force vector, students developed to use “force” as the linguistic marker for the arrow. Similar developments, and with about the same number of alternate names were observed in the other student groups recorded during this investigation (Roth, 1996c).

---

5 In two companion papers which only focused on the verbal utterances, I described in moment to moment detail the emergence of scientific discourse from students early attempts to come to grips with the descriptions and explanations of the microworld events (Roth, 1995a, 1996c).
3.3. Gestures, Talk, and Three-Dimensional Models

Compared to the still and animated drawings in the two previous sections, gestures play a different role in the presence of three-dimensional models. Here, the model may also provide proprioceptive and visual feedback to the speaker. Gestures therefore not only “animate” the model, but gestures become part of the phenomenon.

3.3.1. Setting and ethnographic context

The present episode was recorded in a Grade 4-5 classroom where 9 to 10-year old children learned about forces through engineering design activities. Central to their activities was the design (including planning, constructing, prototyping) of model bridges, towers, and dwellings. Because the focus of the curriculum was on the appropriation of engineering discourse, students presented their design artifacts to the class (and during a public exhibit at the local science museum) followed by an analysis-question-critique session. Two cameras (supplemented by ethnographic observation of two researchers) recorded the 13-week unit. The episode featured here was part of the final sessions where all groups presented their bridge designs. Knowledge was assessed six weeks later, using posttests in the form of semantic maps and glossary entries were used to assess students knowledge. In the episode analyzed here, Jeff (Grade 5) and Jon (Grade 4) presented their bridge design. As they pointed out in the course of their presentation, they were interested in the design per se which would hold up the most weight (as measured by the number of wooden blocks in a bucket) rather than in designing a bridge for a particular purpose. In the course of designing, they went through several cycles of prototyping and testing to increase the load from an initial 50 blocks to almost 357 blocks.

3.3.2. The episode

Figure 5 shows Jeff in the process of describing the function of the posts at each end of the design artifact. In fact, in this episode, he repeated what he considered to be the effect of the upright pillars on the bridge truss. In the first pass, his right hand moved down alongside the pillar ([2-3]) beginning with the utterance of “weight.” The movement of the hand then shifted to describe a horizontal action of another force ([3-5]) and then waved twice more along the horizontal truss. After a 2.17-second pause, he repeated the meaning unit, but this time enacted the weight by pushing with both hands down at the end of the bridge ([6]) which straightened the curved truss (tension on top, compression on the bottom). He then released both hands from the bridge deck, and moved his right hand index finger across the truss, while verbally describing the action of the tension on the bridge deck (“the tension brings it down”).

23
Figure 5. Video-off print from conversation where Jeff presented a bridge design that supported the most weight of any bridges in the class.

3.3.3. Analysis

In contrast to the drawings, the model Jeff constructed and now talked about is physical which affords new possibilities to deictic and iconic gestures. In his first description, the hands rested at the upper part of the pillar, until the utterance of “the weight” (\([2]\)) and then, while uttering “down on the end” moved downwards alongside the pillars stopping at the end until completion of the verbal meaning unit. The first gesture along the horizontal deck fell together with the utterance of “it’s easier for the force” (\([3-5]\)). At this point, the verbal meaning unit was not yet completed. Thus, while completing the meaning unit (“to go across here”), the hand moved twice more back and forth along the truss. Here, the gestures represented the forces. (As one reviewer pointed out, we can also think of the “to go across here” and the repetition of the gesture as assuring the demarcation of the second rhythmic-semantic unit of the sentence.) From a physics perspective, forces are vector rather than scalar quantities. Technically therefore, they cannot be pointed to, but have to be indicated in terms of magnitude and direction. Here, in this first part of the episode, Jeff gestured the direction in which the weight of the pillars acted. The gestures were iconic representations and their extension in space was paralleled by the time it took for the utterance of the corresponding meaning unit. When the gestural meaning unit was completed before the verbal meaning unit (at \([5]\)), the right hand repeated the action of the horizontal force.

Jeff then reiterated the meaning unit of the relationship between pillar weight and tension in the upper part of the truss. In this, the effect of the weight down and force across became even more clear when Jeff actually pushed down on the ends of the bridge.
deck. Again, the downward motion of the hands began with the utterance of “weight.” Now, however, Jeff actually pushed down onto the ends of the bridge ([6]). This straightened the bridge deck. At the same time, although small, the bridge pushed back so that Jeff had feedback from the reaction force to the weight. He then released the bridge deck (thereby allowing the bridge to bounce back) and, while uttering “and then,” brought his right hand into the position for explicating what the consequence of the weight is. (The left hand was simply used to hold the bridge which began to rotate as the right hand moved along the truss.) The beginning of the horizontal gesture then coincided with the utterance of “tension.” In this situation, the end of the verbal meaning unit fell together with the end of the horizontal gesture and there was no repetition as in the earlier cases.

In this case, at the end of 3-week investigation and construction of bridges, Jeff had developed a sophisticated understanding of forces in bridges. Though he was not entirely correct in suggesting that the tension brings the deck down, his gestures enacted the forces and their relationship. His gesture therefore enacted an explanation of the form \( W \rightarrow T \), the downward weight implies or causes the horizontal tension. While Gaalen and his peers were at the beginning of investigating their domain, Jeff was at the very end. It is therefore consistent with the literature on the developmental trajectories of gesture and talk (e.g., Freedman, 1977; Hayward & Tarr, 1995) that Jeff’s verbal scientific argumentation coincided with the corresponding iconic gestures. However, these gestures should not be considered ancillary to the verbal meaning unit. Because of the vectorial nature of the physical concept of forces, it is actually the downward gesture (and the force Jeff felt from the bridge deck tension in truss resisting to bending) which expressed the appropriate topology of this phenomenon. From an audience perspective, Jeff’s gestures in the first part of his explanation are similar to the gestures that were associated with the chalkboard diagrams in Section 3.1. The three-dimensional object, however, makes the identification of the physically significant elements more difficult: whereas in a diagram, only the physics-relevant elements were drawn, the artifact did not embody such discourse-relevant selections. On the other hand, Jeff’s second explanation in which the gesture actually pushed on the bridge showed the effect the downward weight has on the bridge deck. Thus, Jeff’s utterances could be directly compared with the changes in the system.

3.3.4. Discourse appropriation and learning

Although there was only facilitation but no direct instruction by the teachers (e.g., lectures, demonstrations) on this topic, Jeff and his classmates developed sophisticated discourse (Roth, 1998a). Our posttest shows that he competently used scientific representations (drawings, language) to explain tension and compression (Figure 6). Jeff’s verbal explanation for compression suggested that it arises from forces in two directions. His drawing elaborated the nature of these forces. First, they act in the bridge from left to right and right to left along the bridge structure. In the center, then, the downward force of the bridge weight is counterbalanced by the upward ends of the forces that act in the
structure itself. Similarly, the text of the “tension” glossary entry identifies it as something that arises from pulling. Furthermore, both tension forces are indicated by arrows parallel to the ropes (cables) in which they act. It should be noted that in both instances, the gestures were precursors to the signifying arrows drawn later. The arrows are the abstractions of the earlier actions where Jeff’s hands and index finger moved through horizontal and vertical trajectories now represented in the arrows. The multiple arrows appear to reconstruct the repeated movements of his index finger across the truss.

Figure 6. Jeff’s glossary entries in which he constitutes his understanding of compression and tension forces in bridges.

4. Discussion and Conclusion

The episodes presented here portray general patterns of speech-gesture-ground interactions in a large data base of learning in school science laboratories. Each episode makes evident the complex and shifting interactions that exist between expressive means (speech and gesture) and the entities talked about. Because the entities are physically present, they partially represent themselves such that the conversational participants do not have to express what is evident and therefore goes without saying (Roth, in preparation). In the following, I focus in greater detail on the relationship of gesture on the one hand and representation, cognitive development, and schooling on the other. I conclude with suggestions for possible avenues of future research.

4.1. Gesture and Representation

In the three episodes, the role of iconic gestures was different. When they occurred in front of a fixed drawing, gestures provided an image of what might occur if the depicted situation were implemented in the world. Gestures therefore “animated” the drawing. When the “same” gestures occur in front of a “runable” computer display, they can be compared to the events on the screen either retrospectively (such as when gestures
iconically represent a past event) or prospectively (such as when gestures anticipate future events, and therefore constitute hypotheses yet to be tested). In both situations, gestures represented the phenomenon iconically. More so, the gesture engaged students’ sensorimotor system and therefore lead to and arise from proprioceptive representations. Forces and motion are not only encoded in symbolic, but also in kinesthetic form. The difference between the two situations lies in the possibilities for visually comparing the enacted motions and those that occur in the world. These two situations considerably differ from the third situation where gestures were produced in the presence of a three dimensional model. In this case, the gesturing individual actually receives feedback during the execution to the sensorimotor system. Thus, when Jeff pushed down on the pillars, the bridge provided resistance; there was a visible bending in the artifact which bounced back when Jeff removed the pressure. Here the gesture included putting pressure on the bridge which coincided with the physical phenomenon. The gesture not only depicted pressure, but also enacted the pressure.

The problem with static diagrams on chalkboard and other media is that, although individuals can enact and gesturally represent forces, the systemic relationships between the different elements may not be discovered. Thus, although Shaun gesturally enacted the forces by the different parties, the relationship between these forces was not evident in his talk or gestures. Thus, as his classmate Alain pointed out, the concrete model he developed (involving drawing, descriptions, and enactions) did not really work in the sense that it would provide the teacher with an advantage. Such relationships between elements can be tested, and thereby enacted when the models can be run (animated) in computer microworlds. In the case of the computer-animated drawings, hand positions and gestures were important initial descriptions of the events on the monitor which, in their scientific accuracy, preceded those verbally provided by the students. Gestures therefore could always be checked in regards to their fidelity with the events modeled by the computer. On the other hand, my research also showed that these students did not transfer their knowledge from the microworld to real-world phenomena. Here, then, arguing over and about three-dimensional models provide additional affordances to speakers and listeners in that the outcome of gestures that enact forces can be visually checked. Furthermore, the speaker can also feel all the reaction forces that arise within the system in response to the gestures. However, in these three-dimensional models, there are no representations of the physics concepts (as forces and velocities in the computer-animated models) so that the veracity of Jeff’s uttered conceptual talk could not be tested by him or the audience.

From a cognitive perspective, talking over and about physically-present models of scientific phenomena affords codings other than verbal. Thus, all students in the previous episodes used iconic gestures. These, in their very nature, permit kinesthetic coding and knowledge of physical systems. Using gestures as part of making scientific arguments therefore allows for additional modes of coordinating meaning (common ground) in groups. However, whereas the correctness of the kinesthetic representations of physical
events cannot be tested in the presence of drawn models, computer-animated models allow for visual feedback for correcting the kinesthetic expression. Finally, the three-dimensional model, in addition to visual feedback, gave the speaker proprioceptive feedback about the reaction forces acting within the system in response to his/her gestures.6

The results of the present study encourage us to expand notions of knowing which focus solely on articulated knowledge (e.g., Anderson, 1985) and to include kinesthetic coding as a central aspect of cognition and language. For example, the model proposed by Shaun is not entirely located in his head, nor solely his body nor in the diagram. Each of these components is part of a model for the pulley system, so that the cognition is at best distributed across these components. That Shaun did not have a mentally imaged the diagram can also be seen from his repeated look at other diagrams previously drawn by the teacher and his peers, and which were still available on the board. There is no evidence that he had mentally envisioned the diagram as a whole; he enacted the forces, as we noted, which is clearly a physical and physiological representation; and the diagram as a whole clearly first appeared in the diagram on the board where it was visually available in its entirety. The communicative effort was therefore distributed across these different modes which considerably differed from an earlier effort which not a single person present did understand the model Shaun suggested. (In this case, one student who thought that he understood failed to draw Shaun’s model, even with additional instructions provided [e.g., Roth, 1996b]).

There are other affordances of gestures as well. The pointing gestures accompanying verbal deixis (this, it, that) may allow students, during the initial phase of learning, to express themselves more rapidly by unloading aspects of cognition into the environment. Thus, rather than having to search for the appropriate but unfamiliar words, verbal and gestural deixis allows students to communicate rapidly and efficiently. The activity is allowed to continue rather than turning into an exercise of finding the right word. Clark (1996) has argued that indicating and describing-as draw on fundamentally different processes and cognitive resources. There is evidence that cognitive processes related to some model are much faster when they use physical models and recognition rather than representation and mental simulation of the same process (Kirsh, 1995). In the present study, especially the analysis of Gaalen showed that the deictic gesture appeared considerably before the associated verbal associate.

4.2. Gesture and Cognitive Development

6 Coding other than by memory can occur in many places. Many readers will have had the experience that they could not remember and explain how something works on a computer in the absence of the computer. Also, years after completing my Ph.D., I could not remember and say my advisors phone number, but could dial it and knew when I misdialled based on the perception of the touch tones. I could later “dial” in the absence of the phone and, by looking, “recall” the phone number.
There is little research to tell us what the exact psychological mechanisms (particularly biologically-plausible ones) of this change from a topological (iconic) gesture to typological (symbolic) gesture are. There appears to be some evidence that iconic gestures become more symbolic in the process of development; from being reenactments of events they increasingly serve as signs of actions (Hayward & Tarr, 1995). There is also evidence that expression of ideas in words takes longer to produce and is more elaborate and demanding task than the expression in gestural form (Kendon, 1985). In another study, I described the microsteps of such a development over time which were documented in video records of Grade 7 students working with levers. Here, children initially provided locations (and the deictic terms “here” and “there”) of weight where the correct response would have been the distance from the fulcrum (as Piaget had suggested). As the children moved weights on the balance beams, their descriptions evolved to indicate changes in positions (“from here to there”) and later, with more experience indicated differences in position resulting from moving the weights (“between here and there”). Still later, they began to reference the positions of the weights in terms of relative distance from an absolute marker, e.g., the end of the beam before consistently using absolute distance to the fulcrum as the relevant variable to be attended to (see Roth, 1998b). In the light of that study, the bodily experiences allow students who enact the phenomena at hand, to evolve iconic gestures into symbolic gestures; these can then be accompanied or substituted by a typological representations (words, arrows). That is, we can expect some more or less slow transition between the topological representation of iconic gesture to the typological symbolic gesture and its representation in language and drawing. Language therefore appears to have a deep connection with the experience of inhabiting the world with a physical body.

Most iconic gestures are easier to formulate and execute than the words they are to be integrated with; iconic gestures often anticipate verbally equivalent utterances (Clark, 1996). I showed that, especially in the case of Gaalen and computer-animated diagrams, there were substantial delays between gestures and the corresponding speech features that followed at the early stages of investigating motion phenomena. In this context, Clark notes that the rarer a word, the farther it lags behind its gestural equivalent. In developmental terms, this claim may find its equivalent in the observations made here. Because students are unfamiliar with scientific discourse (they have trouble finding appropriate words), their gestures considerably preceded their verbal discourse. The presence of the materials, leading to conversations that are over and about, may facilitate the appropriation of physics knowledge. Clark noted that describing-as, indicating, and demonstrating rely on different cognitive resources for speaker and listener. In the presence of objects (and drawings models), the modality of the discourse is likely to shift to indices and icons. The high frequency of verbal deixis in laboratories may therefore not be a surprise, and is likely disambiguated by the accompanying modalities.

---

7 This was confirmed in another, more recent study about learning abstract concepts in static electricity (Roth & Welzel, 1999); here, students experiences with the phenomena provided the sensorimotor and visual experiences as a ground on which the abstract theoretical descriptions could be mapped.
4.3. Gestures and Schooling

In laboratories and situations where the conversation is over and about graphic representations inscribed in some medium, it is the character of gesture as a form of graphic representation that is being exploited. In these situations, utterances are about the inscription but the gesture is over and about the inscription. Gesture constitutes a support system for making salient those aspects that the talk is about, but it also is an expressive medium that enacts (in iconic similarity), that which the inscription is about. Gestures (pointing, iconic enacting) do double work. They bind the utterance to the inscription, but also express something about the inscription. Gesture allows for the representation of (aspects) of physical experience in the world that can be represented in words at best only indirectly and sometimes not at all (e.g., it is difficult to represent continuous change in words, and spatial arrangements cannot be accomplished in discrete aspects of spoken utterance).

From an educational perspective, gesture may be an underused resource in the early stages of discourse development and for producing expressions when students have not yet appropriated disciplinary forms of talk. Gestures as additional conversational resources are especially important given recent conceptualizations of learning through active student participation in ongoing, domain-relevant conversations (Lemke, 1990). For example, Shaun’s substantial development of discursive competence may have arisen from the possibilities which the artifacts and drawings in his classroom provided to early articulation of scientific ideas. He established the domain ontology by pointing or drawing on the relevant aspect. Shaun “ran” his model by enacting the forces and their direction. Thus, the meaning of this diagram appears in a double sense: at a linguistic level, deictic terms and descriptive labels and, at a sensori-motor level, the force which is experientially realized (if not in magnitude so in direction). Gestures are therefore not additional components of utterance, but integrated because they are referring to or enact what the spoken does not express.\(^8\) It is precisely the mutual cointerpretability of words and gesture that binds them together into single units of utterances. The task of educators would then consist in facilitating students to shift the proportion from gestures to words, and thereby transition to an increasingly abstract representational medium. In the evolutionary process, the gesture itself therefore retreats more and more into the background, taking on functions to highlight the objects and events before the collaborators eyes and thereby take on more sign function which may stand independent of the signifying individual.

This and other research (e.g., McNeill, 1992; Suchman & Trigg, 1993) documented a high incidence of gestures when individuals deal with unfamiliar situations. If gestures are

\(^8\) Mathematicians use symbolic gestures to express, via metaphors, abstract mathematical concepts (McNeill, 1992). McNeill describes a video record where two mathematicians discuss some complex issue where in one situation, the hand gesture belied what the words said, but were correct and consistent with the gesture and words of the other mathematician.
so frequent when individuals (research scientists, science students) tread new conceptual terrain, this may be an indication that these modalities allow individuals to distribute cognition across different modalities. In fact, part of the representations are available in the context and do not require the same cognitive resources (much like a note pad, scratch pad decreases the cognitive load in calculating). However, science is taught almost exclusively in terms of words and concepts thereby privileging the typological dimensions of scientific communication. Although labs are used, these are to “reinforce” learning in the typological domain. However, I showed here that interacting in the presence of artifacts, diagrams (elsewhere for graphs) provides resources that might well be crucial stepping stones in the concept formation.

4.4. Implications and Further Research

In this study, I showed how gestures are important parts of the knowledge students develop and how gestures precede verbally represented knowledge in short and long terms. Research has to be conducted to study the breadth and depth to which this phenomenon is generalizable. For example, to what extent can or does highly developed kinesthetic knowledge (e.g., football quarter back) support the development of conceptual understandings of physics via the bridge of iconic gestures and situations in which proprioceptive feedback during explanation constructions is provided. Furthermore, all of the phenomena studied by the students in the presented episodes are related to motion and forces (i.e., the mechanics and dynamics). However, many phenomena in physics are not visible nor can they be felt. Future research should therefore be conducted to find out if assistance in the use of metaphorical gestures could scaffold the development of scientific language. Finally, the present study suggests that schools (and universities) may be ideal “laboratories” for developmental linguists to study the genesis of formal discourses (e.g., science, mathematics) and the role gestures play during this development. In addition, the relative role of gestures to verbal utterances can be investigated in immersion (science and mathematics) classrooms. Here, students may be doubly handicapped for they are not only required to learn new specialized discourses, but they have to do this in a second language (Roth, 1994). Pilot work and evidence from a heretofore unaanalyzed database suggest that the presence of representations as indexical ground supports the development of the specialized discourse (here, science) by allowing communication despite the preponderance of verbal deixis (e.g., this, that, goes like this).

References


