Journal of Cognitive Systems Research

From Action to Discourse: The Bridging Function of Gestures

Wolff-Michael Roth, University of Victoria·

ABSTRACT: Gestures are pervasive in human communication across cultures; they clearly constitute an embodied aspect of cognition. In this study, evidence is provided for the contention that gestures are not only a co-expression of meaning in a different modality but also constitute an important stepping stone in the evolution of discourse. Data are provided from a Grade 10 physics course where students learned about electrostatics by doing investigations for which they constructed explanations. The data show that iconic gestures (i.e., symbolic hand movements) arise from the manipulation of objects (ergotic hand movements) and sensing activity (epistemic hand movements). Gestures not only precede but also support the emergence of scientific language. School science classes turn out to be ideal laboratories for studying the evolution of domain ontologies and (scientific) language. Micro-analytic studies of gesture-speech relationships and their emergence can therefore serve as positive constraints and test beds for synthetic models of language emergence.

KEYWORDS: Gestures, evolution of communication and language, multi-modal communication, pragmatics, embodied cognition

1. Introduction

Gestures are a pervasive feature of human communication in all cultures even those where gesticulation is frowned upon (Cassell, 1998). The importance of gestures in present-day human communication may have arisen from its role in the historical emergence of language. In the evolution of the human species, communication

· 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

first emerged in the form of symbolic movements (gestures) that were later supplanted by speech; that is, speech piggy-backed on existing neural configurations initially responsible for manipulation of objects (ergotic movements) and gestures (Allott, 1988; Corballis, 1999; Mattingly & Studdert-Kennedy, 1991). Such theories appear especially plausible in those cases where it can be shown that a link exists leading from the manipulation of objects to iconic gestures, and a changeover from iconic to symbolic modalities of contributing to communication. In such a case, the emergence of language and discourse is brought about and tied to embodied experience. One might ask, "Do similar relationships between activity, gestures, and speech exist in the development of individual human beings?" Deictic gestures (pointing), too, may be an important aspect of embodied communication based on the observation that deixis and deictic representation are core aspects of cognition (Agre, 1997; Ballard, Hayhoe, Pook, & Rao, 1997).

To find out whether manual activity, gestures, and discourse are related, and therefore constitute an important aspect of embodied, situated, and distributed cognition (Clancey, 1997; Clark, 1999: Suchman, 1987), communication has to be studied in sufficiently complex environments that also exhibit developmental aspects of discourse. School science (and perhaps mathematics) classes where students engage in hands-on activities constitute an interesting everyday setting. Here, the evolving relationship between manipulation of objects, gestures, and language can be easily studied as an aspect of "cognition in the wild" (Hutchins, 1995). Whereas children develop sensori-motorschemata while acting in and handling objects of the world (e.g., Piaget, 1970), I am not aware of similar research that shows how the learning of abstract-level concepts (which, according to Piaget, are "formal operational") is supported by such activities. Thus, "hands-on" activities are based on the premise that abstractand symbolicforms of representation (e.g., scientific discourse) is supported by handling materials; however, an existence proof for the supportive function of concrete activity in the development of formal language has yet to be made. In this article, I report and discuss data suggestive of the possibility that a developmental trajectory exists, beginning with primary manipulations and sensing of materials (ergotic and epistemic movements. These movements evolve into iconic gestures (symbolic movements) and from there support the emergence of discourse that describes the primary activities and observations; on these primary levels are lodged and evolve forms of talking about abstract entities.

Earlier studies suggested that students, who engage in school-science "hands-on" activities that force them to produce explanation of their observation, progress through a developmental continuum (Roth, 1999, 2000, 2001b). As students become familiar with the domain, their representations move from primarily external and gestural modalities to primarily spoken and written forms of representation. In these earlier studies, however, students investigated devices where concepts were perceptually or phenomenally accessible and could be explained in terms of "basic-level categorizations" (Lakoff, 1987) such as force, velocity, or acceleration. Symbolic communication involving basic-level categories appears to arise from communicative forms in which objects in the world and manipulations support the rise of (deictic and iconic) gestures, which themselves support the emergence of scientific language. In this sense, one can therefore speak of the embodied dimensions of scientific discourse. But do gestures have any function in the development of scientific discourse? This role of gesture is especially unclear when the core issues are microscopic events, or better, when macroscopic phenomena are modeled by means of discourse about microscopic and atomic events inaccessible to human experience.

The purpose of this in-situ naturalistic study was to see if there is any evidence for the evolution of scientific language after student conducted hands-on science activities and of the role of gestures in this evolution. This study thereby contributes to the ongoing debate about embodied aspects of human cognition, that is, whether there is a continuity between situation, gesture, and verbal modality of human communication (e.g., Blake, Olshansky, Vitale, & Macdonald, 1997; Locke, 1997). This study is concerned with understanding the relationship between cognition and environment, especially when subjects are actively involved, have a central role in determining what happens next and what they need to communicate about to complete their tasks.

2. Actions, Gestures, and Cognition

Human evolution has a long history preceding the appearance of language and the concomitant phenomenon of culture. Not surprisingly, there are theories in which cognition, the human body, and its sensori-motor achievements—which are the results of selection processes during the (successful) evolutionary trajectory—play a central part (Clark, 1999; Merleau-Ponty, 1945). Among the different ways in which the body may be involved in communicative acts, my work has focused on gestures. In the following, I briefly provide an overview of some of the more influential work on the gestural aspects of communication.

2.1. Studies in Cognitive and Learning Sciences

2.1.1. Interaction of Gestures, Speech, and Cognition

Gestures and speech have evolved into and constitute highly integrated communicative systems not only among aborigines (e.g., Levinson, 1997) but also members of industrial societies (Iverson & Goldin-Meadow, 1998). Talk and gesture are almost always copresent during communicative action (e.g., McNeill, 1992) across all cultures (Cassell, 1998). Although early research considered talk and gesture as separate components of communication (e.g., Kendon, 1983), more recent work constitutes these features as aspects of the same communicative event (Goodwin, 1986) arising from a common semantic representation (McNeill, 1992). This semantic representation is largely of topological nature. It is therefore more easily expressed in gestural than in verbal form, because the latter requires re-coding into typological (digital) form (McNeill, 1992; Roth & Lawless, in press-b).

There is ample cross-cultural evidence that words and gestures are deep features of human cognition. Linguistic anthropologists who work with people as widely distributed as the Hai||om bushpeople in Southwest Africa, Guugu Yimithirr speakers in Australia, and Mayan peasants reported on the interdependence of verbal and gestural deixis, iconic gestures, and spatial orientation (Haviland, 1993, 2000; Levinson, 1997; Widlok, 1997). In their cultural evolution, these traditional peoples developed orientation skills in which gestures, speech, and other aspects of cognition are deeply integrated systems. 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 directional terms (Widlok, 1997). This irreducible interdependence provides strong support for the claim that gestures and language are not simply aspects of linguistic systems but of a broader communicative one (Levinson, 1997). That is, these peoples have undergone cognitive and cultural evolution in response to the need to survive in the particular (harsh) conditions of their surroundings. (On the irreducible relation between cultural [societal] and cognitive evolution see Holzkamp, 1983.)

In Western (like) civilizations, schooling has supplanted traditional forms of learning. We can consider schools as places where students are asked to evolve new forms of discourse, to learn new linguistic systems. When students learn in science laboratories such as in the present study—where they are explicitly encouraged to develop observational and theoretical discourses based on their experience—a context is set in which new forms of communicating arise by means of evolutionary processes.

2.1.2. Evolutionary Nature of Discourse Evolution

The evolution of discursive forms in science classrooms has almost exclusively focused on student-centered learning environments. Student-centered generally means that students are supported in developing their own expressive means; these may arise from negotiations both at the small-group and whole-class levels (Duit, Roth, Komorek, & Wilbers, 2001). This allows multiple discourses to emerge. Students themselves learn to recognize that only some of the discourses are viable in the long term, which encourages them to abandon the less viable ones. Rarely ever, however, are all scientifically (and mathematically) incorrect discourses abandoned.

Early studies of language development in student-centered science classrooms, showed that even for the same objects (e.g., an arrow representing the concept “velocity”), 10 to 20 different words were used by students before they negotiated and settled on what appeared to be viable ones (Roth, 1996a). The particulars of the environment, negotiation with peers and interactions with teacher and textbook, constrained the proliferation of different discursive forms. First, to get on with their activities and because of the collective nature of the products, students were forced to negotiate and find the most viable language. Second, interacting with the teacher and textbooks provided another constraint that selected among discursive forms. There was therefore a double convergence, one acting within student groups the other at the class level.

Some studies focused on conversations as multimodal events in which gestures and the indexical ground against which they occur have communicative functions and constrain the evolution of viable linguistic forms in student-centered science classrooms (Roth & Lawless, in press-b). These studies show that the emergence of communicative competence is closely tied to the physical arrangement between the participants and relative to the artifacts over and about which the communication takes place (Roth, McGinn, Woszczyna, & Boutonné, 1999). That is, when students can usefully engage in manipulating materials and employ gestures, they quickly develop a viable discourse; their communicative competence remains underdeveloped when the learning environment does not afford the use of gestures in communicative acts (Roth, 1996b).

In summary, sociological, anthropological, and psychological evidence suggests that gestures are central to communication, support the emergence of language, and communicate ideas much earlier than spoken language does. Especially in the kind of classrooms I researched, the emergence of new discursive forms is best described by means of evolutionary analogies, for there is first an enormous wealth of new and different language productions among which only viable ones remain, constrained and selected through interactions with other students and teachers.

2.2. Evolution of Brain and Communication

A key stone in any argument that attempts to link activities in the world with communicative competence has to be some sensori-motor theory of cognition. Sensori-motor theories of cognition hold that representations of sensory and motor experiences constitute the basis for all levels of cognition (including communication); no matter how concrete or abstract a cognitive structure, it is said to make use of the same neural configuration that has been initially formed (e.g., Newton, 1996). Thus, "language as a function is not purely emergent but evolved in relation to central programs for motor activity . . . the protoman made utterances that were coincident with and driven by the same rhythm as the movement in question" (Kinsbourne, 1978, p. 553).

Motor theories of language suppose that the phonological, lexical, and syntactic structures of language have their origin in the structures of motor control, particularly the control of bodily activity. Sensori-motor brain circuits evolved together such that the perception of movements involves the same circuits as those that are responsible for enacting these movements (Decety & Grèzes, 1999). In human evolution, these neural networks evolved to meet the specific needs of an organism that had to feed itself through hunting and gathering. In the process, and particularly with the freeing of the hand through upright locomotion, the hand became free for communicative purposes (Corballis, 1999). The stage was set for language to appear. First, there were proto-languages in the form of gestures. During evolution, the brain made use of existing circuitry in new ways sort of building new cars from old parts (Bates, 1999). Both syntactic and semantic properties residing in gestural communication and already-existing motor capacities were therefore used for the production of speech. Finally, gestures that have had their origins in the iconic relations with the experienced world have evolved more abstract properties and thus into a symbolic form.

At some point, arbitrary sounds (and later iconic inscriptions) may have been linked to gestures and thereby constituted the evolutionary path to language (Corballis, 1999). That is, neural circuitry responsible for sensori-motor activity thereby became a vehicle for signs by means of a process of semiotic extension (Roth & Lawless, in press-b). "The sensori-motor theory takes an organic view of linguistic structures, and explains them as aspects of embodied activity in physical and social contexts. Abstract symbol structures can be understood and used because they are recursively isomorphic with the sensori-motor structures with which humans are primordially familiar" (Newton, 1996, p. 23). Gestures then play the role of intermediaries in the following way. Gestures are spatial events. Iconic gestures, because they can capture topological features and relations of things and events, are more closely related to human sensori-motor experience than words that have an arbitrary relation between form and content (meaning).

Three forms of evidence from human development support the central role gestures play in the development of language. First, children express through gestures a lot more understanding than they are able through speech; that is, gesture as expressive form appears ontogenetically earlier than spoken expression (Goldin-Meadow, Wein, & Chang, 1992). Second, when people learn to express themselves about new topics, that is, when their knowledge is in transition, there is a mismatch between gesture and language; gestures communicate correct understanding before speech (Church & Goldin-Meadows, 1986). Finally, there is evidence that cortical development is an evolutionary process that continues throughout the life span of individual human beings; processes that were traditionally thought of as learning in fact change the very structure (connectivity) of the brain (Quartz, 1999).

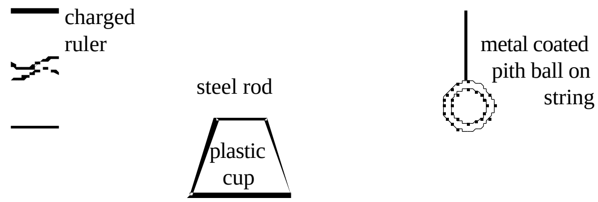
3. Method

This is an in-situ naturalistic study of the evolution of communication in a physics classroom where students did investigations that they later attempted to explain. In the course of their repeated attempts, the relationship between handling materials, enacting gestures, and using speech changed. This required careful tracking of the temporal and contextual features of the recorded conversations.

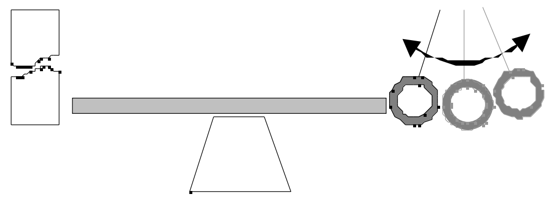
3.1. Research Context and Participants

This study was conducted in an existing database (Welzel, 1995) consisting of 30 videotaped lessons in a German tenth-grade

a. The system

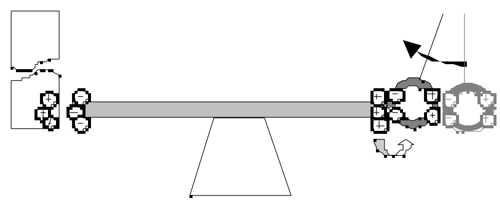


b. The Phenomenon



When a charged ruler is brought close to the left end of the steel rod, th
pith ball moves toward the rod, touches it, and then bounces back and forth even if the ruler has been taken away.

c. The Physicist's Explanation



Physicists explain the phenomenon by means of "influence" and "charg transfer." First, the positive charges on the ruler create a charge separation in the rod, which in turn creates a charge separation in the p ball. Therefore, the negative charges on the pith ball are attracted to the positive charges on the rod. When the pith ball touches the rod, positive charges are transferred to the pith ball which is now effectively charged and can cause charge separation in the rod and subsequent charge transfer continuing the process until the charges on both have equilibrated.

Figure 1. The steel rod and pith ball experiment. a. The system and its parts. b. The phenomenon as seen by the students. c. Physicists explain the phenomenon by means of "influence" and "charge." Students are expected to learn the physicists' language about the phenomenon.

physics class. In this class, students learned about static electricity by engaging in hands-on investigations for which they constructed explanations that they negotiated with their peers. Every so often, students also explained their findings to the entire class. One of these investigations was concerned with the phenomena of electrostatic influence in which a metal coated pith ball is made to bounce simply by approaching a charged plastic ruler near a steel road (see Figure 1). Physicists explain this phenomenon in terms of a temporary charge separation in the steel rod—achieved through electrostatic influence—which itself produces a temporary charge separation in the pith ball. As Figure 1.c shows, this causes the pith ball to be attracted to the steel rod and to become charged. This charge is, after the ruler has been removed, and with the successive bouncing off the steel rod, given back to the rod. The purpose of the activity is for students to gain familiarity with the phenomenon and then to coordinate words and related observations, partially through making inferences from known properties of materials and known physics discourse and partially with the help of teachers.

In this context, two student groups were videotaped. The groups constituted themselves according to student preferences, but were randomly selected for participation in the study. Based on their understanding, the selected students were representative of the population ($t(17) = 1.78$, $p > .05$). One group consisted of 4 male students, the other of 4 female students. Statistical tests showed no gender-related differences in achievement during the unit ($t(17) = 1.20$, $p > .05$); similarly, qualitative interviews showed no difference in conceptual understanding before and after the unit. Finally, there were no differences in the extent to which male and female students used gestures.

3.2. Data Construction

The unit on static electricity was videotaped using two fixed cameras that focused on one group each yielding a total of 30 hours of videotaped materials. Two microphones recorded the conversations at each of the two tables where the students were seated. Videotapes were transcribed such as to contain descriptions of the actions in which students engaged. From the original video, all instances in which students (or teachers) used gestures ($n \approx 100$)

were digitized ranging in length from about 20–120 seconds. The sound track was enhanced to improve audibility and ease of transcription and to make them available to frame-by-frame analysis (using Ulead® MediaStudio Pro 5.2 and QuickTime 3.0).

The computer program Ulead® MediaStudio Pro 5.2 Audio Editor allows the visual representation of volume so that the onset of each speech act was easily identified as it normally corresponds to a rapid vertical increase to a peak that stands above the noise. The height and slope of the curve facilitate the identification of the pitch accents related to the "rhematic elements of utterances" (novel features) that are the reference points for timing the relative occurrence of the associated gestures (e.g., Cassell & Prevost, 1996). Individual utterances were timed to 0.001 s (1 ms), and coordinated with the frames which allowed timing to 0.04 s (40 ms associated with a 25 frames / s rate of the video).

Frames with characteristic aspects of the gestures (e.g., extreme positions where stroke ended) were printed and the salient feature of body position and gesture traced onto transparency film from which the images were scanned. These images were then coordinated with the speech features. The agreement between the time measurement based on the video signal (frame by frame) and that by the sound track fell within the accuracy of the video signal (40 milliseconds).

4. Evolution of Scientific Language in School Science

Students' activities were related to a number of distinctly different investigations (charging materials, testing charges, electroscope, influence and grounding, etc.). During all of these, students were asked to begin with an investigation and to subsequently construct verbal descriptions and explanations. From these different investigations and across the two groups, several patterns emerged that extended the results of previous investigations. In this section, the following points are elaborated:

1. There was an increasing abstraction of the communicative action from the specifics of the setting. Initially, the original materials and equipment

played key roles in the communicative action whereas later, these were more or less replaced by gestures or completely represented in speech.

2. Materials and gestures that were used in communication allowed a level of representational complexity that students were not yet able to reach in speech alone. That is, the presence of the entities and phenomena talked about supported the emergence of gestures, and thereby of subsequent verbal expressions. Over time, communication shifted into the verbal modality thereby becoming independent of direct visual access to the objects.
3. There were several important time-related phenomena: initial temporal lags of speech behind gestures progressively diminished, overall presentation time decreased.

4.1. Evolution of Hand Movements

In this section, I show how students began their first explanations by reenacting parts (or all) of an investigation; this re-enactment served as the indexical ground of their utterances.[1] Later, the materials and equipment still functioned as indexical ground but gestures began to replace actual objects and events. Subsequently, students frequently employed a different object or gesture by itself to represent some relevant aspect of the event they talked about, and finally they represented all relevant aspects of objects and events in symbolic (abstract) form. I distinguish three types of functions that gestures have important in this development: In addition to the evident symbolic function obtaining to gestures during speech acts, gestures have epistemic and ergotic functions (Cadoz, 1994; Weissberg, 1999). As to the epistemic function, the hands (as well as other body parts) permit the human agent to perceive qualitative aspects such as the temperature, form, texture, or movement of objects. The ergotic function relates to the fact that humans change their environment, turn, displace, compress, or pull objects. Epistemic and ergotic movements frequently go together taking on a representative function that increases cog-

[1] The episodes presented here constitute but a small selection of examples that by themselves would not allow the inferences made here. Other episodes supporting the same claims are available elsewhere (e.g., Roth, 1996a, 1996b, 1999, 2000, 2001; Roth & Lawless, in press-a, in press-b; Roth & Welzel, 2001).

nitive capacities by involving gesture and environment in thinking (Kirsh, 1995).

4.1.1. From Ergotic to Epistemic Movements

When asked to provide explanations for phenomena at hand, students initially relied almost exclusively on redoing the investigations as part of their attempts to provide a description of the phenomenon and as a context for constructing atomic-level explanations. Because the associated events were too fast, students simulated the events by moving the objects through the different stages of the phenomenon. This allowed them to describe the observed objects and unfolding (simulated) events in real time allowing for a copresence of expressive means and aspects of the world.

4.1.1.1. Episode

Phil, Matt (pseudonyms are used throughout) and their two peers not featured here have repeatedly done the investigation where they bring a charged object close to (or make it touch) the end of a metal rod opposite to the covered pith ball (Figure 1). As a result, the pith ball bounced wildly. The four students then attempted to construct their written report of what they have done and observed and their explanation. In this first of six recorded episodes, Phil and Matt collectively attempted a description and explanation of the phenomena.

Rather than just talking about his previous investigation, Phil actually ran it again (Figure 2.a i, ii). He discharged the rod and brought the pith ball to its resting position before charging a plastic ruler and bringing it to the steel rod. He attempted to graft a conceptual explanation on top of this observational talk. He used his right hand to point (with ruler) to the part of the steel rod where there should be a surplus of electrons (2.a.iii). His left hand first pointed to the end of the rod where “subfluous” electrons would move so that “it” (the coated pith ball) would be attracted (2.a.iv). He represented this attraction by a swiping movement of the hand parallel to the pith ball rod axis where this attraction is to be expected. In the second part of this episode (Figure 2.b), Matt reiterated what might be happening at the conceptual level. His gestures represented the movements of the electrons (literally

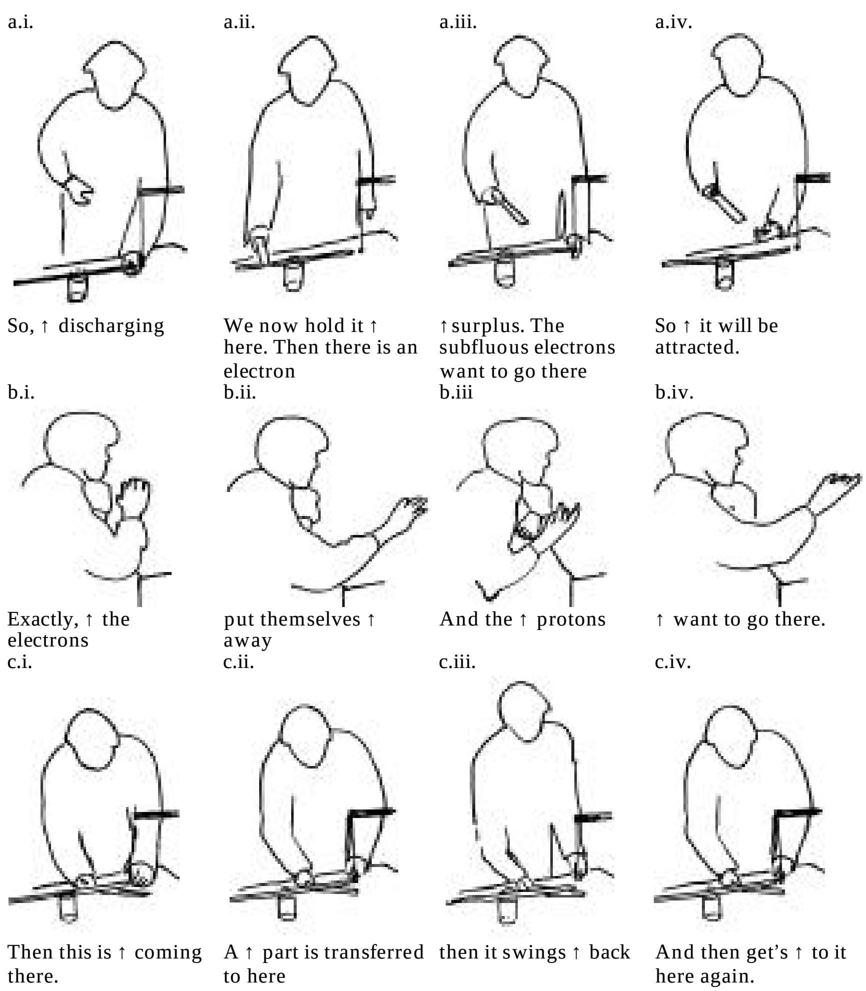


Figure 2. Excerpt from a conversation between two students, Phil (a, c) and Matt (b). Phil (a) constructs an initial explanation as they watch the pith ball bounce. Matt (b) and Phil (c) resort to gestures over the equipment but without actually charging the different bodies. The arrow indicates the timing of the gesture with respect to the speech. (The "up" arrow indicates the point where utterance and image coincide.)

"away"[Fig. 2.b.i,ii] from the pith ball at the opposite end of the table) and that of the protons towards the pith ball (Fig. 2.b.iii). In the final part (Figure 2.c), Phil took the pith ball in his left hand and, paralleling speech and gesture, moved it. Simultaneously, he described observational and conceptual aspects of the event. Accordingly, the pith ball approached the rod (2.c.i, ii), transferred "a part" then swung back (2.c.iii), and returned to the steel rod (2.c.iv).

4.1.1.2. Analysis

There are several typical dimensions for the early stages of communicative competence (cognitive complexity and temporal aspects are discussed in subsequent sections). First, students use equipment and materials, which they describe in observational terms. Second, their conceptual talk is often scientifically inappropriate. Third, in the early stages, students often spoke from the point of view of the inanimate entities involved and thereby portray these entities as animate. Fourth, their communication relied heavily on verbal and gestural deixis.

First, Phil reenacted the investigation while providing a phenomenal description of his actions and observations ("discharging," "hold it here," "it will be attracted"). As the event unfolded, he began a first attempt to layer a theoretical description on top of what was unfolding before the students eyes. His movements were ergotic,causing changes in the world that can be observed. However, these events were often too fast to be simultaneously described. Students began to recreate the events, which allowed them to coordinate speech with the perceived events. Here, Phil's gestures became epistemicas he took the pith ball into his fingers and moved it through the trajectory (Figure 2.c)—a form of cognition distributed across the setting. While guiding the pith ball through its trajectory, he could stop the motion (or do slow motion) so that his talk about coming, transferring, swinging back, and returning was in relative synchrony with the actual position of the pith ball.

Second, given that students were to learn physics and had to begin with the language currently available to them, it was not surprising that they used words inappropriately and described inappropriate physical events—as seen from the current stand of science. In this case, Phil made up a new word "subfluous."[2] Matt

[2] In the original German transcript, Phil used the word "unterschüssigen" which does not exist but is a neologism that builds on the contrast with "überschüssig," superfluous (adjective) in which the same verb root is paired with the contrasting preposition.

described protons as moving in the metal rod which, from the physicists' perspective is impossible for the nuclei, where the protons are located, are fixed in the lattice of a solid. Although these ways of speaking were not appropriate, they can be seen as first attempts in creating a new discourse. Whether these new forms actually survive cannot be determined at that point in the events.

Third, in the early stages of learning about new phenomena and theoretical entities, students' speech and gestures frequently were from the object point of view or portrayed them as animate. In this excerpt, Phil talked about electrons that "want to go" some place and Matt suggested that the electrons "put themselves away" and protons "want to go" some place.

Fourth, in the early stages of students' examination of physical phenomena and their explanations, there was a high degree of verbal and gestural deixis. For example, Phil not only used the deictic terms "here," "there," "this," and "it" but the referents of these terms shifted even in the brief episode displayed here. "It" referred to the ruler (Figure 2.a.ii), pith ball (Figures 2.a.iv, 2.c.iii), and steel rod (Figure 2.c.iv). From Phil's perspective, "there" both referred to the right end of the steel rod where there was an electron surplus (Figure 3.a.ii), the left end where there was an electron deficit (Figure 2.a.iii), and the right extreme of the pith ball's trajectory. The same referent also was designated with different indexical terms: when the pith ball approached the end of the steel rod it was both "here" (Figure 2.c.iv) and "there" (Figure 2.c.i). Despite these variations and apparent inconsistencies at the verbal level, there was no problem apparent in the students' communication. With the materials and equipment actually in front of them, accessible to all students in the group, the respective listeners disambiguated the different indexical terms. Furthermore, (I return to this below), the use of verbal and gestural deixis allowed students to go on rather than engaging in extensive efforts of immediately producing full descriptions and explanations in the verbal modality.

4.1.2. From Epistemic to Symbolic Movements

In a second stage, students used some of the materials from their investigations as ground against which they layered their theoretical descriptions. We can conceptualize the deictic and iconic

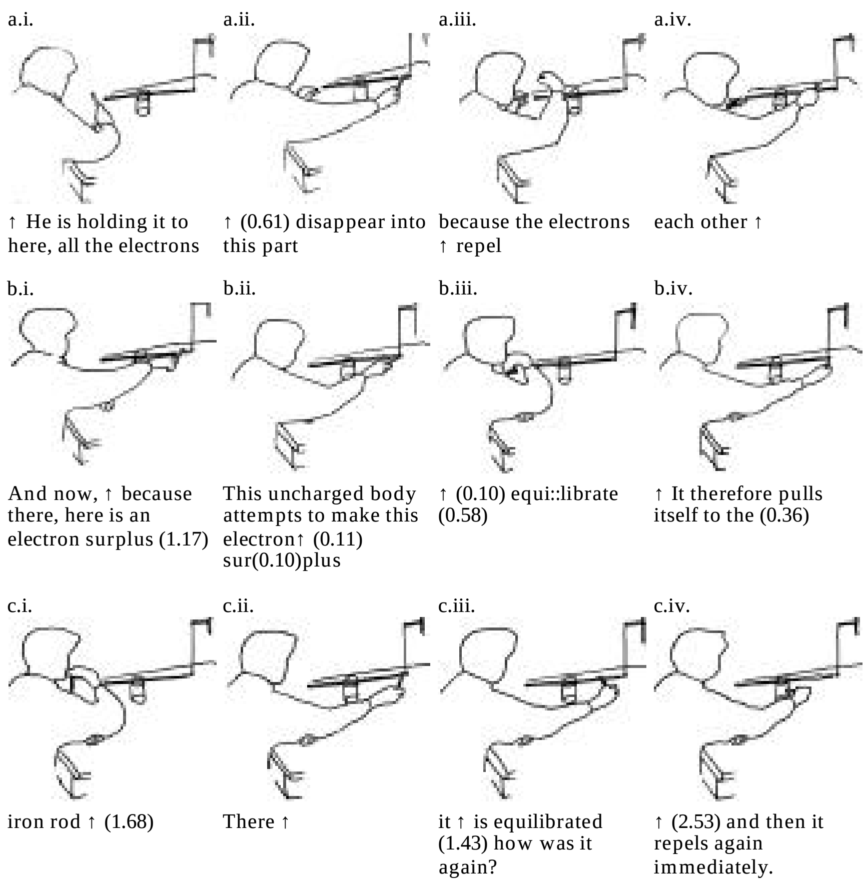


Figure 3. Matt provides an explanation by using the equipment as indexical ground. This ground and the phenomenon that it is associated with provides the phenomenal layer onto which a new theoretical discourse can be constructed (grafted). At this stage, his iconic gestures run ahead of the corresponding (stressed) lexical item. The arrow indicates the timing of the gesture with respect to the speech. (Colon[s] in or at the end of a word indicate that the preceding phoneme is drawn out by about 0.1 second per colon. Numbers [1.68] indicate pauses in seconds.)

gestures as distributing cognition around the setting. Whereas students conducted the investigation or moved parts of the equipment literally around as part of describing what happens and constructing a theoretical description in the previous event, students now simply used talk and gestures in one communicative act.

4.1.2.1. Episode

In this episode, Matt constructs another phenomenal and theoretical description of the pith ball and iron rod investigation. In the first line of Figure 3, he describes how holding the charged ruler to the end of the steel rod makes electrons move to the opposite end because of a repelling effect. His deictic gesture indicates the place in the rod where the electrons will go (Figure 3.a.ii) whereas his iconic gesture embodies the notion of "repelling."

In the second line, Matt described and explained the effect of those electrons that are supposed to be at the (from his position) opposite end of the rod. He suggested that the pith ball ("this uncharged body") made an attempt to cancel the charge surplus and therefore pulled itself to the rod (Figure 3.b.iv-3.c.i). His deictic gestures that accompanied verbal deixis ("this," "it") made the pith ball salient (Figure 3.b.ii, 3.b.iv). Iconic gestures enacted the notions of "equilibrate" and "pull to" as the hand moved from the stretched out position at the end of the rod all the way in front of him. Toward the end (Figure 3.c.iii-3.c.iv), Matt began to hesitate about how to complete the description. His theoretical ("equilibrates" Figure 3.c.iii) and observational descriptions ("repels" Figure 3.c.iv) appear to stand in the air without relating to each other. At this point, he had not yet (as it happened later in the lesson) arrived at a conceptually complete explanation for the process in which observation and theoretical discourses are coordinated.

4.1.2.2. Analysis

In the evolution of students' language, this example is further along than the previous episode where students actually observed the event or moved the objects around to be able to describe them and their relations to each other in the various configurations. Here, what is being done as part of the investigation and what can be observed was described or enacted as gesture. Against the objects present and the description of actions and observations, Matt presented a theoretical description. Materials and equipment served as ground and in some cases were replaced by arbitrary objects. Here, a pencil (Figure 3.a.i) replaced the original ruler and subsequently served as a pointer. The presence of these materials or their substitutes allowed students to point to particular aspects without the need to generate verbal signifiers. Matt did not name the iron rod other than in Figure 3.c.i, although he repeatedly referred to it. Also, he did not name the pith ball; he brought the object in to communication by means of verbal ("this" [Figure 3.b.ii], "it" [Figures 3.b.iv, 3.c.iv]) and gestural deixis.

As in the previous episode, the same indexical items have different referents; because of the logic of the underlying event to be described and explained, these indexicals are therefore disambiguated. Thus, "it" refers to the charged ruler (Figure 3.a.i), pith ball (Figure 3.b.iv, 3.c.iv), iron rod (Figure 3.c.iii), and the explanation itself ("how was it again" Figure 3.c.iii). In the same way, both ends of the steel rod are designated by the deictic term "here" and the opposite end of the rod is also "there" (Figure 3.b.i, 3.c.ii). At this stage, the visible objects and invisible entities are animate and engage in intentional action in the explanation. Thus, the pith ball "attempts to make ... equilibrate," "pulls itself" or "repels ... immediately." Furthermore, the rod "equilibrates" and the electrons "repel each other."

Toward the end of this episode, Matt had become more independent of the indexical ground constituted by the materials, though he still made use of indexical words and gestures to designate the things at hand. He did not need to fully describe those things visible to all. They "went without saying" and it sufficed to designate the relevant objects, entities, and events by pointing to them or expressing them by means of an iconic gesture. Against the material ground, the theoretical discourse took its hold. In these first episodes, speakers took the point of view of the entities involved. In the evolution of communicative forms, it may be easier to talk about the objects and entities as if they were imbued with agency and to move to a dispassionate and non-animate perspective at a stage when very familiar.

4.1.3. From Symbolic Movements to Speech and Text

When students became very familiar with the objects, equipment, and phenomena produced with them, they no longer required the presence of the materials. At this point, arbitrary objects served as signs that stood for some object or entity. For example, in an explanation of how static electricity is produced, a student used two pens to stand in for the two materials to be rubbed. A third, fountain pen stood for an atom which, as part of the rubbing was separated into a positively charged nucleus and (negatively charged) electrons. Her iconic gestures signified the separation as the student pulled the cap of the pen, verbal and gestural deixis associating each of the entities with the macroscopic materials corresponding to the microscopic descriptions. In a similar way, toward the end of the second lesson on the steel rod-pith ball investigation, Phil produced an explanation in which he only used an arbitrary PVC rod to stand for the steel rod.

4.1.3.1. Episode

Phil produced observational and theoretical descriptions against the arbitrary PVC rod, which acted as a placeholder for the materials and equipment as a whole. His gestures and speech represented both the phenomenal events and the underlying entities and their relations (Figure 4). Phil showed with his right hand where the (charged) ruler was held relative to the rod and, with his left hand, how "everything" was repelled inside the rod toward the other end (Figure 4.a.i, ii). As a consequence, the pith ball moved to the end of the rod and then moved away again. At the same time, there were events that his gestures re-enacted over the rod and that therefore were attributed to the steel rod in the original investigation. Here, Phil's right hand enacted movements of "something" (or "they") in the rod (Figure 4.b.i-4.b.iii, 4.c.iv-4.d.iv) with which he sought to explain the intimated movement of the pith ball done twice with the left hand, after it was represented initially with the right hand (Figure 4.a.iii-4.b.i).

4.1.3.2. Analysis

In this episode, the PVC rod provided a material representation of the steel rod that had been part of the original investigation. The PVC rod represented the steel rod and therefore functioned as (material) sign for some other object. Resembling the steel rod in shape and size, it had an iconic relation to the object it referred to. In this, the right end corresponded to that end where the ruler was approached, the left end to that where the pith ball was found earlier. The atomic level events were gestured to occur inside the rod. A crucial and also the most difficult part of the explanation was the articulation of what happened at the atomic level after

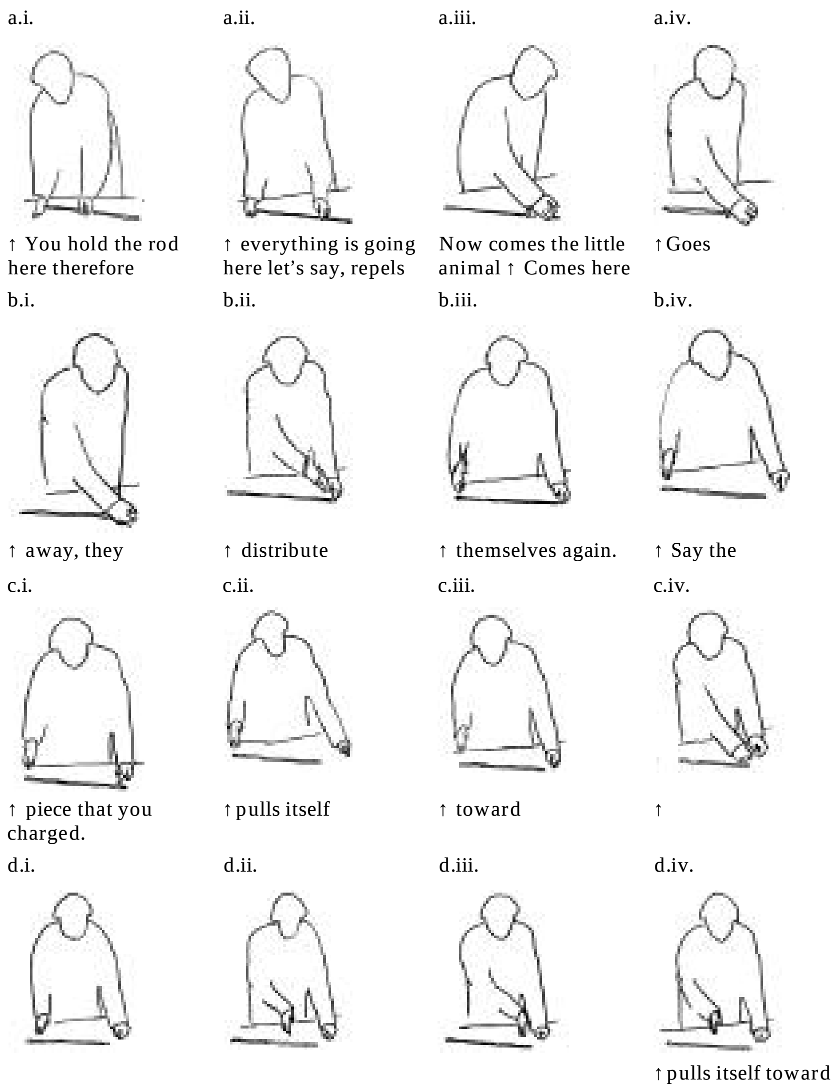


Figure 4. Phil and his group attempt to construct their explanation for induction on the metal rod and the pith ball so that they can describe it in their notebook. He picks up some plastic rod on his desk and then enacts his explanation using gestures over and about the rod.

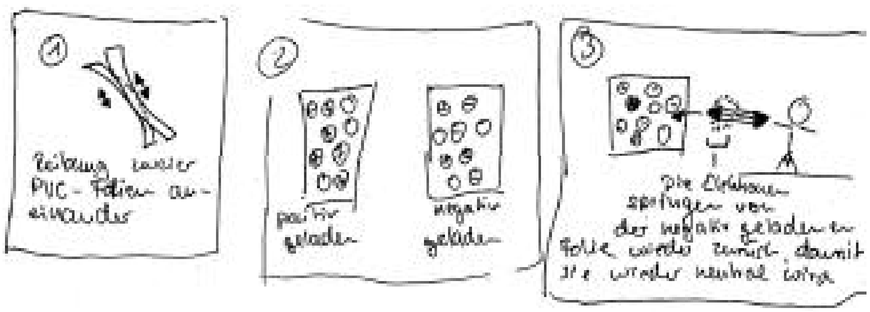
the pith ball was initially attracted and repelled. Here, Phil did not yet have a verbal explanation, but his gesture enacted a process according to which there was a cyclic process of movements (perhaps equilibration) across the rod (Figure 4.c.iv–4.d.iii). Changing to representing the pith ball with the left hand while depicting the atomic level movements with the right hand also clarified the referents. As in the previous episodes, he imbued the objects and entities with agency.

The entities used as part of the gestures took the role of additional representations. They tracked objects and entities and therefore did not need to be represented mentally. They were perceptually available to all participants in the conversation. These objects existed as signs available for future reference and as long as they were needed. They served as indexical ground to gestures without the need of a description. In this way, communication was spread across the setting, reducing the need for internal representation. Each time students changed the representational form, that is, in the way represented in the three episodes, the production of the explanation became more difficult again. Students had to replace previously existing physical representations by discursive or gestural representations.

4.1.4. Discussion

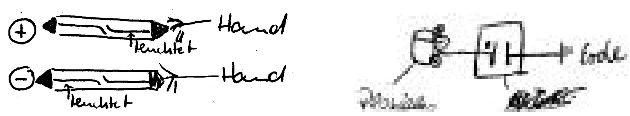
The sequence of episodes illustrated how students became increasingly independent of the actual objects and equipment in the production of their observational and theoretical descriptions. There was an abstraction from the actual phenomenon to the simulated phenomenon with the actual objects, to a simulation where the objects themselves are replaced by other, arbitrary objects. For this investigation, all recorded episodes showed students using some form of material representation. However, there were no notebook entries that could be analyzed to show the continuation of the representative practices. Corresponding data from another student investigation is therefore used (Figure 5).

Figure 5 shows some of the student responses to the test, "How can you test whether the charge on a rubbed transparency film is negative." The results provide a cross section of students at various points in the trajectory of representational competence. The representations changed from more iconic (topological) to more textual (typological) representation forms. In Figure 5.a, the student represented in iconic form the events of rubbing two films a.

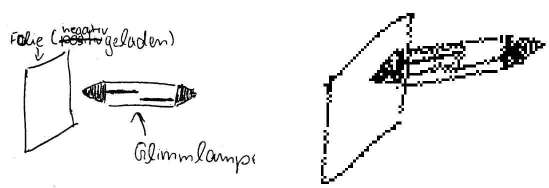


Text: The left hand side of the neon light bulb will glow.

b.



c.



Text: Touching causes glow on the film side.

Figure 5. a. Iconic representation of a producing and testing for charges. There is little other text to accompany the drawings. b. various aspects of the iconic representation are now replaced by words or oth er symbols for charges (left), lamp (right), person (“Erde” < Ger. ground). c. Even more of iconic representations are omitted and replaced by text (not shown).

and how he tests for the charges, recording his observation as a glowing neon bulb. Also represented is the conceptual level explanation in terms of positive and negative charges on the two films. In Figure 5.b, students represented the events of holding a neon lamp to the film by means of a drawing of hand, the word “hand,” the word “ground” (“Erde”) and the electrical symbol of ground. Students’ drawings became increasingly abstract finally omitting most of the detail (Figure 5.c) and disappeared eventually altogether leaving all explanations to the verbal texts (not shown here). In Figure 5.c, the hand and its connection to the ground is no longer salient; rather than object representations accompanied by "+" and "-" symbols as in the other representations, the charges are indicated in verbal form as "negatively charged" ("negativ geladen").

Gestures therefore played an important role in locating objects and entities by means of deixis, and animating the movements of these objects and entities by means of iconic gestures. However, the iconic gestures also had deictic function. If a student moved his hand from left to right across a rod while talking about the movements of electrons, he did not just gesture motion (as this might happen in those communicative contexts where the topical situation is not at hand). Rather, he also represented motion in a particular direction and along specific trajectories. Gestures were over the material objects that are their topic. Gestures thereby encoded information not necessarily available in contexts where these objects (or their representations) are absent.

Taken together, material objects, gestures, speech, drawings, and written text constituted different modalities for situationally representing ideas in communication. At present, schooling is such that it favors verbal (written) representation. This feature of schools therefore constitutes an evolutionary pressure that shifts students' representations from situated (material, gesture) into situation-independent (symbolic, abstract) modalities. In the present learning environment, students could initially draw on all representational modalities. This possibly decreased the cognitive load associated with speech production with subsequent drifts to verbal modes when students were more familiar with the domain and after complex situations had been "compiled" into more inclusive concepts.

4.2. Evolution of Temporal Dimensions of Gesture and Speech

There are important temporal dimensions in the relation of context, gestures, and speech as representational modalities in students' communication. First, students gesture phenomena that they cannot yet express at all in a verbal modality. Second, when students began to express phenomena verbally, the verbal mode lagged behind the gestures. As students became familiar with the phenomena on the one hand, and the production of descriptions on the other, the lag all but disappeared so that gestures and speech were perfectly coordinated. Third, and paralleling the previous development, the overall time for producing observational and theoretical descriptions decreased significantly.

First, throughout the above episodes, there were examples of situations in which students did not produce, or could produce only slowly, scientific correct verbal presentations. However, their gestures already foreboded those presentations in that they embodied the topological and indexical relations to the macroscopic objects. In one particular instance, Phil did not produce a verbal description at all, but his right hand moved across syncopating a temporal succession which paralleled the periodic motion of the pith ball intimated through his previous gesture (Figure 4.c.iv–4.d.iv). Here, the visually available periodic back and forth of the pith ball was represented in the syncopated back and forth movement over the PVC rod of the left, “electron” hand. This can be interpreted as an instance of situations in which learners gesture what they are only (much) later able to express in other modalities (e.g., Goldin-Meadow, Wein, & Chang, 1992).

Second, when students produce descriptions while moving the objects around to produce a slow motion representation of the actual event, talk is normally coordinated with the movement. (These observations are made among adults familiar with the topic they are communicating about.) However, during their attempts in constructing observational and theoretical descriptions of their investigations, there was a lag between the stroke of the gesture and the corresponding word. This lag was from 400 milliseconds to1.40 seconds. In several more extreme situations, delays of 2.01, 2.12, 2.22, and 2.53 seconds were measured. Thus, in Figure 4.a.ii, the gesture representing movement of electrons and deictic gesture indicating their final location was completed 0.80 s before the stressed part of its corresponding word of disappear (“verschwinden”). The end point corresponding to “this part” was reached 1.23 s before the associated utterance. Matt’s gestures ran ahead of his utterances, which appeared to be “facilitated” or “jogged” by the gestures that preceded them. In the present database, there was only one student (out of 8) whose gestures were nearly always perfectly coordinated (within the 40 ms accuracy of the video track) with speech, though there were often long pauses before the simultaneous production of both.

Third, the simultaneous decrease in use of materials and increase in more abstract representations (especially in the verbal modality), was accompanied by a decrease in the amount of time students took to produce an explanation. For example, from his earliest attempts in explaining the phenomenon to the above presentation, Matt reduced the time from 48 to 30 seconds over a 45-minute period. In the same time period, Phil went from 25 to 14 seconds for explaining a phenomenon, paralleled by a change from redoing the investigation to using but one iconic gesture. In a similar way, another student's first presentation of a proposal for an experiment lasted 18 s and involved several gestures including the expected outcomes of approaching a charged film to a water current from the tap. In her second presentation of the same experiment to a student previously absent, she explained the investigation in 11 seconds and the gestures were reduced to the iconic representation of the movement that had to be conducted with the film, and the direction of the water current.

5. Discussion

In this article, illustrative and representative episodes are presented deriving from a naturalistic study of gesture-speech relations as students developed explanations for physics investigations. The episodes illustrate how, with familiarity, there are shifts along several dimensions. First, rather than explaining phenomena, students manipulated the original equipment. They subsequently used the equipment as indexical ground and simulated events by means of gestures. Gestures subsequently substituted earlier presentations. Finally, students used verbal representations that later showed up in the form of written texts, which were accompanied by decreasingly iconic drawings. Second, associated with this shift there was a decrease in the total amount of gesturing over time. Gestures often preceded speech production both when words were totally unavailable and when they were produced but came too late. The delay between gesture and speech decreased as part of the development and disappeared when the two forms of representation overlapped. Taken together, the two types of development can be understood as an evolution whereby a number of embodied representational modalities were coordinated and increasingly shifted into a verbal modality. Thus, although there is some doubt about the continuity of communication as it shifts between different modalities (Locke, 1997), the conditions in the present study lead to observations that are in support of a continuity hypothesis.

5.1. Evolutionary Trends

In evolutionary terms, the individual-level changes normally associated with the notion of "development" reproduced gesture-speech relations that were hypothesized to characterize human (cultural) "evolution." Whether or not one agrees with the "ontogeny repeats phylogeny" assumption, the data from this and other studies support the notion of language development as an evolutionary process (at least in the type of settings researched).

In the shifting levels of representation from things and movements in the world, that is ergotic and epistemic action to symbolic gestures to spoken and written text, students follow evolutionary trends. Some cognitive theories regard environment-individual as a unit that integrates the two into an irreducible, dynamically and structurally coupled unit (e.g., Mandelblit & Zachar, 1998; van Gelder, 1998; Varela, Thompson, & Rosch, 1993). The present data illustrate cognitive events whose emphasis shifts from environment to increasing symbolic representation within the individual agent. Initially, communication relies on the ergotic and epistemic function of gestures accompanied by little speech often inappropriate to the situation. Those aspects that constitute representations outside the human agent were increasingly shifted to internal representations, spoken and written text and drawn images, that were of a different logical type than the gestures and things in the world (on this view of internalization see Leont'ev, 1978). More advanced forms of cognition are made possible to the students by communicating in contexts that afforded other modalities than speech alone. Thus, because material objects and instruments were present and visually available to all participants, they "went without saying" and therefore did not need to be talked about (unless there is conversational trouble). At the same time, the material setting put constraints on how gestures could be understood, because in communication, gestures and aspect of the ground to be made salient have to coincide. These constraints by the material setting therefore operated in ways similar to those that determined the next moves in the actions of short-order cooks and therefore constitute a form of knowing that does not need to be mentally represented (Agre & Horswill, 1997).

In this study, students produced many different descriptions, but only some were viable and survived. Others altogether disappeared, such as the notion of protons that move throughout the material. Still others, such as the "subfluous electrons" changed into "electron holes." The results of the present study are therefore consistent with those of other studies that focused on the emergence, stabilization, and extinction of discourse in science classrooms (e.g., Roth, 1996a; Roth & Duit, 1997). Whether a particular description survived depended on the number and quality of other descriptions of the same phenomenon, the degree to which it was taken up by others, the level of support it received from teachers. (Elsewhere I show how such distributed processes can be modeled using constraint satisfaction networks [Roth, 2001a].) That is, there were clearly identifiable situational factors that allowed at a minimum a descriptive model of the linguistic changes in a physics classroom.

For such evolution to occur, evolutionary "pressures" or "constraints" are required for change to occur (Steels, 1997). In the present context, the most obvious pressure is associated with students' meta-knowledge that at some later point, they have to produce written or verbal explanations. For example, laboratory notes and written tests required context- and gesture-independent descriptions and explanations of phenomena. This constrains students to evolve verbal forms of communication. There was therefore a pressure towards a specific form of communication built into the particular form of enacting curriculum that provided students initially with great freedom to evolve communicative forms, but ultimately constrained them to allow only one (or a small number of variations of verbal forms).

5.2. Cognitive Complexity and Modeling

Materials and gestures as representational forms allowed students to externalize thought. That is, these presentations can be considered as distributed and situated cognition. Using the world as its own representation and later to gesture over and above the objects visibly available to the conversational participants reduced the mental effort that would have been required by speech production and verbal representation alone. That is, parallel channels including natural language, gestures, visual representations, and natural objects allowed different types of communication integrating the typological (designating) and topological (dynamic) aspects of the phenomena and the theoretical discourse laid over them (cf. also Lemke, 1998, 1999). The multimodal representation was thereby made available for inspection in its dynamic parts in a much better way than if the participants were exposed to the symbolic, verbal product only. Finally, there were constraints on the objects and events when they were enacted through ergotic and epistemic movements, or appeared as symbolic gestures above the materials than by manipulating mental images.

An important issue in the modeling appears to be the representational point of view. In their (verbal and gestural) communication, students often took the perspective of the objects, which appear to be engaged in intentional actions. Sometimes researchers treat such representational forms as primitive and animist (diSessa, 1993). Such a characterization may miss the point. First, even the most abstract of scientists, physicists, are known to use object-centered talk and the researchers become part of the world of entities that their work is about (Ochs, Gonzales, & Jacoby, 1996). Second, even accomplished scientists take such perspectives: the Nobel Prize winning Barbara McClintock used alternative, animate language invites the perception of nature as an active partner in a more reciprocal relation to an observer, equally active, but neither omniscient nor omnipotent (Fox-Keller, 1983). I therefore suggest that the subject-centered perspective may provide cognitive advantages over abstract perspective because it decreases the mental effort necessary for producing a non-indexical representation. I-centered perspectives and indexical representation characterize everyday cognition (Agre, 1997)—perhaps becausethey require less cognitive effort.

5.3. Synthetic and In Situ Studies

There are interesting parallels between this study and those that take a synthetic approach to language origins (e.g., Steels, 1997). In this classroom, as in science classrooms more generally students evolved new domain ontologies. That is, what they saw as entities and how these were constituted by part-part and part-whole relations was a matter of evolutionary processes in which both viable and non-viable perspectives emerged (e.g., Roth, 1996a, 1996b; Roth, McRobbie, Lucas, & Boutonné, 1997). It is through the feedback from and negotiations with other students and from the input and constraints on language by the representatives of formal science (teachers, textbooks, scientific equipment) that the ontologies and language games were constrained. Thus, as in Steels' synthetic modeling of language origin, I observed students who had to negotiate, before they could make sense of the phenomena at hand, what the relevant objects and events were, that is, where the relevant joints are that their perceptual fields had to be cut. The second problem was then one of finding a suitable language. Here, although teachers and textbooks provided words that could be embedded in natural language structures to produce new forms of discourse, the mapping and coordination of this new course with the objects was not an easy task. There is substantial evidence from many studies on everyday "scientific misconceptions" that despite (sometimes extensive) instruction employ scientifically inappropriate language to many natural phenomena. Careful studies of language acquisitions in situations such as those I have described here should be run in parallel with synthetic modeling studies to provide evolutionary constraints in their respective theoretical development.

6. Conclusion

This naturalistic study provides support for the contention that in certain situation—hands-on activities in school science—gestures take a transitional role between in the acquisition of scientific language. That is, verbal communication arises from and piggybacks on physical engagement with the world. This is not a strong existence proof, for qualitative studies such as this are not suited to test causal relationships. However, in this database alone there is a large number of examples that substantiate the developmental trajectory from ergotic and epistemic movements to symbolic movements (gestures) and scientific language. This trajectory appears to foster students' acquisition of new discursive forms despite the fact that they have to learn both domain ontology and (observational and theoretical) language to describe it. The present study illustrates these trends even though the realm of entities talked about was no longer accessible to direct experience.

In school science laboratories, students are required to evolve new language and articulate previously unknown objects and events in a short period of time (relative to human evolution). School science and mathematics classrooms may therefore constitute natural and still little used laboratories for (experimental) linguistics and cognitive science concerned with the embodied and situated nature of cognitions.

Acknowledgment

This research was made possible in part by a Fellowship in the area of Cognitive Neurosciences and Information Sciences from the Hanse Institute of Advanced Studies (Delmenhorst, Germany) and Grant 410-96-0681 from the Social Sciences and Humanities Research Council of Canada. I thank the professors of the Institute for Physics Education at the University of Bremen for providing me with access to the database. They and their postdoctoral fellows and Ph.D. students also provided me with many opportunities for discussing my emergent analyses.

References

Agre, P. E. (1997). Computation and Human Experience, Cambridge, Cambridge University Press.

Agre, P., & Horswill, I. (1997). Lifeworld analysis. Journal of Artificial Intelligence Research, 6, 111–145.

Allott, R. (1988). The motor theory of language origin and function. In J. Wind, B. Chiarelli, B. Bichakjian, A. Nocentini, & A. Jonker (Eds.), Language Origin, A Multidisciplinary Approach, Dordrecht, Kluwer Academic Publishers, pp. 105–119.

Ballard, D. H, Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. Behavioral and Brain Sciences, 20, 723–767.

Bates, E. (1999). Plasticity, localization and language development. In S. Broman & J.M. Fletcher (Eds.), The Changing Nervous System, Neurobehavioral Consequences of Early Brain Disorders, New York, Oxford University Press, pp. 214–253.

Blake, J., Olshansky, E., Vitale, G., & Macdonald, S. (1997). Are communicative gestures the substrate of language. Evolution of Communication, 1, 261–282.

Cadoz, C. (1994). Les réalités virtuelles [Virtual Realities], Paris, Flammarion.

Clancey, W. J. (1997). Situated Cognition, On Human Knowledge and Computer Representation, Cambridge, Cambridge University Press.

Clark, A. (1999). Where brain, body, and world collide. Journal of Cognitive Systems Research, 1, 5–17.

Cassell, J. (1998). A framework for gesture generation and interpretation. In R. Cipolla & A. Pentland (Eds.), Computer Vision in Human-Machine Interaction, Cambridge, Cambridge University Press, pp. 191–215.

Cassell, J., & Prevost, S. (1996). Distribution of semantic features across speech and gesture by humans and computers. Proceedings of the Workshop on the Integration of Gesture in Language and Speech, pp. 253–266.

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

Corballis, M. C. (1999, March). The gestural origins of language. American Scientist 87(2). [http,//www.amsci.org/amsci/articles/99articles/corballis.html]

Decety, J., & Grèzes, J. (1999). Neural mechanisms subserving the perception of human actions. Trends in Cognitive Sciences 3, 172–178.

diSessa, A. A. (1993). Toward an epistemology of physics. Cognition and Instruction 10, 105–225.

Duit, R., Roth, W.-M., Komorek, M., & Wilbers, J. (2001). Fostering conceptual change by analogies: Between Scylla and Charybdis. Learning and Instruction, 11(4-5), 283–303.

Fox-Keller, E. (1983). A Feeling for the Organism. The Life and Work of Barbara McClintock, New York, Freeman.

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

Goodwin, C. (1986). Gestures as a resource for the organization of mutual orientation. Semiotica 62, 29–49.

Haviland, J. B. (1993). Anchoring, iconicity, and orientation in Guugu Yimithirr pointing gestures. Journal of Linguistic Anthropology, 3, 3–45.

Haviland, J. B. (2000). Pointing, gesture spaces, and mental maps. In D. McNeill (Ed.), Language and Gesture: Window into Thought and Action, Cambridge, Cambridge University Press, pp. 13–46.

Holzkamp, K. (1983). Grundlegung der Psychologie [Foundations of Psychology], Frankfurt, Campus.

Hutchins, E. (1995). Cognition in the Wild, Cambridge, MA, MIT Press.

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. Nature 396, 228.

Kendon, A. (1983). Gesture and speech: How they interact. In J. M. Wieman & R. P. Harrison (Eds.), Nonverbal Interaction, Beverly Hills, Sage, pp. 13–45.

Kinsbourne, M. (1978). Evolution of language in relation to lateral action. In M. Kinsbourne (Ed.), Asymmetrical Function of the Brain, Cambridge, Cambridge University Press, pp. 553–565.

Kirsh, D. (1995). The intelligent use of space. Artificial Intelligence 73, 31–68.

Lakoff, G. (1987). Women, Fire, and Dangerous Things: What Categories Reveal about the Mind, Chicago, University of Chicago Press.

Lemke, J. L. (1998). Multiplying meaning, Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Eds.), Reading Science, London, Routledge, pp. 87–113.

Lemke, J. L. (1999). Typological and topological meaning in diagnostic discourse. Discourse-Processes 27, 173–185.

Leont'ev, A. N. (1978). Activity, Consciousness and Personality, Englewood Cliffs, NJ, Prentice Hall.

Levinson, S. (1997). Language and cognition: The cognitive consequences of spatial description in Guugu Yimithirr. Journal of Linguistic Anthropology 7, 98–131.

Locke, A. (1997). The role of gesture in the establishment of symbolic abilities, continuities and discontinuities in early language development. Evolution of Communication 1, 159–192.

Mandelblit, N., & Zachar, O. (1998). The notion of dynamic unit, Conceptual developments in cognitive science. Cognitive Science 22, 229–268.

Mattingly, I. M., & Studdert-Kennedy, M. (Eds.). (1991). Modularity and the Motor Theory of Speech Perception, Hillsdale, NJ, Lawrence Erlbaum Associates.

McNeill, D. (1992). Hand and Mind: What gestures Reveal about Thought, Chicago, University of Chicago.

Merleau-Ponty, M. (1945). Phénoménologie de la perception [Phenomenology of Perception], Paris, Gallimard.

Newton, N. (1996). Foundations of Understanding, Amsterdam, John Benjamin.

Ochs, E., Gonzales, P., & Jacoby, S. (1996). "When I come down I'm in the domain state", Grammar and graphic representation in the interpretive activity of physicists. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), Interaction and Grammar, Cambridge, Cambridge University Press, pp. 328–369.

Piaget, J. (1970). Genetic Epistemology, New York, Norton & Company.
Quartz, S. (1999). The constructivist brain. Trends in Cognitive Science 3, 48–57.
Roth, W.-M. (1996a). The Co-evolution of situated language and physics knowing. Journal of Science Education and Technology 5, 171–191.
Roth, W.-M. (1996b). Thinking with hands, eyes, and signs, Multimodal science talk in a grade 6/7 unit on simple machines. Interactive Learning Environments, 4, 170–187.
Roth, W.-M. (1999). Discourse and agency in school science laboratories. Discourse Processes, 28, 27–60.
Roth, W.-M. (2000). From gesture to scientific language. Journal of Pragmatics 32, 1683–1714.
Roth, W.-M. (2001a). Designing as distributed process. Learning and Instruction, 11, 211–239.
Roth, W.-M. (2001b). Situating cognition. The Journal of the Learning Sciences 10, 27–61.
Roth, W.-M., & Duit, R. (1997). Learning in real time, How understandings emerge from physics students' laboratory activities. In M. G. Shafto & P. Langley (Eds.), Proceedings of the Ninth Anual Conference of the Cognitive Science Society, Mahwah, NJ, Lawrence Erlbaum Associates, p. 1028.
Roth, W.-M., & Lawless, D. (in press-a). Scientific investigations, metaphorical gestures, and the emergence of abstract scientific concepts. Learning and Instruction.
Roth, W.-M., & Lawless, D. (in press-b). Signs, deixis, and the emergence of scientific explanations. Semiotica.
Roth, W.-M., McGinn, M. K., Woszczyna, C., & Boutonné, S. (1999). Differential participation during science conversations: The interaction of focal artifacts, social configuration, and physical arrangements. The Journal of the Learning Sciences 8, 293–347.
Roth, W.-M., McRobbie, C., Lucas, K. B., & Boutonné, S. (1997). The local production of order in traditional science laboratories: A phenomenological analysis. Learning and Instruction 7, 107–136.

Roth, W.-M., & Welzel, M. (2001). From activity to gestures and scientific language. Journal of Research in Science Teaching, 38, 103–136.

Steels, L. (1997). The synthetic modeling of language origins. Evolution and Communication, 1, 1–34.

Suchman, L. A. (1987). Plans and Situated Actions: The Problem of Human-Machine Communication, Cambridge, Cambridge University Press.

van Gelder (1998). The dynamical hypothesis in cognitive science. Behavioral and Brain Sciences, 21, 616–665.

Varela, F. J., Thompson, E., & Rosch, E. (1993). The Embodied Mind: Cognitive Science and Human Experience, Cambridge, MA, MIT Press.

Weissberg, J.-L. (1999). Présences à distance. Déplacement virtuel et réseaux numériques, pourquoi nous ne croyons plus la télévision [Presence at a Distance. Virtual Displacement and Numerical Networks, Why we no longer Believe in Television], Paris, Éditions l'Harmattan.

Welzel, M. (1995). Interaktionen und Physiklernen, Empirische Untersuchungen im Physikunterricht der Sekundarstufe I [Interactions and Physics Learning, Empirical Studies of Learning Physics in a Secondary 1 Class], Frankfurt, Germany, Lang.

Widlok, T. (1997). Orientation in the wild: The shared cognition of Hai||om bushpeople. Journal of the Royal Anthropological Institute, 3, 317–332.