The Mathematics Enthusiast

Volume 8
Number 1 *Numbers 1 & 2*

Article 4

1-2011

Transcriptions, Mathematical Cognition, and Epistemology

Wolff-Michael Roth

Alfredo Bautista

Transcriptions, Mathematical Cognition, and Epistemology

Wolff-Michael Roth & Alfredo Bautista
University of Victoria

Abstract: The epistemologies researchers bring to their studies mediate not only their theories but also their methods, including what they select from their data sources to present the findings on which claims are based. Most articles reduce mathematical knowing to linguistic/mathematical structures, which, in the case of embodiment/enactivist theories, undermines the very argument about the special nature of mathematical knowing. The purpose of this study is to illustrate how different transcriptions of mathematics lessons are generally used to support different epistemologies of mathematical knowing/competence. As part of our third illustration, we provide embodiment/enactivist researchers with an innovative means of representing classroom interactions that are more consistent with their theoretical claims. We offer a comprehensive transcription, which, when treated by readers in the way musicians treat their scores, allow them to enact and feel the knowledge that the article is about.

Keywords: Transcribing • Epistemology • Enactivism • Performance

1. Introduction

1.1. The problematic: theories and research data

Our theories about knowing and learning mediate how we look at the world generally, and at the data sources we collect as part of mathematics education research more specifically. The currently most dominant theories have come to us through a lineage of work from Kant to Piaget and (radical, social) constructivism. In these theories, knowing is thought of in terms of a mind that constructs itself (e.g., von Glasersfeld, 1991), or as a "collection of minds" that first construct knowledge together before constructing it individually (e.g., Cobb, 1999). More recently, embodiment (Lakoff & Núñez, 2000) and enactivist theories (Davis, 1995) have been proposed to mathematics educators. In these theories, knowing is not supposed to be reduced to the mind that constructs itself but is to be considered in terms of mind that arises from intentional bodily engagements with the world.[1] Embodiment theorists tend to focus on the relation between sensorimotor schemas – e.g., the source-path-goal schema – and similar structures in language. The transition between the two, that is, the transformation, is said to occur by metaphorization processes. Empirical support for each of these theories is provided by particular data produced in and

[1] It has been shown that the very framing of embodiment/enactivist theories in terms of *intentions*, *material body*, and *world* gets us further into metaphysics and body mind distinctions rather than out of it (Henry, 2003). A way of framing a non-metaphysical theory of mathematical cognition has been proposed (Roth, 2010a, in press).

through mathematics education research, presented in the form of transcriptions of communicative situations – e.g., clinical interviews, classroom conversations, or written tests. In this article, we show that some of these transcription forms do not support the theories they are intended to support and other forms of transcriptions contain interactional detail that some but not other theories can explain. In the following section, we provide an example of enactivist/embodiment theories.

1.2. Data and epistemology: the case of enactivist/embodiment theories

Enactivist scholars tend to encapsulate their theories around the diction *knowing is doing*. Many mathematics educators do not buy into enactivist/embodiment theories. Thus, for example, one critic (rightfully) questions the sources of the metaphors offered by Lakoff and Núñez: "Do they really form a natural basis for our thinking, or are they the logical creations of the authors, who are trying to develop a consistent epistemology" (Dubinsky, 1999, p. 557). For embodiment/enactivist theories to become reasonable alternatives to going conceptualizations of mathematical knowing – those fundamentally based in Kant's analyses – they have to show that there is a *necessary* link between moving about (and sensing) in the world, on one hand, and understanding mathematical concepts, on the other. However, the nature of their data and way in which embodiment/enactivist mathematics educators present these works against them. This idea constitutes the starting point of the present article.

To sharpen the problematic of the relation between data and theory, consider the following example. The paper that introduced many mathematics educators to embodiment presents the mathematical idea of continuity as a case study (Núñez, Edwards, & Matos, 1999). Paradoxically, their article consists entirely of text and mathematical formalism – e.g., the statements "$\lim_{x \to a} f(x) = L$" and "if $0 < |x - a| < \delta$, then $|f(x) - L| < \varepsilon$." In that article, therefore, knowing mathematical continuity is reduced to language and language-like formulations. That is, despite the rhetoric about the embodiment of mathematics, the authors only appeal to our mind and obliquely point to the embodied dimensions of knowing without *directly* addressing or appealing to them. Moreover, it may be that culturally and historically these formulations have been derived from embodied experiences; but this does not *necessitate* similar experiences on the part of mathematics learners who live today (Husserl, 1997).

It is not surprising, therefore, that mathematics educators ask what embodiment theories – to take but one example – have to offer to the teaching and learning of mathematics (Dubinsky, 1999). Children may learn about *cylinders* without having had the same experiences as early Greek mathematicians and mathematics learners, for whom the concept arose from the experience of rolling objects metaphorically extended to the concept "cylinder." The ancient Greek used this experience, associated with the word *kúlindros*, roller, derived from the verb *kulínd-ein,* to roll to develop the mathematical-deal *concept* of the cylinder. In fact, the Greek word has even more ancient roots in the Proto-Indo-Germanic *(s)kel-*, to bend, crooked. That is, for the Greek, the word *kúlindros* (cylinder) was an active rather than a dead metaphor, a term that has been carried (Gr. *férein*) across (Gr. *meta*) from the everyday experience of rolling things to the mathematical entity.

In our viewpoint, the main argument of embodiment/enactivist researchers would be much stronger if the data they produce actually forced readers to mobilize forms of knowing that cannot be reduced to linguistic/mathematical structures. Similarly, perception constitutes a form of consciousness that reflects reality differently than intellectual (verbal) consciousness, leading to the fact that the former cannot be reduced to latter (Merleau-Ponty, 1945; Vygotsky, 1986). A verbal transcription of an event, therefore, never renders those aspects in which perceptual consciousness differ from intellectual consciousness. On the other hand, more advanced forms of transcriptions just might exhibit structures that (radical, social) constructivist can no longer explain, or for which they need to develop extensions of their theory so that it continues to provide a viable account of mathematical knowing.

1.3. Purpose

In this article, we present different approaches to representing mathematical communication (knowing) and we show how the resulting transcriptions offer different forms of data that support some but not other epistemologies. Besides, , and most relevant to our own work and theoretical commitments, we develop a means for embodiment/enactivist mathematics educators to show which aspects of the body are *necessary* for understanding formal mathematics. Our representations of lesson fragments relate to knowing mathematics as musical scores relate to the performance of a symphony. That is, we suggest that if someone is capable to read a score, this does not mean that the person knows, or knows how to play, the music with an instrument. This reader does not inherently know what the person referred to in the score has exhibited in his/her performance. Just as the (practical) performance of the music cannot be reduced to the symbols of the score (notes, figures, etc.), the mathematical performance cannot be reduced to the words that appear in transcriptions.

2. Knowing and representations thereof

Historians (e.g., Kuhn, 1970) and sociologists of scientific and mathematical knowing (e.g., Barnes, Bloor, & Henry, 1996) have shown that there exists an interactional relationship between theories and observation. This relationship has been captured in the diction that "If observation is 'theory-laden,' theory is 'observation-laden'" (p. 92). Such is not only the case for mathematics and science but also for research in mathematics (and science) education. Our (authors') own commitments are to embodiment and enactivist theories of cognition. But we have realized only of late that the real issue in the debate may be due to the nature of the data: enactivist/embodiment researchers do not produce the kind of data that would show the *necessity* of the body in and to mathematical knowing. We therefore present the background to the present problematic of data and theory in terms of our own theoretical commitments.

2.1. Practical understanding and formal knowledge

On both cultural-historical and ontogenetic scales, knowing-how in (practical understanding of) the world precedes formal theories. Thus, everyday understandings and the measurement of objects and places preceded and constituted the grounds of formal geometry in ancient Greece (Husserl, 1939). Children learn to speak their mother tongue without knowing any formal grammar whatsoever. High-performance athletes, such as football or soccer players, do not have to know an ounce of physics to make a successful pass even under the most adverse, weather-related conditions. Practical mastery generally does not require symbolic mastery. However, when tennis or golf players do want to change the way in which they play their balls, then they often seek a different form of understanding. They think *about* their play; and this thinking requires signs for a mediated access to their practical understanding. Yet it is also widely known that while they are conscious of their play, these athletes tend to play worse than they have done before or will afterward. That is, symbolic (conscious) access interferes with the playing itself, which tends to be based on unmediated relations between players and their lifeworld. However, the symbolic access to practice is required to *think about* what one is doing.

In the history of human practices, these symbolic forms of knowing – i.e., symbolic mastery – began to separate from the practical understanding of the world. Thus, for example, formal architecture began to develop and separate from master craftsmanship around the time that the great Gothic cathedrals were built (Turnbull, 1993). Prior to the separation, the craftsmen had no plans or knowledge of structural mechanics. The cathedrals were built based on the bodily embodied design skills of the master artisans, working with templates, strings, and embodied geometry in the context of a community of artisans. From the occupation of master craftsmen evolved architects, and craftsmen no longer did design. The new architects concentrated on designing buildings, including the ways in which the strength and stability of the walls had to be increased to make them larger and larger. There is therefore a separation between practical mastery of building cathedrals and symbolic mastery underlying the construction thereof. In a similar way, the peoples around the world developed and played different forms of music before developing means of representing music in a formal way (Treitler, 1982). The point that enactivist/embodiment and practice theorists make is that formal mastery requires some form of practical (embodied and enacted) understanding of the world that is *always present* and in fact *required* by formal mastery. However, it is precisely this latter part that scholars in the field *do not* make apparent and evident in their presentations.

In the theory of textual interpretation, it is well known that explanation *requires* practical understanding of the world (e.g., Ricœur, 1991). Thus, the practice of textual interpretation involves two moments that mutually constitute each other. On one hand, there is practical understanding that we evolve while and through participating in the world. For example, children learn to speak a language and to count before knowing grammar or arithmetic. On the other hand, there is explanation. The point theorists of hermeneutics make is that explanation *cannot* occur without practical understanding, which precedes, accompanies, and concludes explanation. That is, practical understanding completely envelops explanation; but it is through explanation that practical understanding is developed. Thus, children already have to speak language before they can engage in explaining how language works – that is, before they learn grammar. It is evident that to know formal grammar, one has to know language – without language, there would be no need to theorize something like language, there would be no way of asking the question of formal *versus* practical understanding, and so forth.

The same point has been made in a study of categorization in the social sciences (Garfinkel, 1967). Graduate students in sociology had been asked to categorize medical records according to a set of criteria that the supervisors of the research project had created. The purpose of the project was to find out how hospitals worked based on the records that the various personnel created in the course of a patient's trajectory. It turned out that the graduate students, in their classification work, drew on the very type of knowledge that the study was to yield from an analysis of the hospital records. That is, the graduate students drew on their practical understanding of hospital work and organization to classify the records such that the researchers could find out about the practical understanding that makes hospitals work the way they do. The medical records simply constitute formal representations; and to understand them, the practical understanding of how hospitals work is required.

2.2. Mathematical representation and mathematical work

The relationship between practical understanding and formal representation thereof has been conceptualized as the relation between practical action – i.e., *work* – and its formal representation – i.e., the ways in which it is *accounted for* (Garfinkel & Sacks, 1986). Formally, this relation, for the proof of the sum of the interior angles of a triangle, is represented in the form of "doing [proofing that the sum of the internal angles of a triangle is 180 degrees]." Here, "doing" designates the *work* for which "proofing that the sum of the internal angles of a triangle is 180 degrees" are the notational particulars. Take the diagram in Figure 1. It can be taken as the notational particulars of a proof that the sum of the internal angle of a triangle is 180 degrees. But these notational particulars constitute only the formal representation. They do not denote the actual work of *doing* the proof. That is, the formal representations stand in as accounts of the work but do not denote the work itself, and, therefore, they do not denote the knowing underlying the production of the account (Garfinkel, 1996). Knowledgeable readers will easily show, using Figure 1, why the sum of the internal angles of a triangle has to be 180 degrees. And it is precisely this bodily and embodied *work* they do in such a showing that constitutes practical understanding of mathematics (geometry). It is precisely this work that embodiment/enactivist mathematics educators do not sufficiently analyze, show the structure off, and theorize. If this work requires forms of knowing that are not present in the account (e.g., Figure 1), especially, if it involves embodied forms of knowing (e.g., sensorimotor knowing) that have to be enacted in the process of doing, then there exists the *necessary* condition for formal mathematics. But these are precisely the kinds of data lacking in current enactivist/embodiment accounts of mathematical knowing because the transcriptions offered do not point readers to or require the enacting of the *work*. It is only in doing such work that a person can *feel* what it means to *do* mathematics.

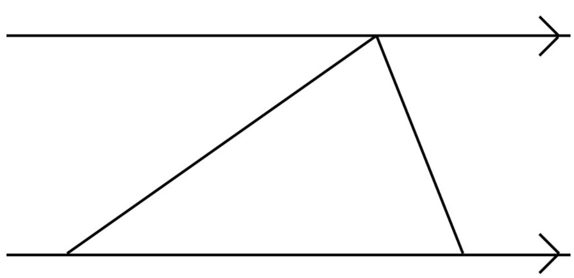


Figure 1. Account of proof that the sum of the internal angles of a triangle (on the Euclidean plane) is 180°.

This way of thinking about mathematics also allows us to understand the debate between Núñez (e.g., 2009) and his critics (e.g., Goldin, 2001). The former points out that the structures of mathematics – e.g., the different notions of continuity – are the results of cultural-historical contingent metaphorization processes whereby practical, bodily and embodied understandings of continuity lead to formal, objective mathematics that anyone can reproduce anywhere in the world. The critics however focus on the formal representations, the diagram (Figure 1) and the fact that the sum of the internal angle of a triangle is 180 degrees. This representation is objective in the sense that the proof can be reproduced over and over again, and each time the result is 180 degrees. This constitutes the objective part of geometrical science (Husserl, 1939). For Núñez it is the embodied work that matters; but it is precisely the work that is not represented in or pointed to by his transcriptions. Thus, we (authors) find that embodiment/ enactivist mathematics educators have by and large failed to provide accounts in which the nature of this work has become available. They have failed because they offer up *formal* properties (e.g., Núñez and colleagues on continuity) and verbal descriptions rather than the non-formal properties of mathematical communication that underlie and ground the formal ones. What such scholars must offer to be more convincing are representations of mathematical activity that allows access to and shows the necessity of the practical, bodily and embodied dimensions of mathematical *work*.

The purpose of this article is to exhibit a form of transcribing mathematical communication that provides readers with access to the bodily and embodied work that one can feel when *doing* mathematics. We propose a kind of transcription that is something like a recipe, which does not in itself represent the work but provides guidance for action. In doing what the transcription denotes, through, and with their own embodied performances, readers perform the mathematical communication presented in the transcription. Whether they have successfully followed the transcription can be established only after the fact. That is, like with any recipe or musical score, the formal representation is not a causal antecedent of the work, though it is a resource in and for the practical action (Suchman, 1987). A simple word-by-word transcript of a lesson may not be sufficient to exhibit what students in a mathematics classroom actually know. It will exhibit even less the didactical skill of a teacher, who may know, just because of the way a student speaks in an interaction, whether the student speaks with certainty, whether she likely or unlikely knows, and so on. This, then, is precisely our point of departure for developing transcripts, which we suggest should be used as scores that readers have to enact rather than just read – much like a musician who picks up the instrument and plays a tune rather than read sheet music and much in the way a (hobby) cook actually makes a dish rather than just read a recipe book and marvel at the accompanying images.

3. Representing mathematical communication/knowing

In this section, we provide a fragment from a second-grade geometry lesson to exemplify the kinds of data that different forms of transcriptions make available. We provide sample analyses that the proposed transcription supports and that the analyses can explain. We show, for example, how a particular kind of transcription supports constructivist claims about stable knowledge structures; we also show that this requires particular reductions where any temporality is removed from the transcription. We are specially interested in producing transcription and transcription use that lead to a better understanding on the part of researchers of precisely *what* the students' knowledge consists in. Our contention is that if researchers only focus on what can be presented in text, they know very little about what precisely the interaction participants know.

The fragment was randomly selected from 30 hours of recordings in a second-grade mathematics class in the process of completing a unit on three-dimensional geometry. It derives from a lesson in which children were provided with a shoebox containing a "mystery object." The object could be reached and touched through a hole in the shoebox but not seen, as there was a plastic bag taped to the inside. That is, the children could only touch/feel the object by sticking their hand through the hole and into the plastic bag, which separated their hands from the object. The video shows the three girls – Sylvia (S), Jane (J), and Melissa (M) – at a large, round table on which their shoebox is placed (Figure 2). The research assistant Lilian (L) videotaping this group also participates in the conversation transcribed. From the beginning of the modeling task, Melissa has repeatedly said that she feels a cube; and she has built a cube from her lump of plasticine. Jane and Sylvia have formed rectangular prisms of similar shape from their respective plasticine lumps. But the teacher explicitly has instructed the students to produce one and the same model and, if there is disagreement about its shape, to discuss until they reach agreement. The fragment picks up when Melissa asserts once again that she "thinks it is a cube" just as she pulls her right hand back from the shoebox after another trial of feeling the mystery object. In the following, we provide three takes on the fragment leading up to a different form of representing the events with consequences for the kinds of conclusion that can be made and are supported by the fragment.

Figure 2. Sylvia, Jane, and Melissa (from left to right) are in the process of building models of the mystery object inside the shoebox.

3.1. Take 1: logocentrism

Most transcriptions that appear in mathematics education journals reduce events – lesson, interviews, or problem-solving sessions – to the transcription of the words said, augmented by ethnographic descriptions of actions and context where necessary. Moreover, the words are not taken for and by themselves but rather as indices pointing to something else not directly present: "meaning," "conception," or "idea." It is precisely these two strategies that lead to the separation of body and mind and lend themselves to Kantianism and other constructivist theories (Henry, 2003; Nancy, 2007).

Transcribing videotape by using only words flattens the observed events into language. The ancient Greek originally used the term *logos* for language and word; they later also used it to denote reason, a use that has survived to the present day sedimented in the term "logic" (Heidegger, 2000). By transcribing events into words, we obtain a representation thereof where everything that exists is named and, being in the form of words, is reduced to the form of intellect and reason. In the philosophical critique of metaphysics, this tendency to reduce everything to words and reason (i.e., *logos*) has come to be denoted by the term *logocentrism* (Derrida, 1967), a way of thinking about *being* that has its origin in the ancient Greek culture and has shaped the Western way of relating to the world. That is, the idea of rational thought apart and independent from the material world, metaphysics, is bound up with the practice of reducing complex situations to words and verbal description.

3.1.1. Producing the transcription

To produce transcriptions of this first type requires little else than playing a video and noting the words heard. Generally, we produce such transcriptions using a digital video file (.mov format) and then transcribe the words we hear directly into a word processing program. Where transcribers hear someone speaking but without being able to make out specific words, question marks are used to indicate the approximate number of words (e.g., <??> to indicate two words). The transcriber also inserts verbal descriptions of actions where appropriate or necessary. Many transcribers/researchers also insert punctuation that follows common grammatical practices. That is, where the transcriber hears a question, a question mark will be inserted at the end of the sentence independent of the fact how participant listeners have heard the current speaker as evidenced in their subsequent turns.

Transcript 1

```
01 M:  ((after putting her hand in the box for a while)) I still think it is a
       cube.
       ((The whole group pauses))
02 S:  Let me check ((puts her hand into box)).
03 L:  Why do you think it is a cube?
04 M:  Because it's the same; it's the same ((turns her model over in her
       hands)).
```

3.1.2. Reading, analyzing, and theorizing the transcription

Characteristic of this form of transcript is the removal of temporality of all dimensions of participants' action, not only regarding the production of their talk but also regarding their physical behavior (e.g., gestures, body position, transactions with physical object/s, gaze orientation). As readers can see, the transcript presented above is reduced to the order in which words have been pronounced. The verbal description of the hand/arm movement no longer renders the temporality of the movement and is not coordinated with the temporal unfolding of the speech. Because temporality has been removed, the forms of thought said to be "behind" the utterance are taken to be relatively constant over the length of a typical lesson or interview. Such a description, by and large static, facilitates making claims about "conceptions" and "conceptualizations" that can be sampled unproblematically in an interview. Researchers tend to make no difference between some word used at the beginning, in the middle, or at the end of an interview.

Most mathematics education researchers take such transcriptions and infer "meanings" and "mental structures" that somehow are in the speakers' minds and that have led them to say what they said. For example, a mathematics educator interested in our work took the video and transcript, concluding from the episode that "Melissa (initially) conceptualizes the mystery object as a cube. She bases her conclusion on the tactile observations she makes by turning the object over and 'checking the sizes' of its faces." Here, the verbal articulations and descriptions of movements become indices for something that is not directly available. On one hand, there is Melissa saying, "I still think it is a cube," and on the other the mathematics educator claims that "Melissa (initially) *conceptualizes* the mystery object as a cube." The relation between word and thought (mind) is taken to be as a rather simple one, the former providing access to the latter. Thus, in mathematics education research, verbal transcriptions of interviews and classroom videotapes are regularly used to find out what and how students think, how they solve problems, or how they "construct" their mathematical mental structures (or, conceptions, representations, or even identities).

3.1.3. Discussion

Nearly 80 years ago,, it has already been suggested that "thought is not merely expressed in words . . . the structure of speech does not simply mirror the structure of thought (Vygotsky, 1986, p. 218–219). All three – speech, thought, and the relation between the two – are *processes*. We do not see any evidence for a conceptualization, unless simple word use is taken to be synonymous with conceptualizing something. Instead, there is evidence for the fact that students and adult talk about phenomena even before they have thought about and reflected upon some idea (phenomenon, topic), and, therefore, *could not have formed* (i.e., "constructed") a concept (Roth et al., 2008). Rather, thought is the *consequence of* speech, comes to existence through speech. Moreover, whereas it might be appropriate to say that Melissa "turned over the cube," the simple description of this action in words may overstate the issue. For Melissa may have turned the cube in the way we walk or scratch an itchy spot: it does not require our conscious intentional thought. We also do not know whether Melissa was intentionally "'checking the sizes' of its faces." Rather, we observe her using the thumb and index of the right hand in apparently the same or slightly changing configuration along three different edges of the cube while articulating that some "it" – which we do not know whether it is an edge, a face, her cube, or the mystery object – "is the same." That is, as soon as something is articulated in words, it is moved from the realm of Being, presence, and presentations into that of beings, present of the present, and re-presentations (Heidegger, 2000). Moreover, in this realm, it is subject to verbs that inherently embody intentionality (Henry, 2003).

This kind of transcript is consistent with a constructivist approach, which, at least since Kant, is concerned with abstractions and abstract thought. In Piaget's theory, we find this gesture((what gesture??)) in the development from concrete operations that lead to formal thought as embodied sensorimotor schema are abstracted and become the pattern for logical thought. It is also a description that runs counter to the epistemologies of embodiment and enactivism because it emphasizes a conscious mind and mental structures in situations that may not be appropriate. Thus, whereas it is evident that we would not characterize a person as *consciously* placing feet in walking, there is a tendency in mathematics education research to use an *intentionalist* discourse when it comes to describe what children/ students do in the mathematics classroom: "construct meaning," "develop conceptions," "acquire knowledge," "position themselves," "construct identity," and so on. Interestingly, though, scholars interested in mathematical cognition from both embodiment and enactivist camps, too, make use of such transcriptions, thereby doing a disservice to their argument. It is not surprising then that many mathematics educators opposed do not buy into embodiment and enactivist theory, as everything there is made available in such transcription is at the verbal level itself an image of the concepts thought of in metaphysical, linguistic terms.

3.2. Take 2: sequential analysis of turn taking

The afore-described constructivist inferences are inconsistent with social/cultural-historical theory that theorizes speech (communication) and thought as continuously developing processes that mediate their respective developments (Vygotsky, 1986). That is, thought and speech are different, incompatible expressions of some higher order unit; and they are processes. Thus, from such a perspective we have to take Transcript 1 as a temporal event in which not only speech unfolds from top to bottom but thought as well. Moreover, in such a theory, gesture and speech are dialectically related; they are manifestations of a higher order communicative unit rather than precisely corresponding to each other (McNeill, 2002). That is, as speech unfolds so do gestures; and speech and gesture mediate their mutual development in the same way as speech (communication) and thought. In this section, we provide a form of transcription and approach that lends itself to viewing thinking, speaking, and their relation as processes.

We begin this second take by representing the fragment in an augmented way typical of conversation analysis. This transcription form includes all the sounds produced, pauses, hesitations, respiration, prosodic information, and emphases (see Transcript 2). The approach is grounded in a history of ideas of language philosophy that what matters to understand language are not "meanings" but the ways in which words are used (Wittgenstein, 1958). Subsequent developments in language philosophy focused on speech acts (Austin, 1962). A speech act consists of three parts: *locution, illocution*, and *perlocution*. Locution refers to the act of saying something, illocution to the intent (asking, ordering, responding), and perlocution to the effect. In any concrete analysis, the effect that a locution has on others in the setting is available only in and through their subsequent acts. Consequently, to understand a speech act, researchers have to take the *turn pair* as the minimal unit of analysis. That is, it is no longer possible to attribute speech to an individual because a speech act is inherently spread across multiple participants, across speakers/audiences. This is consistent with a conceptualization of discourse in which any utterance straddles speaker and listener, where any word – spoken for the benefit of another – belongs to both speaker and listener (Bakhtine [Volochinov], 1977; Derrida, 1996). This way of approaching transcription and its interpretation therefore focuses on understanding this event as unfolding event, as something living and lived, rather than on purported structures of individual minds whose contributions to the conversation are independent of those of others.

3.2.1. Producing the transcription

Notice how Transcript 2 adds features that were not present in the first transcription. (The differences in the text itself derive from the fact that the original transcription was done by someone else, and subsequent enhancements revealed problems in the original hearing.) For example, pauses within speaking turns and between speaking turns are measured and indicated to 1/100th of a second. The transcription also marks emphases (capitalization), partial sounds ("sti"), mispronunciations ("cob")[2], extended sounds (colons), and trends of the pitch (punctuation). Thus, the transcription renders aspects of the real time production of speech; that is, it contains the mumbles, stumbles, stutters, breathings, malapropisms, metaphors, and tics characteristic of everyday speech. Conventions to produce this kind of transcripts can be found in Appendix A.

Transcript 2

```
01 M:  ((pulls rH out of box, pushes it away)) I sti (0.18) I s::TILL think it
       is a cube.
02     (1.66)
03 S:  ((S picks the box, turns it, reaches in)) LET me CHECK.
04 L:  WHY do you think its a CJOB (.) CUBE.
05     (0.20)
06 M:  CAUSE like (0.31) the SAME ((turns cube and has caliper grip with
       thumb/index)) (1.13) its the SA::ME shape.
07     (1.55)
08 S:  WHERE i:s IT; ((reaches into the box))
```

The production of such transcriptions begins with word-by-word renderings such as those in Transcript 1 but with punctuation removed, as it is used to mark the pitch tendency within the locution. We export the sound from the video into an audio format (.aif) so that it can be imported into a program for linguistic analyses. A freely downloadable, multi-platform package frequently used by linguists is PRAAT (www.praat.org). It allows precise timing of pauses in speech, measurement of speech intensities (volume), pitch (F0) levels, and speech rates. Speech emphases can be heard and – because these are produced by means of changing intensity, pitch, or rate – can be verified by visual inspection of the PRAAT display. The display also allows identification of pitch jumps and within-word movements, which are indicated in the transcript using specific signs. The conventions used follow published conversation analytic conventions that are enhanced for the analysis of prosody (Selting et al., 1998).

[2] The research assistant, Lilian, is a native Portuguese speaker. In that language, cube is *cubo*. An interference might have occurred between the pronunciations of cube (IPA: kju:b) and *cubo* (kubó).

In those instances where visual information is relevant, screen prints or drawings are imported into the transcription or provided in an accompanying figure (see below). The precise timing of the visual information with the speech is indicated in the transcription. When drawings are used instead of screen prints – which may be to implement confidentiality or to feature only essential information while dropping gratuitous detail – the off print is imported into Photoshop. A second layer is created and an outline copy of the essential information is produced using the "paint brush" (see Figure 2, 3). To make essential elements stand out even further than they would in a pure line drawing, different degrees of shading may be used.

3.2.2. Reading, analyzing, and theorizing the transcription

Focusing on the second transcript presented, we first note that the locution in turn 01 is not fluent. There is a beginning "I sti," a pause, another beginning with drawn out "s" before the remainder of the word "still" is completed followed by "think it is a cube" that will have completed the locution. (We never know whether some word constitutes the end of a locution or speaking turn until some next speaker begins to speak, or until the same speaker takes another turn at talk.) Both the repeated articulation of the personal pronoun "I" and the second part of the word "sTILL" are articulated with emphasis (as indicated by the capitalization). This utterance cannot be understood on its own because, from a conversation analytic and speech act theoretic perspective, it is only the second part of a unit, the first part of which is not available in this transcription. In a fuller consideration of the entire episode, a researcher would focus on the emphases, which produce contrasts to the different claims that Sylvia and Jane have made and which make salient that Melissa already has repeatedly made statements about the mystery object as a cube.

Melissa's turn is the first part of what turns out to be two turn pairs. Sylvia says, "Let me check," which allows us to hear the pair of turns as a constative/verification speech act. In fact, Sylvia not only says "let me check," but also pulls the shoebox over close to herself and sticks her hand into it. Her verbal articulation is a *formulation* of the action: Sylvia not only reaches into the box but she formulates for others what she is doing, that is, the she articulates the *intent*. She makes explicit and available to her audience a verbal description of the illocutionary act. Her reaching into the box is formulated as an action that has the intent of checking. Because of the pairing of turns, the checking is heard with respect to the constative "it is a cube."

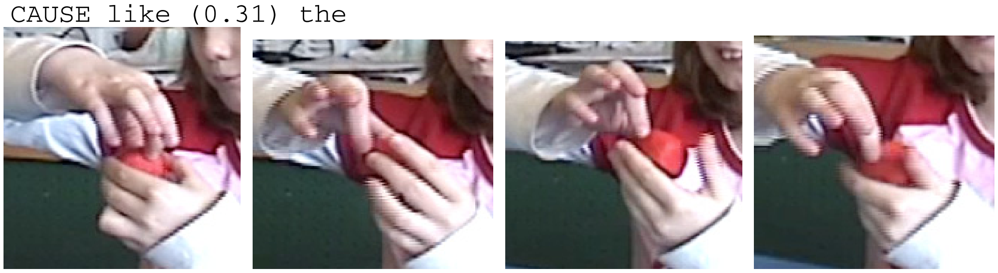
The second turn pair exists in the sequence with Lilian, the research assistant, who is also acting as the teacher of this small group of students. We can hear turns 01 and 04 to constitute a sequence, because Lilian's locution "Why do you think it is a cube" picks up on and repeats the contents of Melissa's utterance. Interestingly, the transcription indicates that the pitch is falling toward the end of the locution, which is typical for constative phrases. But the fact that the interrogative adverb "why" is articulated with emphasis allows us to hear a constative/request-for-justification speech act: "I still think it is a cube" is followed by "Why do you think it is a cube?" This hearing is consistent with the next turn sequence, which we can hear as a question/response pair: "Why do you think it is a cube" is followed by a coordinating conjunction "[be]cause," which introduces a reason, "like the same . . . it's the same shape."

This form of transcript in the hands of conversation analytically informed researchers, therefore, allows readings that focus on the unfolding nature of the event. Such researchers also focus on analyzing pairs of turns, that is, on the effect a locution has on the other participants as their actions make it available to everyone else. There is a focus on the sequential enchainment of locutions (utterances), where turn pairs constitute the minimal unit. This kind of analysis is *process oriented*, allowing us to understand the constitution of this segment. What matters is – consistent with Wittgenstein's (1958) language philosophy – *how* words are used rather than purported and never accessible "meanings" behind the word. Moreover, from a discursive psychological perspective, Melissa's and Lilian's reference to thought processes ("I still think," "Why do you think?") are taken to be everyday ways of reasoning where psychological concepts are invoked for the purposes at hand. Such researchers are little interested in purported contents of the mind; instead, they focus on the mobilization of psychological discourses for the purposes of the situation at hand (Edwards & Potter, 1992).

From a conversation analytic perspective, Melissa's "I think" is taken to be a formulation of the work she is/has been doing at the instant, and Lilian is taking up the self-description as a way of referring to the same work description. It is not the researcher who imputes thought processes – as in the preceding section, where a mathematics educator imputes conceptions – but it is one of those ongoing descriptions that interaction participants provide to articulate the situation together and *for one another* with the content. Here, the content is the nature of the model Melissa has built, and its relation to the mystery object. It is the situation itself that suggests the use of the "thinking" as a description, and the available language form to describe what she has been doing is that she is "thinking." An alternative might have been to say, "I feel it to be a cube" or "I believe it to be a cube."

In this transcript, because the gestures are described in words, their contribution to the communication comes to be evaluated purely in terms of the linguistic sense ("meaning") that researchers attribute to them. In classical conversation analysis, gestures were not attended to – in part because the research was based on audio-recorded conversations on the telephone. But many conversation analytically and ethnomethodologically oriented studies of this nature focusing on mathematics – following the ground-breaking work of the applied linguistic Charles Goodwin (e.g., 2000) – now include precise studies of gesture. In our own work on the role of gestures in science learning, we precisely coordinated information about gestures with speech because, as it turned out, the changes were related to familiarity and expertise of the speaker within the domain talked about (e.g., Roth, 2000). These studies included transcriptions such as the following rendering of turn 06, in which vertical lines indicate at which point a particular hand/arm configuration occurred. (Vertical bars coordinate speech and image.)

```
06 M:  CAUSE like (0.31) the
```



```
       SAME |      (0.66)      |     (0.47)   |        its the | SA::ME shape.
```

In this transcript, we observe the rotation of the cube held in the left hand and an associated movement of the right hand, the thumb and index finger of which grab the plasticine "cube."[3] The transcription clearly shows that three bodily configurations precede the articulation of the predicate "it's the same shape," and the fourth configuration also precedes the second, key part of the predicate "same shape." This key part is further of interest, as the word "same" is drawn out (see colons in transcription), which might be heard – depending on context – as an emphasis or as a delay in the verbal performance. Psycholinguists often focus on the relation between gestures and the contents of speech that is said to correspond to the former (Roth, 2003). It turns out that developmental studies of mathematics, for example, show that gestures expressing a new developmental level *precede* verbal expressions at the same conceptual level (Alibali, 1999). That is, words and gestures manifest *very different forms* of knowing. In fact, when the conceptual content of the gestures is different from those of speech, it is taken as an indicator of developmental readiness (Church & Goldin Meadow, 1986); and without training even teachers and undergraduate students glean information from children's hands (Alibali, Flevares, & Goldin-Meadow, 1997). Using words instead of images to depict children's communication *falsifies* what they are communicating to the teacher or researcher. Moreover, studies in science education show that the alignment between gestures and corresponding speech during conceptual transitions, which may be out by up to three seconds, decreases with students' familiarity in the domain (Roth, 2002). When alignment is achieved, observers tend to assess as competent the explanations of the phenomena that are the current topic. It matters that language and gesture are different in nature, have different content and form, and that they may contradict each other.[4] This form of transcription therefore provides support to theoretical approaches that assume the continuous development of both speaking and thinking at the moment-to-moment and ontogenetic scales, but they are inconsistent with those approaches that theorize stable mental structures.

In this instance, the hand movements may actually not be purely symbolic. The left hand holds the cube rather than gesturing a cube, and the right hand produces a configuration that is applied with little change to the cube that turns underneath it. The situation does not symbolically represent the events that have occurred just seconds before while Melissa has had her right hand in the shoebox, but her left hand remained outside. We do not know what happened inside the shoebox, how and even whether the mystery object has been turned. This is of particular importance later given that the mystery object turns out not to be a cube. But in the present instance, the configuration is repeatedly applied to the different dimensions (x, y, z) of the plasticine model ("cube"). The configuration, therefore, especially when it occurs the first time, constitute an epistemic (knowledge-seeking) movement designed to "check the faces," as the mathematics educator referred to above suggested to us. During the same and other lesson of this geometry curriculum, we did observe purely symbolical movements when the same hand configurations were used in communication in the absence of cubes.

[3] It is a cube but not in the sense of geometry, which only deals with ideal objects. Rather, it is a figure of the kind that preceded geometry (Husserl, 1997).

[4] In dialectical psychology and philosophy, speech and gesture *inherently* are contradictory, each manifesting the communicative content in a one-sided way (e.g., Roth & Lee, 2007).

3.2.3. Discussion

Transcription 2 exhibits temporal features characteristic of human interactions; it also features some of the details of the actual production of communication, including hesitations, false starts, emphases, and so on. This type of transcription – embodied in our conversation analytic reading above – lends itself to theories that include temporal features between thinking and speaking and to theories that focus on the interactional nature of human life and its continually unfolding nature where subsequent states are unavailable to the actors. Moreover, theories that take the actor perspective on social events find such transcription useful, as these contain implicit and explicit information that participants use in the pragmatic conduct of social/societal events, including interviews and mathematics lessons.

One of the questions one might ask is this: Is there something *behind* these performances, some structures, that drive/cause what we observe? In other words, is there knowledge of some kind in the brain that causes the vocal track and the hands/arms to do what they do in order to externalize something that is hidden from direct observation in the brain? Or should we take the verbal and gestural performances as the knowing itself? If the second is the case, as researchers informed by embodiment and enactivist theories claim, then this and the preceding form of transcription are insufficient in two ways. First, because these contain too little information about the communicative productions and expressions themselves; and, second, the relation between knowing as represented and knowing-how of what the representation refers to is the same as knowing to read a recipe and knowing-how to make the dish. We contend that mathematics educators who read transcripts do not (necessarily) have the know-how of these performances; someone who reads a musical score does not (necessarily) know how to play the tune on the musical instrument it was intended for. And it takes precisely the cooking or playing to *know what it feels to cook or play*. In the following section, we address the first of these questions and then make a proposal about how to address the second.

3.3. Take 3: interaction rituals

Recent developments in philosophy and sociology (of emotions) focus on temporality, periodicity, and resonance as fundamental phenomena for the constitution of (common) sense (Collins, 2004; Nancy, 2007). Thus, we can observe an increasing alignment of prosody across speakers within turn pairs among teachers who are working together over several months; and these alignments are coextensive with the sharing of sense in and of the situation (Roth et al., 2005). For example, pitch misalignment is associated with conceptual dissociation and conflict; and rhythmic alignment across speakers and listeners can be observed even when listeners cannot see the speaker's rhythmic body movements (e.g., Roth, 2010b). These rhythmic alignments are sources of emotional alignment and a sense of solidarity (Collins, 2004). Pitch and rhythm are of interest because speakers are not conscious of it. That is, these features of speech and body movement *determine* sense, but, because consciousness is not involved, words only one-sidedly represent the content of communication. This also tends to be the case for speech intensity, though under certain circumstances speakers are conscious of their speech intensity and increase or decrease their volume. In contrast, as part of outbursts of anger, they do not voluntarily control speech intensity. Because these are non-conscious features of communication, these cannot be theorized in the same way as verbal consciousness. Transcriptions including these features therefore lend themselves to provide support to embodiment and enactivist theories and to theories that track the real-time evolution of events from the perspective of the participants (Roth & Pozzer-Ardenghi, 2006).

Our recent work in mathematics classrooms also exhibits the importance of prosody and rhythmic features in the voice, gestures, and body movements. In Figure 3, we provide a more extensive transcription. In the following, we articulate the possible readings it affords consistent with a radical approach to embodiment that has been termed "incarnation" (Roth, 2010a). The following dimensions are represented in the transcript: intensity and pitch of the participants' talk, duration of their utterances (see black boxes), the sounds/words they pronounce, and other relevant embodied dimensions that emerged during the entire episode such as hand gestures performed with the object, body position, and gaze orientation. Because the variable "time" is the main criteria to display our empirical evidence, we suggest below that this transcript is to be treated in the way musicians treat a musical score: as an occasion for playing a particular tune in a particular way. In this way, the rate and total time of playing themselves become performative aspects. As a result, readers will *feel* the type of knowing observed when they re-play the transcript rather than merely look at and read it.

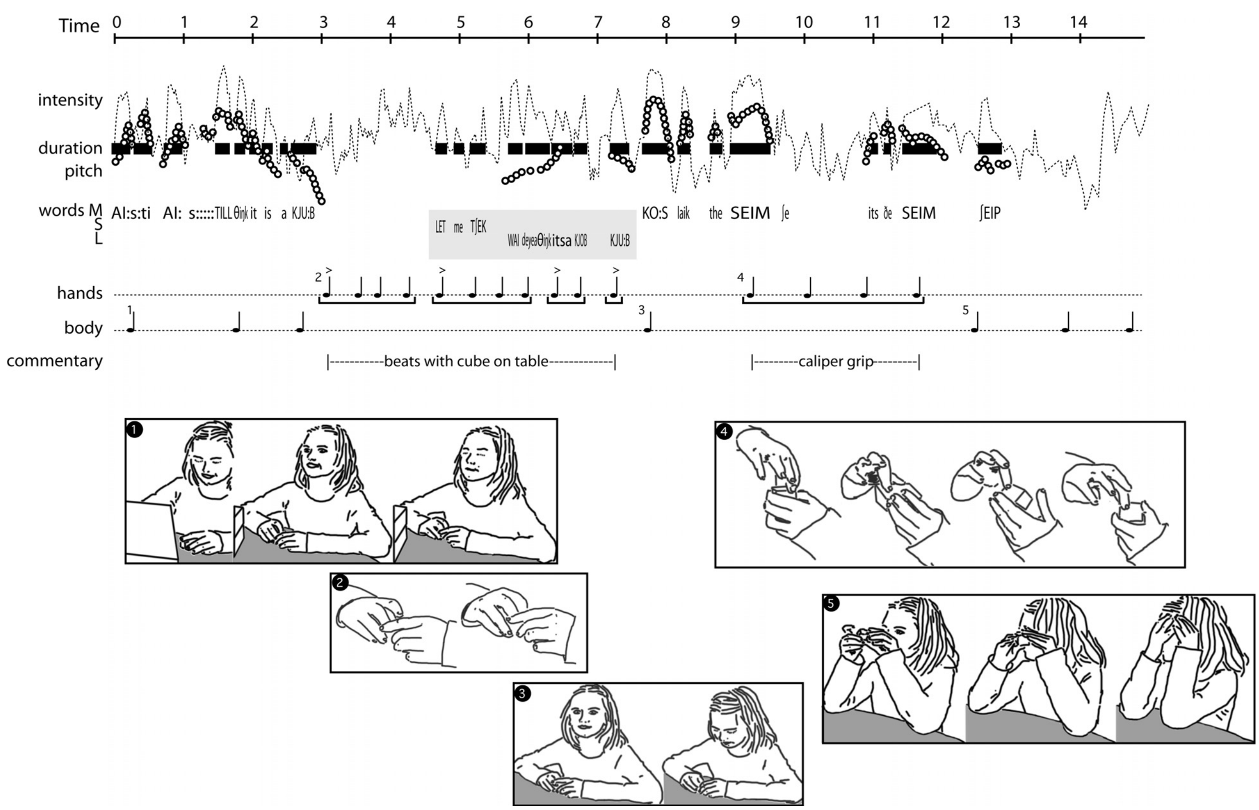


Figure 3. The extensive "transcription" includes prosodic features (pitch, volume, rate), rhythm, verbal, and visual information. The "words" are transcribed using the conventions of the International Phonetics Association

3.3.1. Producing the transcription

As can be observed, this type of transcription uses information that was presented in the preceding types of inscriptions (e.g., words). In addition, the transcription directly maps the sound (phonemes), using the conventions of the International Phonetics Association, onto the prosodic information (Figure 3). Because the phonemes are directly mapped against the prosodic information, changing speech rates, emphases, and rhythms also become visible. We used a graphics program into which the PRAAT display was imported. Using horizontal black bars, the length of the phonemes is indicated. Each word is typed at a specific font size and then changed in horizontal extension until the transcribed phoneme has the same length as the black bar. Moreover, as in a musical score, the melodic line (pitch) and changes in intensity – indicated in musical terms (e.g., piano, pianissimo, forte, diminuendo) in the second type of transcriptions – are given quantitative expression. In addition to the coordination of visual information already present in the augmented version of Transcript 2, these now are associated with the information about repeat patterns. This, therefore, allows exhibiting the rhythmic aspects of a performance, which also would be available in a musical score.

3.3.2. Reading, analyzing, and theorizing the transcription

This transcription (Figure 3) exhibits some striking differences with respect to the preceding Take 1 and Take 2. First, it makes explicit the temporality of all the dimensions of the students' and the teacher's verbal/physical action. Not only is speech in time, it makes time as "words," phonemes, and even individual letters are drawn out or speed up; there are pauses; and there are emphases that punctuate what is being said. For example, Melissa stresses "I," "still," "cube," "cos," "same," "same," and "shape." These stresses with the interspersed more rapid deliveries punctuates the utterance as it unfolds in time; it gives it a particular rhythm. In actual listening, (a) perceiving the rhythm requires a consciousness very different from intellectual consciousness and (b) perceiving the rhythm means producing the rhythm (Abraham, 1995). In Lilian's utterance, the "words" run together making out of "do you think it's a cube" one single sound complex.

We note that the pitch moves up and down, sometimes producing spikes with individual words (e.g., ""cos," "like, "same") and producing overall tendencies (e.g., the pitch drops with the production of "still think it is a cube." Such information is important, as research shows that in harmonious exchanges, speakers tend to latch onto the pitch of the preceding speakers, whereas in conflictual situations, the pitches tend to be significantly apart. In fact, in conflict, the pitch levels tend to rise, each speaker "trumping" over the preceding one so that both may be speaking with fundamental frequencies three to four times above their normal pitch (e.g., Goodwin, Goodwin, Yaeger-Dror, 2002). Thus, for example, one study in a science classroom showed such a phenomenon as a teacher and her student argue about chemical valences, and their argument over conceptual differences come to be reflected in the differing pitch levels; appeasement was associated with falling pitch levels across a number of speaking turns also involving other students (Roth, 2010b). Speech intensity, too, contributes to the way we understand what and how someone else speaks, as interaction participants tend to hear much louder than normal speech as "shouting," in many situations heard as an expression of anger. Much lower than normal speech intensity, in the case of a student who also speaks slowly, may be heard as a sign of timidity, not knowing the answer, or as a tentative exploration of ideas. Teachers use such hearings routinely in their assessments of teaching, yet at present, mathematics education research does not account for these embodied features.

The transcript includes visual information similar to the one we presented in the preceding subsection. For example, the fourth image sequence exhibits the same four hand/finger configurations introduced previously. Here, however, we also mark with a "□" on the temporal axis the precise instant when the configuration is produced. The musical notation exhibits the highly rhythmic feature of the gestural production. That is, the four configurations that exhibit mathematical features – sameness of the length of the edges – are produced in a highly rhythmic fashion, which constitutes a very different manifestation of sameness across the dimensions. Melissa is transacting with a solid characterized by the idea of even number (such as 4 and 2, as demonstrated in the stresses of the beats she produces on the table), and vice versa – the object is transacting with her as well. To a certain extent, it might be argued that the idea of "evenness" emerges from Melissa's physical action while she transacts with the plasticine model.

Comparison with the verbal production shows that the first gestural beat falls together with the emphasized "same"; the second beat falls at the beginning of the pause which in speech, as in music, is an important feature; the third gestural beat coincides with the restart of the verbal "melody"; and the forth beat falls on the second "same." We might expect another beat corresponding to the verbal production of "shape"; but, as our transcription shows in the change of the bodily configuration where the gaze, heretofore exclusively oriented to the hands and cube, now is raised to meet that of Lilian, the person who has requested the justification Melissa has just ended producing. Melissa then turns to gaze at Jane, and finally appears to complete her presentation by enclosing her cube in a gathering movement that also brings the elbows close together. This, therefore, constitutes a continuation of the rhythm but in a different modality, that is, on a different "stave" of our "score of mathematical communication."

Returning to the beginning of the transcription, we note that the changing orientations constitute a rhythmic phenomenon as Melissa orients from her cube to others and back to her cube (image sequences 1, 3, and 5). Between these sequences there are long pauses of speech. The second of these "pauses," as shown above, occurs when Melissa rhythmically produces the four gestures that constitute an integral aspect of the (unconscious) embodied/enacted justification why the mystery object is a cube. The first "pause" in the shift of orientation is associated with a pause in Melissa's speech. There is a long pause, which Sylvia breaks announcing that she is going to check, followed in turn by Lilian's request for a reason. During this pause in speech, Melissa hits the table repeatedly with her plasticine model (in the sound wave, there are spikes that mark the precise instant that the cube hits the table). As our transcription shows (Figure 3), there is a rhythmic beat that is produced and that we can perceive. Not only is this performance rhythmic, but the transcription shows that the beats fall together with the beats in Sylvia's talk; it also coincides with the beginnings of the major segments in Lilian's talk as exhibited by the speech intensity profile (i.e., where she says "deya [do you]," "it's a," "cob," and "cube"). That is, the same rhythm can be perceived in all three speakers, or, if Melissa were to be taken as the main figurant in this instance, the others would be found to have aligned themselves with the beat she has initiated. But, because perception of rhythm means production of rhythm, all of these rhythmic features produce interactive interference that leads to entrainment into the same rhythm. This is precisely what we have observed both in mathematics (Roth, in press) and in science classrooms (Roth, 2010b) where there are rhythmic features in speech and other bodily productions across individuals; and these beat frequencies change across individuals. Thus, it is not that the same beat occurs by chance. Rather, when the speaker changes the beat, others follow, sometimes imitating it and sometimes improvising on the original beat. This is so even though the beat is not accessible to verbal consciousness but constitutes a very different form of consciousness (Abraham, 1996; Nancy, 2007). The perception of beat is a form of active resonance that allows for the alignment through entrainment.

The rhythmic aspects together with the prosody emphasize ritualistic aspects of human interactions. Our transcription therefore is consistent with social theories that focus on interaction *rituals* (Collins, 2004) and sense as a resonance phenomenon (Nancy, 2007). Sense cannot be reduced to words, as integral aspects of sense manifest themselves in and are expressed by non-verbal means. Moreover, the ritualistic moments also are tied to emotion, finding both their expression in the performance and driving this performance.

3.3.3. Discussion

Readers unfamiliar with such analyses might ask why this is important. It is because these changes in rate and intensity are associated with what we hear as main and subsidiary clauses of a sentence (Roth, in press). Whether something is a main or subsidiary clause goes right to the heart of competence in mathematical communication and mathematical understanding. Thus, the prosodic and rhythmic aspects, which appear to have nothing at all to do with the mathematical content – they do not appear in mathematics textbooks – nevertheless are integral and irreducible aspects of mathematical communication and the practice of mathematics. That is, the difference between mathematical content and purely performative dimensions of communicative production is undecidable. They constitute one and the same phenomenon. These analyses therefore are important for those who adhere to embodiment/enactivist perspectives on mathematics education. Mathematics is not embodied because bodily gestures (hands, hand/arm, other body parts) exhibit logical structures that may be seen as parallel to and exhibiting the same verbal-conceptual content. Rather, mathematics is embodied because there are features in mathematical communication and practice that play integral and central role of producing mathematical distinctions, but they are not part of the verbal-linguistic register. More importantly, the two registers are irreducible to each other, each constituting a one-sided and therefore partial manifestation of a higher-order phenomenon of mathematics and mathematical communication. And it is precisely this irreducibility of mathematical linguistic features and purely embodied features (prosody, rhythm, bodily gestures) that support enactivist/embodiment theories.

We propose taking our transcription differently than transcriptions normally are taken in the literature. We suggest that our transcription relates to the performative of mathematical communication as a cookbook recipe relates to cooking or in the way a musical score relates to a musical performance. That is, to really feel the knowing and understanding in Melissa's communication, readers need to perform our "score." Such performances relate to Melissa's in the way one musician's rendering relates to that of another; this relation is different from the one between score and performance. This is especially so because the performative dimensions (such as prosody and the rhythmic performances) are irreducibly involved in the mathematical sense even though they cannot be rendered in terms of linguistic consciousness. Rhythm has to be performed to involve and make it accessible to rhythmic consciousness in the same way that the visual aspects (e.g., hand gestures) require a form of consciousness different from and irreducible to verbal consciousness (Vygotsky, 1986). Performing the transcription, therefore, amounts to a process of *reterritorialization* (Deleuze & Guattari, 1991/2005), whereby something said to be transcendent and metaphysical comes to return to the real world. This very same thematic exists in the biblical literature under the phenomenon of incarnation with its image of the word (a representation) becoming flesh. It is precisely this idea of incarnation that we have recently offered as a way out of the problematic presentation of the enactivist/embodiment literature (e.g., Roth, in press).

4. General discussion

There is a close relationship between the format in which researchers present the data (e.g., transcription) they extract from the data sources (e.g., videotape) and the theories they use to interpret or (try to) explain these data. Some data are such that they cannot be explained by particular theories. In such cases, researchers of the standard paradigm likely do not accept the data as valid, explain unwanted effects away, or introduce hidden variables to the theory (Kuhn, 1970). Here, we present the case of different forms of transcriptions that use classroom video as their source that researchers collect to develop their findings. Such transcriptions stand in a mutually constitutive relation with the claims that researchers (can) make. On one hand, the transcription is the source material from which claims are (inductively) developed. On the other hand, in research publications, the transcriptions function as evidence in support of the claims made.

In this study we show how different forms of transcription render visible different aspect of mathematical communication and therefore support different kinds of claims and the associated theories. We show that transcriptions that make use of words only and omit all information about the actual production of communication (*Take 1*) lend themselves to support constructivist arguments that make claims about stable knowledge (structures) in the mind somehow abstracted from the physical world. As soon as gestures and other perceptual aspects, for example, are rendered in terms of verbal descriptions, they no longer constitute embodied dimensions. Aspects of a situation produced and recognized by perceptual consciousness have been reduced to the verbal consciousness. Even talk about sensorimotor schemas does not get us any further because this talk is consistent with a Kantian position that makes mind a metaphysical entity – the embodiment theorist Johnson (1987) acknowledges having borrowed his conception of the schema from Kant – to the point that there is nothing outside (verbal) understanding (Henry, 2003). Because “the presuppositions of the Kantian ontology remain closed to the being of life” (p. 45), no constructivist account of knowing is able to capture the essence of embodiment/enactivist theory.

The preceding sort of claims are impossible if a researcher takes the stance that we present in *Take 2* as the production of communication that can no longer be reduced to individuals. The minimum unit of analysis is the turn pair, which means – consistent with a range of theories – that each word pertains both to the speaker and to the listener.

Moreover, in this second kind of approach the temporality of the production matters, because what is said at some time takes into account what has been said before but may be entirely inconsistent with what is said thereafter. The approach therefore is consistent, for example, with Vygotskian (1986) theory, which stipulates communicating and thinking to be continually changing processes. Any word uttered therefore no longer is the same when it is uttered again. Even an individual word repeated once or more no longer has the same function and therefore cannot be analyzed in terms of a constant sense or "meaning." In fact, researchers taking this stance no longer worry about "meaning" that somehow is indicated but not really present because the only thing that counts, consistent with Wittgenstein's position, is word-use and how consecutive speakers employ, re-employ, or change employment of words. Because temporality and time are important, this second approach much better than the preceding one can account for the continual changes that we observe in language and culture in a mathematics classroom over time, even though individual students and teachers do not think about or are conscious of such changes.

If it is the case that others are entrained into the collective pitch and bodily rhythms – as our example here shows consistent with other research (Auer & Couper-Kuhlen, 1994; Szczepek Reed, 2010) – then the production of the individual locution no longer is reducible to the speaker. Thus, more so than articulated in the context of the second case, where the word is a feature common to speakers and listeners, the *production* of the locution no longer is independent from other productions in the setting. Each locution then has to be theorized as an integral *part* of a more complex situation. This situation that cannot be reduced to its parts, for the parts are produced as a function of the whole, and this whole only exists in and through the production of the parts. In this manner, our work also suggests a link between the individual and the collective through completely embodied phenomena inaccessible by and irreducible to mental phenomena (mind). Other than articulated by the enactivist theorists, bodily phenomena are *collective* rather than the result of individual sensorimotor actions.

The most difficult phenomenon to explain with (radical, social) constructivist theories is real-time production of mathematical communication. This is so because there are aspects that are central to the sense that participants mark and re-mark in and through their communicative contributions but that have no place in mathematics in the form we can find articulated in a textbook. But Kant (1964) did realize that the separation between the purely mental and the purely bodily may be impossible. Thus, at the very end of his life he wrote an analysis of jokes where the intellectual recognition of the pun occurs at the same time and indistinguishable from laughter.[5] His explanation involves both: The tension within the set up of the joke that addresses the senses creates a disequilibrium of the innards, which, when released, creates laughter. The two aspects are irreducible because the mind does not need laughter because it could simply analyze the pun (and perhaps find nothing funny about the story). A "joke" that is not funny is not a joke and is not associated with laughter. We suggest that precisely the same irreducible aspects between the conceptual and the purely bodily come to be sensed and experienced when readers perform our transcription (score). This transcription then is nothing other than an *account*

[5] Actually, in the early part of his work, Kant (1960) thinks of wit only in mental terms and uses the example of the well-known mathematician and founder of the mathematics curriculum in Germany, Christopher Clavius, to suggest that someone can be intelligent but dull (no wit).

(recipe, plan; manual of instructions); the performance involves the actual mathematical *work*. After the fact the performance can be judged to be a more-or-less adequate rendering of the account/score/plan – much like we might judge a musical performance to be inconsistent with the score or the dish to be inconsistent with the recipe. As a result, knowing to perform what the transcription *refers* to, readers are enabled to *feel* the work of mathematics that leads participants in an episode to produce what we see. But the transcription itself does not get us to this feel. The purpose of the present article is precisely to provide "scores" for the performance of the mathematical knowing that researchers write about in their studies. This is especially important for those mathematics educators adhering to enactivist/embodiment theories, which require very different forms of data than the alternative constructivist-cognitive accounts.

Acknowledgments
This study was supported by two grants from the Social Sciences and Humanities Research Council of Canada. All opinions are those of the authors.

Disclosure Statement
There are no conflicts of interest of any kind.

References

Abraham, N. (1995). *Rhythms: On the work, translation, and psychoanalysis*. Stanford, CA: Stanford University Press.

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology, 35*, 127–145.

Alibali, M. W., Flevares, L. M., & Goldin-Meadow, S. (1997). Assessing knowledge conveyed in gesture: Do teachers have the upper hand? *Journal of Educational Psychology, 89*, 183–193.

Auer, P., & Couper-Kuhlen, E. (1994). Rhythmus und Tempo konversationeller Alltagssprache [Rhythm and speech in conversational everyday speech]. *Zeitschrift für Literaturwissenschaft und Linguistik, 24*, 78–106.

Austin, J. (1962). *How to do things with words*. Cambridge, MA: Harvard University Press.

Bakhtine, M. [Volochinov, V. N.] (1977). *Le marxisme et la philosophie du language: essai d'application de la méthode sociologique en linguistique*. Paris: Les Éditions de Minuit.

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

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Mathematical Thinking and Learning, 1,* 5–43.

Collins, R. (2004). *Interaction ritual chains*. Princeton, NJ: Princeton University Press.

Davis, B. (1995). Why teach mathematics? Mathematics education and enactivist theory. *For the Learning of Mathematics, 15* (2), 2–9.

Deleuze, G., & Guattari, F. (2005). *Qu'est-ce que la philosophie?* Paris: Les Éditions de Minuit. (First published in 1991)

Derrida, J. (1967) *De la grammatologie,* Paris: Les Éditions de Minuit.

Derrida, J. (1996). *Le monolinguisme de l'autre ou la prothèse d'origine* [Monolingualism of the other; or, the prosthesis of origin]. Paris: Galilée.

Dubinsky, E. (1999). Book review: Mathematical reasoning: Analogies, metaphors, and images. *Notices of the AMS, 46*, 555–559.

Edwards, D., & Potter, J. (1992). *Discursive psychology*. London: Sage.

Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Garfinkel, H. (1996). Ethnomethodology's program. *Social Psychology Quarterly, 59,* 5–21.

Garfinkel, H., & Sacks, H. (1986). On formal structures of practical action. In H. Garfinkel (Ed.), *Ethnomethodological studies of work* (pp. 160–193). London: Routledge & Kegan Paul.

Goldin, G. A. (2001). Counting on the metaphorical. *Nature, 413,* 18–19.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics, 32*, 1489–1522.

Goodwin, C., Goodwin, M. H., & Yaeger-Dror, M. (2002). Multi-modality in girls' game disputes. *Journal of Pragmatics 34*, 1621–1649.

Heidegger, M. (2000). *Vorträge und Aufsätze* [Presentations and essays]. Frankfurt: Vittorio Klostermann.

Henry, M. (2003). *De la phénoménologie. Tome 1: Phénoménologie de la vie*. Paris: Presses Universitaires de France.

Husserl, E. (1939). Die Frage nach dem Ursprung der Geometrie als intentional-historisches Problem. *Revue internationale de philosophie, 1,* 203–225.

Husserl, E. (1997). *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie: Eine Einleitung in die phänomenologische Philosophie* [The crisis of European sciences and transcendental phenomenology: An introduction to phenomenological philosophy]. Hamburg: Felix Meiner.

Johnson, M. (1987). The body in the mind: The bodily basis of imagination, reason, and meaning. Chicago: Chicago University Press.

Kant, I. (1960). *Werke Band I: Vorkritische Schriften bis 1768* [Works vol. 1: Pre-critical writings until 1768]. Wiesbaden: Insel.

Kant, I. (1964). *Werke Band VI: Schriften zur Anthropologie, Geschichtsphilosophie, Politik, und Pädagogik* [Works vol. 6: Writings on anthropology, philosophy of history, politics, and pedagogy]. Wiesbaden: Insel.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Lakoff, G., & Núñez, R. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being.* New York: Basic Books.

Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Paris: Gallimard.

Nancy, J-L. (2007). *Listening*. New York: Fordham University Press.

Núñez, R. E. (2009). Gesture, inscriptions, and abstraction: The embodied nature of mathematics or why mathematics education shouldn't leave the math untouched. In W.-M. Roth (Ed.), *Mathematical representation at the interface of body and culture* (pp. 313–332). Charlotte, NC: Information Age Publishing.

Núñez, R. E., Edwards, L. D., & Matos, J. F. (1999). Embodied cognition as grounding for situatedness and context in mathematics education. *Educational Studies in Mathematics, 39*, 45–65.

Ricœur, P. (1991). *From text to action: Essays in hermeneutics, II*. Evanston, IL: Northwestern University Press.

Roth, W.-M. (2000). From gesture to scientific language. *Journal of Pragmatics, 32*, 1683–1714.
Roth, W.-M. (2002). From action to discourse: the bridging function of gestures. *Journal of Cognitive Systems Research, 3*, 535–554.
Roth, W.-M. (2003). Gesture-speech phenomena, learning and development. *Educational Psychologist, 38*, 249–263.
Roth, W.-M. (2010a). Incarnation: Radicalizing the embodiment of mathematics. *For the Learning of Mathematics, 30* (2), 2–9.
Roth, W.-M. (2010b). *Language, learning, context: Talking the talk*. London, England: Routledge.
Roth, W.-M. (in press). *Mathematics in the flesh: The origins of geometry as objective science in elementary classrooms*. New York, NY: Routledge.
Roth, W.-M., & Lee, Y. J. (2007). "Vygotsky's neglected legacy": Cultural-historical activity theory. *Review of Educational Research, 77*, 186–232.
Roth, W.-M., Lee, Y. J., & Hwang, S.-W. (2008). Culturing conceptions: From first principles. *Cultural Studies of Science Education, 3*, 231–261.
Roth, W.-M., & Pozzer-Ardenghi, L. (2006). Tracking situated, distributed, and embodied communication in real time. In M. A. Vanchevsky (Ed.), *Focus on cognitive psychology research* (pp. 237–261). Hauppauge, NY: Nova Science.
Roth, W.-M., Tobin, K., Carambo, C., & Dalland, C. (2005). Coordination in coteaching: Producing alignment in real time. *Science Education, 89*, 675–702.
Selting, M., Auer, P., Barden, B., Bergmann, J., Couper-Kuhlen, E., Günthner, S., Meier, C., Quasthoff, U., Schlobinski, P., & Uhmann, S. (1998). Gesprächsanalytisches Transkriptionssystem [Conversation analytic transcription system]. *Linguistische Berichte, 173,* 91–122.
Suchman, L. A. (1987). Plans and situated actions: The problem of human-machine communication. Cambridge: Cambridge University Press.
Szczepek Reed, B. (2010. Speech rhythm across turn transitions in cross-cultural talk-in-interaction. *Journal of Pragmatics, 42,* 1037–1059.
Treitler, T. (1982). The Early History of Music Writing in the West. *Journal of the Americal Musicological Society, 35*(2), 237-279.
Turnbull, D. (1993). The ad hoc collective work of building gothic cathedrals with templates, string, and geometry. *Science, Technology, & Human Values, 18*, 315–340.
von Glasersfeld, E. (1991). Abstraction, re-presentation, and reflection: An interpretation of experience and of Piaget's approach. In L. P. Steffe (Ed.), *Epistemological foundations of mathematical experience* (pp. 45–67). New York: Springer.
Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press.
Wittgenstein, L. (1958). *Philosophical investigations* (3rd ed.). New York: Macmillan.

Appendix A

Typical conventions used for transcriptions such as those presented in Take 2

Notation	Description	Example
(0.14)	Time without talk, in seconds	`more ideas. (1.03) just`
()	Pause of less than 0.10 seconds;	`kay. () bert`
((turns))	Verbs and descriptions in double parentheses are transcriber's comments	`((nods to Connor))`
::	Colons indicate lengthening of phoneme, about 1/01 of a second per colon	`si::ze`
[ ]	Square brackets in consecutive lines indicate overlap	`S: s[ize ]` `T: [colby]`
this one	Underlined part coordinates with a gesture described; lH and rH indicating left and right hand, respectively	`this ones:? ((rH moves down, up, down right face, Fig 4.1b))`
((:B))	Colon prior to letter in double parentheses: The speaker directly addresses another person "B"	`57 T: ((:B)) hOW did`
<<p> >	Piano, words are uttered with less than normal speech volume	`<<p>um>`
<<pp> >	Pianissimo, words are uttered with very low, almost inaudible volume	`<<pp>this>`
<<f> >	Forte, words are uttered with greater than normal speech volume	`<<f>that> makes`
<<ff> >	Fortissimo, much louder than normal speech volume	`<<ff>hU::::::ge.>`
<<all> >	Allegro, faster than normal speech rate	`<<all>[whawould]> that`
<<len> >	Lento, slower than normal speech rate	`<<len, drawn out>but () its like a flA:Tcube.>`
<<confidently> >	Ethnographic description of speech that is enclosed in brackets	`<<confidently>because its like a sort of (0.60) vertex>`
ONE bert	Capital letters indicate louder than normal talk indicated in small letters.	`no? okay, next ONE bert.`
.h, hh	Period before "h" indicates in-breath; "h" without period is out-breath	`.hhi, hh hh`
(?cular)	Question mark with whole or part word in parentheses indicate	`(serial?), (?cular)`

	possible hearings of words or missing sound	
(??)	Question mark(s) in parentheses: Inaudible word(s), the approximate number given by number of marks	`i (??)`
,?;.	Punctuation is used to mark movement of pitch toward end of utterance, flat, slightly and strongly upward, and slightly and strongly downward, respectively	`T: so can we tell a shape by its color?` `T: does it ‘belong to another ‘group (0.67) O:r.`
=	Phonemes of different words are not clearly separated	`loo::ks=similar`
↑↓	Arrow up, down: Significant jump in pitch up or down	`is ↑sort, ↓<<all>so thats`
`‘^ˇ	Diacritics indicate movement of pitch within the word that follows—down, up, up-down, and down-up, respectively	`` `um; ‘sai:d; ^Cheyenne; ˇsquare ``