

Multimodal Revoicing as an Interactional Mechanism for Connecting Scientific and Everyday Concepts

Virginia J. Flood

University of California, Berkeley, CA, USA

Keywords

Multimodal revoicing · Scientific concepts · Everyday concepts · Interactional mechanisms · Ethnomethodological conversation analysis

Abstract

A perpetual problem learners face is identifying which aspects of embodied experiences are relevant for appreciating the world in culturally specific ways. Vygotsky argued that social interactions with more competent cultural members provide arenas for linking *everyday* and *scientific concepts*. However, the precise interactional mechanisms of how these linkages are forged remain underexamined. I argue that understanding these mechanisms requires examining how intersubjectivity is built and maintained. I propose that ethnomethodological conversation analysis and the co-operative action framework provide a uniquely suited analytic orientation for this project because they focus on the fine details of the *actual* practical methods people use to procedurally achieve intersubjectivity. To illustrate the utility of these approaches, I show how the fine details of *multimodal revoicing interactions* present semiotic challenges that allow learners to link everyday and scientific concepts. Specifically, I examine the role *dialogic gesture* plays in reformulating a multimodally expressed idea about what it means to “go faster.”

© 2018 S. Karger AG, Basel

It is now widely accepted that scientific and mathematical ways of thinking grow out of our embodied experiences in the world [e.g., Abrahamson & Lindgren, 2014; Goodwin, 2018; Saxe, 2012]. However, the social interactional mechanisms through which learners come to organize the particularities of their embodied experiences (e.g., a feeling of going faster) with standard, generalized formalisms of disciplinary mathematics and science (e.g., faster as more distance traveled in the same period of time) are still largely unexamined [Saxe, De Kirby, Kang, Le, & Schneider, 2015].

In this paper, I explore how *dialogic gesture* [Arnold, 2012] between adults and children serves as an important interactional mechanism for reconciling embodied experiences with culturally specified ways of organizing those experiences. Drawing on Wertsch's and Koschmann's elaborations of Vygotsky and Bakhtin, I propose that to understand this process, we have to better understand the basic *dialogic* mechanisms people have for building and maintaining intersubjectivity together. I argue that ethnomethodological conversation analysis and Goodwin's co-operative action framework [2013, 2018] provide a uniquely suited analytic orientation for uncovering these mechanisms through the discovery of the fine details of the *actual* practical methods people use to procedurally achieve intersubjectivity. To advance this argument, I demonstrate how revoicing, when treated as a multimodal form of co-operative action, can take a variety of different forms that pose *semiotic challenges* [Wertsch, 1985], inviting children to organize the world in culturally specific ways.

Vygotsky's Spontaneous and Scientific Concepts and Wertsch's Semiotic Challenges

For Vygotsky, the process of reconciling subjective experiences with culturally sanctioned methods for organizing those experiences was fundamental to human development [Vygotsky, 1986]. Vygotsky saw this process as a *bidirectional* integration of two distinct forms of awareness – *spontaneous* and *scientific* concepts – which reciprocally shape one another throughout a person's lifetime. Spontaneous concepts describe the “everyday,” natural, or unschooled ways of organizing and interpreting experiences. Scientific concepts, on the other hand, are formally specified, sanctioned ways of organizing the world that belong to particular communities or disciplines and must be handed down through cultural history from generation to generation.

Vygotsky never developed specific, mechanistic accounts of precisely *how* the bidirectional linkage he described takes place [Stone, 1993; Wertsch, 1985]. However, he did specify an arena for this process: The two forms of awareness can only be brought together through social encounters with more culturally competent members of society. Sociocultural studies of cognitive development have since produced numerous accounts of how newcomers (e.g., children, apprentices, students, etc.) gradually come to engage in a community's valued practices and adopt its cultural forms (e.g., counting systems [Saxe, 2012]) through social participation with old-timers (e.g., caregivers, instructors, more experienced practitioners, etc. [Lave & Wenger, 1991; Newman, Griffin, & Cole, 1989; Rogoff, 1990; Saxe, 2012]).

Much of this work characterizes broad patterns of reoccurring interaction, activity, and features of learning communities that give rise to development, especially over diachronic scales of time. For example, Saxe et al. [2015] provide a longitudinal illustration of how reoccurring, routinized collective classroom practices like “correcting the teacher” provide repeated opportunities for students to connect embodied experiences with cultural formalisms over time. However, they argue that there is much left to uncover about how social interactions between adults and children can create ecologies for scientific and spontaneous concepts to shape one another. Saxe et al. [2015] join a growing group of scholars who argue that more microanalytic investigations of the fine details of social interactions are necessary to understand this process [e.g., Koschmann & Zemel, 2014; Macbeth, 2011; Roth & Thom, 2009].

Wertsch [1980, 1985] long ago argued for the importance of microanalytic investigations of the interactions between adults and children. His work contributes some of the most important elaborations of Vygotsky's scholarship [Saxe, 2008] and draws on Bakhtin and Rommetveit to extend Vygotsky's claims about the role of social interaction in development. In particular, Wertsch proposed that the key to understanding how children adopt culturally appropriate perspectives is a better understanding of the ordinary (and not necessarily pedagogical) mechanisms for achieving and maintaining intersubjectivity in interaction [Wertsch, 1985]. He argued that understanding these processes requires revealing and understanding the precise *communicational mechanisms* that adults and children use to establish the mutual intelligibility necessary for coordinating activity together when they approach situations with distinct perspectives.

Following Rommetveit, Wertsch argued that "temporarily shared social realities" are "established and continually modified by acts of communication" [Rommetveit, 1979, as quoted by Wertsch, 1985, p. 160]. In any situation within the social and material ecologies people find themselves inhabiting together, there are innumerable potentially relevant features and ways of interpreting objects and events. Through dialogue, adults and children negotiate, "create and foreground perspectives on reality" [Wertsch, 1980, p. 150]. In particular, they are "constantly defining and redefining the *relevant* aspects of reality" [Wertsch, 1980, p. 161; my italics] for understanding the situation at hand from a more culturally competent perspective. Particular organizational features of interactional sequences – e.g., modulating ways of referring to objects – pose what Wertsch calls *semiotic challenges* that "lure" children into culturally sanctioned ways of organizing their experiences (i.e., adopting culturally appropriate definitions of situations [Wertsch, 1985, p. 176]).

To illustrate how modulating references to objects poses a semiotic challenge, Wertsch describes the case of a mother and her young child working together to assemble a truck puzzle from a collection of wooden pieces. Mother and child demonstrate different interpretations of what they are doing to one another: The child invites his mother to see the wood pieces as "crackers," while his mother points and refers to them by their intended role within the truck puzzle (e.g., disk-shaped pieces can be seen as "wheels"). The negotiation of how to refer to the wood pieces creates a semiotic challenge that provides an opportunity for the child to see these objects in a culturally specific way: how they are meant to be seen *as pieces* for assembling the truck puzzle [Wertsch, 1985]. The child is thus invited to perceive these objects from the adult's perspective as part of a cultural form (a truck puzzle, not make-believe food) as they build the puzzle together.

Stone [1993] joined Wertsch in advocating for a more careful articulation of the precise communicative mechanisms involved in adult-child interactions while critiquing the popular oversimplified metaphor of scaffolding. He described several overlooked interactional mechanisms in ordinary conversation that have the potential to pose semiotic challenges. One of these mechanisms is prolepsis, first documented by Rommetveit as a general conversational practice where listeners, when faced with an initial underspecified utterance, are able to infer the perspective of the speaker by seeking out the meaning of the original utterance in the subsequent, unfolding interaction. Stone noted that when adults and children are engaged in a task, an adult may give a child a directive that presupposes the child has the same perspective of the task goals and methods for achieving those goals as the adult. If the child

does not follow through, the adult may further specify the nature of the directive, breaking it down into smaller components. Through this juxtaposition of a vague directive and subsequent specific instructions for accomplishing the task, the child can begin to appreciate the sense of the initial directive from the adult's perspective on the task.

Wertsch's and Stone's classic arguments strongly motivate the need to better understand the vast array of mundane interactional work participants do to build intersubjectivity in order to understand how scientific and spontaneous concepts become linked in social interaction. Erickson – another early advocate for the microanalyses of interactional processes in learning and development – argues that only a very small amount of the interactional mechanisms involved in learning, teaching, and development has been characterized [Erickson, 2011]. Echoing an argument made by Schegloff [1996] about interaction, Lindwall and Lymer [2008] also lament that the bulk of what is going on during the mundane, not easily classifiable social activity of instructional interactions has been left largely unexamined. It remains like 90% of the physical matter in the universe – dark and unaccounted for. Ethnomethodological conversation analysis and co-operative action offer a critical framework for illuminating this dark matter.

Ethnomethodological Conversation Analysis and Co-Operative Action

The central concern of ethnomethodological conversation analysis (EMCA) is uncovering the fine details of participants' actual, in situ methods for building and maintaining intersubjectivity [Heritage, 1984; Mondada, 2012]. As such, EMCA provides a powerful analytic orientation for Wertsch's and Stone's specifications for how to investigate the bidirectional linkage of scientific and spontaneous concepts in social interaction.

One of EMCA's greatest contributions has been revealing the architecture of producing intersubjectivity that is built into how we conduct interaction together [Heritage, 1984; Schegloff, 1991]. In focused encounters [Goffman, 2013], we carefully monitor one another and shape our contributions accordingly. Every utterance or action we make provides a witnessable display of our current analysis of what we take to be going on and what we project is relevant to happen next. We design our own utterances or actions in response to our interlocutor's displayed analyses that we have witnessed in their prior or ongoing utterances. In this way, the mutual intelligibility of situations is an ongoing achievement that unfolds through time, moment by moment, built both simultaneously and sequentially with the local resources at hand (e.g., turns at talk, spatial arrangements, materials, and bodies – including gestures, posture, and facial expressions) through a relentless process of incremental, public displays of revisions and ratifications of meaning [Garfinkel, 1967; Sacks, Schegloff, & Jefferson, 1974; Streeck, Goodwin, & LeBaron, 2011].

While EMCA has outlined this architecture as the general machinery for achieving intersubjectivity in interaction, EMCA's true project lies in discovering and characterizing the fine particulars of how this machinery is embodied in *actual circumstances*: How are situated, interactional resources mobilized to create scenic displays of the sense *these participants* are making of *this situation* in order to coordinate *this specific* course of action together? As Garfinkel, the founder of ethnomethodology,

argued again and again, these particulars cannot be imagined or hypothesized in the abstract, they can only be discovered “just-in-any-actual case” [Garfinkel, 1996, p. 7] of interaction (e.g., in the case of the mother and her child assembling the truck puzzle). Uncovering the specific resources and precise procedures that allow us to engage in mutually intelligible courses of action together makes it possible to understand how intersubjectivity is achieved and maintained in any actual case.

In particular, the extensive work of Koschmann and colleagues in teaching hospitals demonstrates that EMCA is a powerful tool for investigating the procedural, praxeological nature of intersubjectivity in instructional settings. By examining the fine details of participants’ interactional work to coordinate their activity together, Koschmann and colleagues show how occasions are produced where less experienced practitioners are able to develop understandings of the world that begin to more closely resemble their mentors’ [Koschmann, LeBaron, Goodwin, & Feltovich, 2011; Koschmann, Lebaron, Goodwin, Zemel, & Dunnington, 2007; Koschmann & Zemel, 2014; Zemel & Koschmann, 2014].

Goodwin’s co-operative action framework [Goodwin, 2013, 2018] elaborates EMCA and posits that in order to coordinate *any* concerted activity together in the world (e.g., performing a surgical procedure, assembling a puzzle), participants build new actions and utterances by recycling and reassembling elements of each other’s prior actions and utterances. With respect to learning, Goodwin argues that “the process through which skilled, knowing members are shaped within communities... is frequently accomplished through *the more general practices* participants use to build *co-operative action* in concert with each other” [Goodwin, 2018, p. 12; my italics]. In any interaction, participants work to make *how they are understanding the world in front of them* available to one another by mobilizing a diverse array of mutually elaborating multimodal semiotic resources (e.g., gesture, demonstrative action with tools, gaze, prosody, talk, etc. [Goodwin, 2000]). Each new, subsequent action is created by decomposing and reusing with transformation the public semiotic resources made available through prior (and sometimes simultaneous) actions. Goodwin argues that through these general practices, “pedagogy... emerges as a systematic consequence of the organization of co-operative action” [Goodwin, 2018, p. 12].

Goodwin, like Wertsch and Stone, argues for the investigation of the ordinary interactional mechanisms of building intersubjectivity to understand instruction in the wild. Goodwin’s co-operative action framework extends Wertsch’s and Stone’s work by further illuminating how newcomers become knowledgeable, competent, and skilled members of their communities: In his investigations of senior and junior scientists working together in the field (geologists [Goodwin, 2018]; geochemists [Goodwin, 1997], and archeologists [e.g., Goodwin, 1994]), Goodwin has uncovered a number of specific interactional mechanisms through which participants coordinate their attention to consequential structures in the environment, generating opportunities for an “education of attention” [Gibson, 1979].

Goodwin argues that educating the attention of newcomers involves (a) positioning the environment for relevant forms of perception and action, (b) highlighting relevant features of newcomers’ embodied experiences in the environment for community-valued ways of perceiving the world, and (c) differentiating which aspects of sensuous experiences in the world are relevant for making certain typifications (i.e., the work of classifying and coding – determining “what kind of thing is this?” [Goodwin, 2018, p. 361] – which involves determining which features of sensuous experi-

ence – qualia – make something count as a token of an abstract type or category [Goodwin, 1994, 1997, 2007, 2018]). Each of these activities is an interactional accomplishment. While working together in face-to-face arrangements to carry out a community’s consequential work (e.g., identifying patterns in the soil as evidence of ancient buildings [Goodwin, 2007]; performing a chemical reaction [Goodwin, 1997], or scrutinizing a rock to decide whether more samples should be taken from the area [Goodwin, 2018]), newcomers and more senior members closely monitor both each other and the environment they act in (e.g., a patch of soil, a vat of a chemical reaction in a lab, a rock selected from a mountainside). Through this mutual monitoring, seniors can attend to newcomers’ public displays of attention as they participate in activities together and can work to finely calibrate newcomers’ attention to just what is relevant in that moment for the task at hand [Goodwin, 1994, 1997, 2007, 2018]. These calibrations lead to progressive refinements in meaning, allowing newcomers to reconcile their experiences with culturally specified ways of acting in and perceiving the world that are necessary for being seen as a competent member of that community. This is only possible in the moment through interactions that emerge from organic efforts to build and maintain intersubjectivity. Thus, Goodwin views this interactional process of reconciling spontaneous and scientific concepts as a “simultaneous pedagogy” [Goodwin, 2018, p. 317].

Underlying Wertsch’s and Stone’s classic elaborations of Vygotsky, as well as Koschmann’s and Goodwin’s contemporary investigations of instruction, is what Koschmann [1999] calls a *dialogic theory of learning*. From this perspective, learning is understood to be a dynamic process where new meanings and temporarily shared realities are co-constructed and emerge moment by moment through ongoing courses of co-operative action. Thus, new meanings are *dialogically* organized through the process of different voices coming into contact and reciprocally shaping – or interanimating – each other [Bakhtin, 1981]. Adopting this Bakhtinian perspective on meaning-making, learning becomes a near constant, dynamic, transactional process that emerges from the mundane methods and resources participants use to interactionally achieve intersubjectivity in the moment.

In particular, we can see how in Goodwin’s investigations, Wertsch’s semiotic challenges arise through co-operative action: Through calibration and progressive refinements, newcomers are afforded opportunities to reconcile their embodied experiences (e.g., perceiving a variegated colored patch of soil) with a community’s consequential ways of organizing those experiences (e.g., recognizing a specific part of the patch as what an archeologist would call a “plough scar”). This calibrating of perception and action is always an interactional achievement, highly contingent on the unfolding task and the displays of understanding the newcomer makes available in the moment. By exploring how instruction unfolds through the fine details of co-operative action, we can shed light on the missing interactional mechanisms for Vygotsky’s bi-directional linking of spontaneous and scientific concepts.

Synthesizing these dialogical approaches to understanding learning, I propose that to identify how semiotic challenges arise, it is crucial to focus our attention on the mundane, fine details of the interactional work participants do to shape each other’s perception and action in the world. Careful examinations of the unfolding, moment-by-moment structure of interactional sequences in face-to-face encounters between more experienced cultural members and less experienced members inspired by EMCA and co-operative action are essential for this project.

Revoicing Interactions as Fertile Territory for Investigating Semiotic Challenges

To demonstrate the utility of EMCA and the co-operative action framework for illuminating how semiotic challenges emerge, I analyze a familiar instructional sequence, the *revoicing interaction*. I show how the unexamined and hidden complexities of revoicing interactions – which only come into relief when we examine how they unfold in their situated detail – provide opportunities for newcomers to connect spontaneous and scientific concepts.

First discovered by O'Connor and Michaels [1993, 1996], revoicing describes any combination of *repeating* (reporting or restating verbatim), *reformulating* (modifying content of), and/or *elaborating* (adding new content to) learners' contributions during interactions. Revoicing occurs in many educational contexts (from kindergarten [Yifat & Zadunaisky-Ehrlich, 2008] to undergraduates [Flood et al., 2015]) and across a wide variety of subjects (from mathematics [Krussel & Edwards, 2004] and science [e.g., Ruiz-Primo & Furtak, 2007] to second-language learning [e.g., Park, 2013] and liberal arts seminars [Parsons, 2017]). Scholars argue that revoicing is pedagogically advantageous because it can (a) promote deeper full-class exploration of student-generated ideas [Forman & Ansell, 2002], (b) highlight particular elements of student ideas while backgrounding other elements [Nam, Ju, Rasmussen, Marongelle, & Park, 2008], (c) extend and reshape the content of student contributions to resemble disciplinarily normative concepts [Eckert & Nilsson, 2017], (d) help students adopt disciplinarily normative language and representations [Forman & Larreamendy-Joerns, 1998], and (e) promote participation by explicitly valuing and soliciting student contributions [Strom, Kemeny, Lehrer, & Forman, 2001].

Forman and Larreamendy-Joerns [1998] were first to demonstrate one way revoicing interactions can provide a means for students to connect spontaneous with scientific concepts about argumentation. Drawing on Grice's [1989] work on conversational implicatures, they show how everyday ideas about sufficient explanation or argumentation are vastly different than what is considered a sufficient *mathematical* or *scientific* argument. In the revoicing interactions they examined, a teacher reformulated students' arguments by replacing indexical terms with specialized vocabulary and adding clearer articulations of premises to the initial contributions. Forman and Larreamendy-Joerns proposed that these reformulations function as meta-messages for students, helping students recognize the need to be more explicit when making mathematical arguments, even though this hyper-explicitness violates normal conversational expectations [Grice, 1989]. After their arguments were revoiced, students designed subsequent contributions to more closely resemble the style of formal mathematical arguments modeled in the teacher's revoicing turn.

Forman and Larreamendy-Joern's investigation shows how revoicing interactions can create semiotic challenges. Revoicing interactions are a clear case of co-operative action: semiotic resources are recycled and transformed to build new contributions and provide opportunities for students to reconcile everyday understandings with the specialized practices and ways of thinking that belong to specific cultures. Thus, the mundane interactional resources and methods that accomplish revoicing interactions warrant closer investigation.

From the perspective of EMCA and the co-operative action framework, however, the investigations reviewed above are significantly limited because they ap-

proach revoicing as a strictly *verbal phenomenon*. By focusing on talk alone, the role of the body and the material environment in revoicing is left unexamined. EMCA and co-operative action studies have established that participants constantly mobilize a variety of multimodal semiotic resources in conjunction with verbal content to coordinate concerted courses of intelligible action together, including prosody, gesture, body posture, manipulating materials, and gaze [Goodwin, 2000; Luff & Heath, 2015; Mondada, 2016; Streeck et al., 2011].

Past studies of revoicing tend to reduce interactions to participants' talk alone. The role that bodily communicative resources and material features of the environment play in shaping these interactions is entirely ignored or only sporadically included. This omission reflects a general theme in investigations of interactional sequences in classroom discourse, tutoring, and dyadic caregiver-child interactions. Without considering embodied action in the material environment, many resources that participants use while working to achieve intersubjectivity are not available for analysis and remain unexplored. Following McDermott, Gospodinoff and Aron [1978], we need to create more "ethnographically adequate" accounts of revoicing and other familiar instructional interactional sequences. Such accounts seek to recover and preserve the role *all* semiotic resources play in how participants achieve "temporarily shared social realities." To better understand how semiotic challenges arise through co-operative action in familiar interactional sequences like revoicing, we must reexamine these sequences with an eye towards ethnographic adequacy.

Towards an Investigation of *Dialogic Gesture* in Teaching and Learning

A growing body of literature in STEM (science, technology, engineering, and mathematics) education illuminates how embodied and material resources play a critical role in social meaning-making practices. Learners use embodied semiotic resources to share, negotiate, and refine emerging ideas with peers and teachers [Crowder, 1996; Flood et al., 2015; Kastens, Agrawal, & Liben, 2008; Nemirovsky & Ferrara, 2009; Wittmann, Flood, & Black, 2013]. The body becomes a particularly valuable resource for imagining together [Flood, Harrer, & Abrahamson, 2016; Nemirovsky, Kelton, & Rhodehamel, 2012] when students do not yet share conventionalized vocabularies, symbols, or diagrams for articulating emergent ideas [Roth & Lawless, 2002; Roth & Welzel, 2001]. Radford [e.g., 2003, 2009], in particular, has demonstrated that gestures are key for achieving temporary yet durable stabilizations of meaning as part of the process of objectifying and generalizing emergent mathematical ideas. Researchers have also documented numerous ways in which instructors use gestures and other forms of embodied, multimodal semiotic resources. For example, mathematics teachers create visible links between diagrams and other inscriptions with gesture [Alibali et al., 2014] and repeat their own gestures [McNeill, 2000] to build coherence through long explanations [Nathan & Alibali, 2011]. In science, instructors use gestures and bodily performances to animate and highlight features of inscriptions and models while explaining complex scientific phenomena [e.g., Kress, Jewitt, Ogborn, & Tsatsarelis, 2001; Pozzer-Ardenghi & Roth, 2007].

While the investigations reviewed above illustrate how teachers and learners use a diverse array of multimodal semiotic resources in instructional settings, the bulk of this scholarship focuses on how *individual utterances in isolation* convey teacher-

generated insight or *student-generated insight*. The ways in which instructors' and learners' multimodal semiotic resources come into contact and interanimate one another – that is, how they are *dialogic* [Bakhtin, 1981; Arnold, 2012] – has received little attention in educational studies. Following Goodwin, we must explore how instructors and students reuse parts of multimodal utterances and actions and transform them into something new. These forms of multimodal dialogicity in instructional interactional sequences require further examination to appreciate how semiotic challenges arise to provide opportunities for linking spontaneous and scientific concepts.

Studies of situated, embodied social interaction in a variety of settings – from family dinners to scientific laboratories – suggest that participants are responsive to multimodal semiotic components in one another's utterances and regularly use them to build their own utterances. For example, participants recognizably repeat gestures produced by antecedent speakers to create visible demonstrations that they are tracking and grasping prior utterances (named a "return gesture" by De Fornel [1992]). Lerner [2002] also documented how participants exhibit recognition and understanding of speakers' ongoing stories by reproducing embodied depictions that match the original story tellers'. Kimbara [2006] coined the term *gesture mimicking* to describe this practice. Gesture mimicking is not always affiliative: Repeating another's gestures can be used to intentionally mock someone [e.g., Goodwin & Alim, 2010]. Gesture mimicking is also found in professional settings to build cohesion across complex, collaboratively negotiated ideas [Wardak, 2017] (e.g., scientists proposing and building theories together [Becvar, Hollan, & Hutchins, 2005] or confirming they are interpreting physical phenomena in the same way [Goodwin, 2018]).

Gesture mimicking is also found in a wide variety of instructional settings. Medical students repeat each other's gestures to build collaborative explanations [Koschmann & LeBaron, 2002]. Students also repeat each other's gestures to build mathematical arguments together [Rasmussen, Stephan, & Allen, 2004] and develop shared representations like maps [Enyedy, 2005]. In addition, students also mimic teacher gestures: LeBaron and Streeck [2000] found that students copied their instructor's gestures in discussions of building features to demonstrate their understanding. Arnold [2012], studying bicycle repair workshops, documented a similar phenomenon where trainees recognizably reproduced gestures to match their trainers' gestures during explanations of repair technique. She argued that gesture mimicking is a way in which embodied action and gesture are dialogic, coining the terms *dialogic embodied action* and *dialogic gesture*.

Instructors also mimic students' gestures and embodied actions as part of revoicing. This practice serves a variety of different purposes in different settings. For example, dance instructors use *bodily quotes* – full-body reenactments – of students' dance moves to correct and adjust students' technique [Keevallik, 2010]. Instructors also repeat students' gestures to check their understanding of their students' ideas in the moment [Flood et al., 2015, 2016].

Another important purpose of teachers' gesture mimicking is to make connections salient between embodied action and specialized terminology. In second-language classrooms, Majlesi [2015] found that instructors repeat students' gestured actions to connect new or missing vocabulary terms to the actions they depicted. Shein [2012] discovered a similar phenomenon in a geometry class for second language learners and was the first to label this practice *multimodal revoicing*. She found that

Table 1. A 4×4 matrix of possible forms of multimodal revoicing

	Repeat gesture	Omit gesture	Elaborate gesture	Modify gesture
Repeat talk				
Omit talk				
Elaborate talk				
Modify talk				

when students used gestures while explaining their mathematical procedures, their teacher repeated their gestures and supplied mathematically normative verbal descriptors for the items they depicted. Outside of second-language learning contexts, Arzarello, Paola, Robutti, and Sabena [2008] have also described how mathematics instructors recognizably reproduce students' gestures to affix them with a more conventional mathematical term for the depicted phenomenon.

In the form of multimodal revoicing that previous authors have described, teachers "translate" students' gestures into culturally standard terms by repeating the gesture and elaborating the student's original verbal description. When the normative vocabulary for the gesture depicted is supplied, students are afforded an opportunity to make connections between the new term and their original idea via the visible bridge supplied by the repeated gesture [Shein, 2012; Majlesi, 2015].

Towards Understanding New Forms of Dialogicity in Multimodal Revoicing

In each of the previously observed cases of dialogic embodied action described above, a first participant's gesture or embodied action was reproduced essentially intact and without significant variation by other participants. That is, the reproduced gestures were designed to be seen as *identical* to a preceding gesture in efforts to either build cohesion across an explanation, demonstrate understanding, affix new terminology to, or mock an action. However, the co-operative action framework suggests that there are likely many other subtle ways in which participants might decompose, reuse, and transform embodied, semiotic resources.

Because there are multiple possible forms that dialogic embodied action can take, the pedagogical multimodal revoicing interactions Shein, Arzarello, and others have described represent just *one* among many possible forms of multimodal revoicing: A student's gesture is repeated, and the student's original verbal description is elaborated by adding new terminology. We could call this form of multimodal revoicing *gestural repetition co-timed with elaborated verbal content*. However, if we consider that the content of gesture and talk in multimodal revoicing may each vary along four dimensions, i.e. (a) repeated verbatim, (b) omitted totally, (c) elaborated, or (d) selectively modified, we could imagine at least 15 possible new forms of multimodal revoicing emerging in a 4×4 matrix (Table 1). However, even this is just a coarse ap-

proximation of possibilities because there are many ways to decompose, modify, and reformulate a single gesture, and each deserves consideration as I discuss below.

When students share scientific and mathematical ideas, they often render them through elaborate choreographies of semiotic resources [Goodwin, 2018; Flood et al., 2016; Nemirovsky & Ferrara, 2009; Radford, 2009]. Gestures, in particular, can be highly complex and often contain several distinct segments of movement, called *phrases* [Kendon, 2004] or *strokes* [McNeill, 1992]. During each gesture phrase, the hands can transform into different shapes, adopt various orientations (e.g., palm down, palm up), be held at different distances from the body, incorporate objects, follow different trajectories through space, move at various speeds, or even momentarily freeze [Streeck, 2009].

Because gestures can be so complex, there are a variety of ways that they could be recycled and transformed by others. A multi-phrased gesture can be reformulated by leaving out certain phrases or elaborated by adding new phrases. Individual phrases can be modified by changing the hand shape, altering the trajectory, or modifying the delivery speed. On top of this, any number of repeated, omitted, modified, or elaborated verbal phrases can be coordinated with any number of repeated, omitted, modified or elaborated gesture phrases. When we consider all of the intricacies of the embodied choreographies learners and instructors have at their disposal in revoicing interactions, we are confronted with an uncharted wealth of possibilities. A variety of new forms of multimodal revoicing are possible, whose affordances in pedagogical interactions are as yet unknown. In particular, how these forms of multimodal revoicing might be able to pose semiotic challenges for students by creating opportunities for them to link spontaneous ideas with scientific ideas is open for investigation.

Connecting Spontaneous and Scientific Concepts through Multimodal Revoicing

In the remainder of this paper, I will present an example to illustrate how analyzing a familiar type of instructional sequence using EMCA and the co-operative action framework can reveal how a semiotic challenge arises through the dialogicity of multimodal semiotic resources.

The episode is drawn from a video corpus of 23 task-based teaching interviews during which middle school students and adult tutors work with a technology-enabled embodied learning device for mathematics, called the Mathematical Imagery Trainer for Proportions (MIT-P [Abrahamson, Gutiérrez, Charoenying, Negrete, & Bumbacher, 2012]). The MIT-P was designed to provide an interactive context for users to explore ideas about proportionality through sensorimotor activity. Users operate two independent, hand-held remotes to manipulate the heights of two corresponding cursors on a computer screen (Fig. 1). The system generates feedback by turning the screen green when the cursor heights correspond to a particular preset, concealed ratio (e.g., 1:2 in Fig. 1). Whenever the cursor heights do not fulfill the secret ratio, the screen is red.

As they explore different methods for making the screen green, children are asked to share different patterns they notice and articulate different rules to explain their strategies. Tutors introduce a Cartesian grid and then numerals to label that

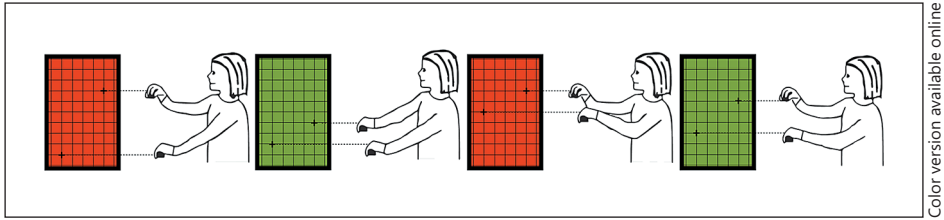


Fig. 1. When the Mathematical Imagery Trainer for Proportion (MIT-P) is set to a 1:2 ratio, the screen is green only when the right hand remote is twice as high as the left hand remote; otherwise it is red.

grid. Students are encouraged to use these tools to generate more sophisticated quantitative models that account for the behavior of the device [Abrahamson et al., 2012].

The MIT-P constitutes a complex, dynamic perceptual field of potentially relevant phenomena. Users discover many perceptual patterns (e.g., haptic, visual) in this phenomenal field of potential relevancies, but a key challenge is getting others – either peers or adult tutors – to perceive these collections of features as significant patterns, as well. While working with the device, children share a variety of interesting discoveries and tackle many productive problems. For example, some children notice that there is a relationship between iterative addition and multiplication [Abrahamson, Lee, Negrete, & Gutiérrez, 2014]. Other children make discoveries related to the idea of covariation [Flood et al., 2016]. One problem related to covariation that frequently comes up is explaining why the two cursors move at different speeds to maintain a green screen. This diverse array of topics makes this video corpus a perspicuous setting [Garfinkel, 2002] for investigating interactional work between adults and children where cultural forms (e.g., scientific concepts like mathematical definitions) are connected with phenomenological experiences (e.g., everyday concepts like the sensation of speediness).

Methods for Examining Multimodal Revoicing Interactions

To study revoicing interactions with EMCA and the co-operative action framework, I used the Max Planck Institute’s video annotation software ELAN [Lausberg & Sloetjes, 2009] to elucidate the fine details of tutor and students’ practical interactional methods and the resources they use to build action together. A highly refined level of granularity is possible with ELAN. Short sequences of video are annotated across multiple “timeline tracks” to record how actions and utterances unfold frame by frame, through time. Each utterance is decomposed into a set of separate tracks for each parallel-unfolding semiotic resource. A track is made for talk, gaze, gesture, body posture, writing, recruitment of tools or object manipulation, and any other resources that emerge. A timeline track for a semiotic resource is then broken down into events over time. For example, a complex gesture can be separated into distinct phrases, hand shapes, hand orientations, and trajectories. Both simultaneously occurring and sequentially occurring relationships between different semiotic resources and their components can be recovered. For example, the gesture track might trace the

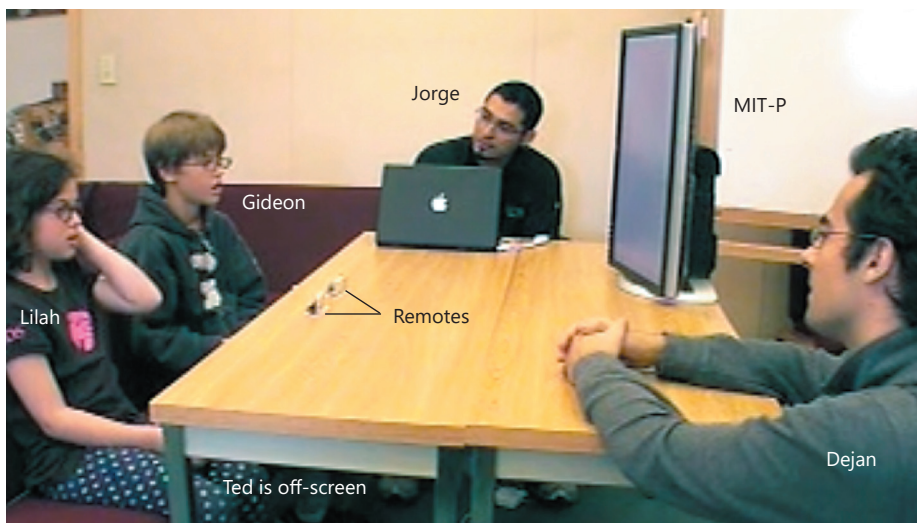


Fig. 2. Lilah and Gideon take a break from operating the MIT-P to discuss their strategy with tutors Jorge, Dejan, and Ted who is, at the moment, off-screen.

distinct phrases and dynamically evolving hand shapes of a participant’s gesture over time. This track can then be compared vertically with other tracks to examine, for example, the precise temporal coordination of the gestural components and talk within this utterance. In addition, prosody was analyzed using Praat (version 6.0.37).

Decomposing revoicing interactions with ELAN was what initially enabled my discovery of the nuanced ways in which multimodal revoicing can take place. The rich variations and organization of these revoicing interactions, which had originally all been classified as “echoing” student ideas in a previous qualitative coding scheme [Abrahamson et al., 2012], only became apparent when annotated in fine detail.

Connecting the Embodied Experience of Faster with a Scientific Description of Faster

The revoicing episode I examine occurs during a discussion about speed while a pair of fifth-graders – Lilah and Gideon – and three tutors – Ted, Dejan, and Jorge – are working with the MIT-P (Fig. 2; see Appendix 2 for a full transcript). Lilah and Gideon have been exploring and experimenting with the MIT-P for about 30 minutes. They have operated the remotes both independently and also sometimes as a team, where each operated a single remote at the same time. Unknown to the children, the MIT-P has been set to a 1:2 ratio where the right cursor must always be twice as high on the screen as the left cursor for the screen to turn green.

So far, Lilah and Gideon have articulated several different strategies they discovered for turning the screen green. One strategy Gideon contributed involves moving the cursors continuously up the screen and always keeping the right-hand cursor “go-

ing up faster” (Gideon) than the left-hand cursor. The second strategy, which Lilah and Gideon described together, involves finding sets of discrete positions on the screen by ensuring that the right-hand cursor is always “twice as high” (Lilah and Gideon) as the left-hand cursor. After the children summarize their strategies, Dejan asks Lilah and Gideon how their height and speed strategies might be related: “Do you think there is some connection with how this one, the right one, is double the left and speed?” He points to the right-hand remote resting on the table in front of Lilah.

Lilah is first to volunteer an explanation about how the strategies are connected. She says: “Maybe because like if- because if you’re gonna to do it like at the same time that one would have to go faster t’like end at the same time, that one would have to lift higher and faster.” However, the idea she shares is more complex than the words above convey. It is composed through an elaborate system of interdependent linguistic and embodied semiotic resources, including a multipart gesture that includes a variety of distinct phrases, each with unique trajectories and handshapes. The transcript of her 10-second multimodal utterance appears in Figure 3 of Excerpt 1 (see Appendix 1 for transcript conventions).

The multipart gesture Lilah coordinates with her talk has six separate gestural phrases (1.3, 1.5, 1.6, 1.9, 1.11, and 1.12) and three different hand shapes (1.3, 1.6, 1.9, 1.11 – curled hands; 1.5 – extended index finger; and 1.12 – fists). First, when Lilah says “Maybe because like if- because if you’re gonna do it like at the same time” (1.1, 1.2), she brings both of her hands to an equal height at chest level (1.3) as she says “the same time.” Her hands are curled in a prehensile posture [Streeck, 2009] as though she is holding invisible objects. Next, Lilah points to the right remote lying on the table (1.5) as she says “that one” (1.4). Then Lilah raises her right hand above her left hand so that her right hand moves approximately twice as fast as her left hand and ends up approximately twice as high as the left hand (1.6). During this gesture phrase, her hands continue to be curled in a prehensile posture. As Lilah performs this gesture phrase, she says “would have to go faster” (1.7). Then, she continues, saying “t’like end at the same time” (1.8), and she again holds both curled hands level at chest height (1.9). Next, she says “that one would have to” (1.10) and repeats a larger amplitude version of the gestural phrase from 1.6, lifting her right hand higher and faster than her left, and keeping her hands in prehensile posture (1.11). As she raises both hands together in fists level with her forehead (1.12), she says “lift higher” (1.13). Then, she lowers her hands to her lap. She finishes saying softly “(and) faster” (1.15). The monosyllabic word Lilah says between higher and faster is difficult to make out, but a good approximation is that Lilah says “and” based on the shape her mouth makes.

Gesture phrases 1.3 and 1.9 are morphologically similar (holding curled hands level at chest height) and phrases 1.6 and 1.11 are morphologically similar (raising curled hands in 1:2 height ratio). Gesture phrase 1.12 is unique, as is the pointing gesture Lilah makes in 1.5. In sum, there are six distinct gesture phrases to this fluently unfolding gesture, but there are only four unique morphological forms. Although the utterance is broken down for analysis, it is important to note that it is delivered as one continuous, fluent stream by Lilah.

Lilah’s response constructs a multimodal, embodied rendering of what “faster” means, using a variety of different semiotic resources and drawing on her prior physical experiences of operating the device. How does she make this idea available to her onlooking audience? Her embodied performance is intricate. It juxtaposes several animated illustrations, evoking the abstract notion of *simultaneity* and integrating

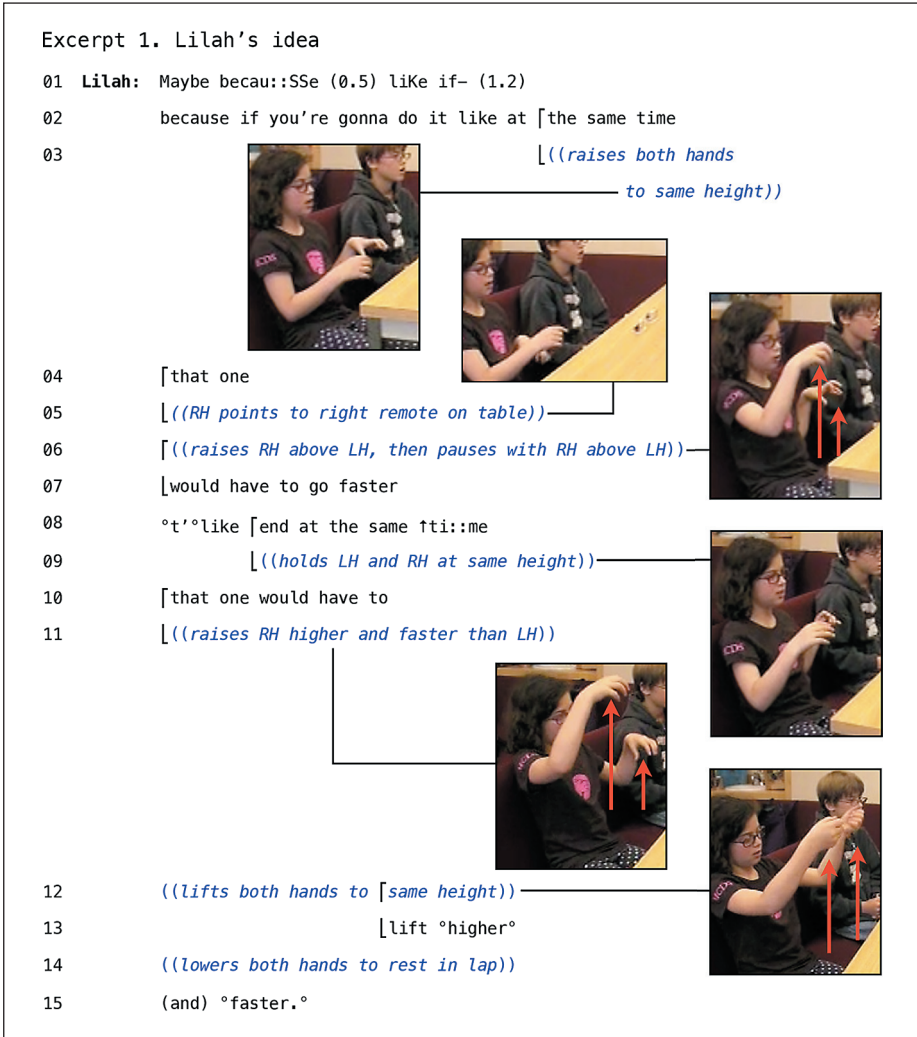


Fig. 3. Lilah explains her idea for how the strategy of keeping the right hand “twice as high” is connected to the strategy of keeping the right hand “moving up faster” (for explanations of transcript conventions, see Appendix 1).

this with a depiction of the physical activity of operating the remotes to highlight the different heights traveled.

In gesture phrases 1.3 and 1.9, Lilah’s movements create a material version of an immaterial entity: time. By bringing her hands together at a level height, co-timed with the speech “at the same time,” Lilah develops an embodied construal of the abstract notion of *simultaneity*. She evokes simultaneity by manually enacting a spatial metaphor of two physical objects (her hands) being at the same location. Streeck

[2008, 2009] has called this process of gesturing *ceiving*, where the hands give material, dynamic form to ideas that have no literal, physical presence in the world. Lilah uses this multimodal imagery of simultaneity to evoke the premise of the remotes *moving at the same time and ceasing motion at the same time*. Thus, she depicts an abstract property of their trajectories over time.

In 1.5, Lilah points to the right-hand remote on the table, providing a referent for “that one” (1.4) and projecting that her audience should seek out the role this (right-hand) remote plays in her forthcoming description. Then, in gesture phrases 1.6 and 1.11, Lilah’s performance creates a recognizable demonstration of how the body must move to operate the remotes so that the screen stays green, and how the right remote moves faster to end up higher. The prehensile posture of her hands makes it possible to see the empty space as though she is grasping phantom remotes [Nemirovsky et al., 2012]. Through the position of her body, the movement of her arms, and her hand shapes, it is possible to see Lilah *as if* she is making the screen green on the MIT-P. These performances make available information about the necessary trajectory, relative vertical displacements, and speed of each remote.

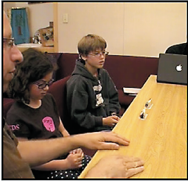
Finally, in the last gestural phrase – raising both hands to her forehead (1.12) – Lilah does not seem to be depicting her operation of the remotes anymore. The remotes never move upwards at the same height at the same time to make the screen green. Lilah’s hands are now shaped into fists, which do not clearly evoke holding phantom remotes. Instead, this move evokes the abstract comparative notion of “higher” in reference to the right-hand remote’s behavior (“that one”) with respect to the left-hand remote. When she lowers her hands to rest, she adds “(and) faster” (1.15).

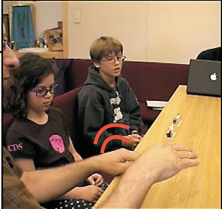
What Lilah accomplishes with her embodied performance can be interpreted as a demonstration that *to keep the screen green, the right remote has to move faster because it has to lift higher than the left remote if both remotes start and stop moving at the same time*. However, Lilah’s idea is open to interpretation and her listeners may not understand precisely what she is trying to express. Her complex multimodal utterance lays fertile ground for Ted to construct his revoicing turn (Fig. 4 of Excerpt 2).

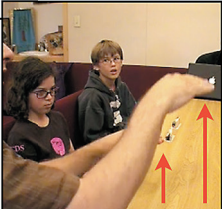
Ted initially responds to Lilah’s idea, saying, “that’s really cool.” He then asks Gideon if Lilah’s idea made sense and invites him to repeat it. When Gideon declines after a pause, Ted responds by multimodally revoicing Lilah’s idea. He says: “Okay, so if I understood correctly what Lilah is saying, is, well, the one on the right has to be double as high as the one on the left, if they’re moving up at the same time, the only way it will get there if it is- if it- it has kind of more- it has to, it’s got- this is my word, it has more ground to cover.” However, like Lilah’s turn, Ted’s response is multimodal and contains a complex, multipart gesture (Excerpt 2).


As Ted begins his response, he designates what is to follow as a recounting of Lilah’s idea by attributing it to her by name (2.2) [O’Connor & Michaels, 1996]. Then, he places his hands flat on the table in a visible display that they will be relevant in his forthcoming account (2.3). When he has secured both children’s gaze, Ted raises his right hand palm flat up above the table, as he says “the one on the right” (2.2, 2.4). Then he lowers his right hand and raises his left hand half as high as he says “has to be” (2.4). When he says “double as high as the one on the left” (2.6), he raises his right hand above his left hand in a 1:2 height ratio. As he finishes, he bobs his left hand up and down in the air (2.7). Here, Ted elaborates Lilah’s original performance by making the

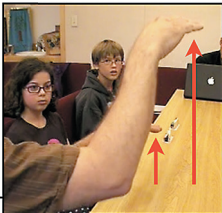
Excerpt 2. Ted revoices Lilah's idea

01 **Ted:** Okay so if I understood correctly
 02 what Lilah is saying is: s, (0.8) [well (0.4) the one
 03  [((lays both hands flat on the table))

04 [on the ri:ght (0.7) [has to be (1.0)
 05 [((lifts RH off the table then lowers it)) [((lifts LH off table))
 06 [double as high >as the one on [the left,<
 07 [((raises RH above LH)) [((bounces LH)) 

08 [if they're moving up at the same time,
 09 [((begins raising RH faster than LH)) 

10 [>the #o::nly# way it will get there=
 11 [((continues raising RH faster than LH))
 12 [=if it is- if it-< it has kind of more-
 13 [((lowers hands and holds RH above LH))
 14 it has to, (0.5) [it's got- this is my word,
 15 [((brings RH to chest)) 

16 it has [more ground to cover. (1.0)
 17 [((raises RH [faster than LH)) 

18 **Lilah:** [Ye::ah=it's more ground to cover if you wanna
 19 finish at the same time.

Fig. 4. Ted revoices Lilah's idea about how the strategy of keeping the right hand "twice as high" is connected to the strategy of keeping the right hand "going up faster" (for explanations of transcript conventions, see Appendix 1).

premise more explicit: *In general, the right remote has to be double as high as the left remote.* By lifting his right hand off the table, lowering it, then lifting his left hand and raising the right hand again, he visually highlights the difference in their heights.

As Ted begins narrating “if they’re moving up at the same time” (2.8), he moves his right hand, still flat, upwards twice as fast as his flat left hand and ends in a position that is twice as high as the left hand (2.9/2.11). As he says this, he raises the pitch of his voice. He continues “the only way it will get there” (2.10), while straining his voice and stretching out the word “only.” He interrupts himself to insert “this is my word” (2.14), touching his chest (2.15). Then, he finishes with “it has more ground to cover” (2.16), while he again enacts the right hand traveling twice as fast as the left hand and ending up twice as high (2.17). Lilah displays her attention to Ted’s performance by keeping her gaze on Ted’s hands throughout his turn. She only briefly looks down when Ted lowers his hands in 2.13 but returns her gaze just after Ted touches his chest (2.15).

Ted’s multimodal response is recognizable as revoicing because (a) he explicitly attributes the idea to Lilah (2.1), (b) his gesture shares morphological similarities with Lilah’s, and (c) his words closely parallel Lilah’s. Ted’s use of “the one on the right” can be heard as tied to and referring back to Lilah’s “that one” (the right remote). When Ted holds his hands in a 1:2 height ratio (2.7) and then raises them (2.9/2.11 and 2.17) to maintain this height ratio, he is visibly repeating Lilah’s gestural phrases produced in 1.6 and 1.11. His narration “if they’re moving up at the same time” (2.8) closely resembles Lilah’s stated premise “if you’re going to do it like at the same time” (1.2).

However, Ted’s turn subtly but significantly reformulates aspects of Lilah’s original idea, both in gesture and in speech. First, Ted elaborates the premise, highlighting and quantifying the double height difference with added gesture and speech in 1.2–1.7. Ted also adds iconic prosody [Shintel, Nusbaum, & Okrent, 2006] as a semiotic resource [Goodwin, 2018] to emphasize part of Lilah’s idea: by increasing the pitch of his voice to correspond with his verbal and gestural description of “moving up at the same time” (2.8) and then straining and elongating his voice on the word “only” when he refers to the right remote (2.10), Ted’s prosody and voice quality evoke the extra effort or work that the right hand must exert relative to the left when they are rising simultaneously and ending in a 1:2 height ratio at the same time.¹

Ted notably does not reproduce any of Lilah’s gestural phrases where she holds her hands at equal heights. We cannot know definitively why Ted omitted Lilah’s *simultaneity* gesture phrases (1.3 and 1.9) and her *higher* gesture phrase (1.12). One possible reason is that the omitted gesture phrases are potentially misleading: when Lilah holds her hands at equal heights, she continues to keep her hands curled or in fists as though she is still holding phantom remotes. However, producing green feedback is never technically possible when the remotes are held at the same height with the MIT-P set to a ratio of 1:2. Another possibility is that Ted leaves out Lilah’s phrases to simplify the imagery because they are redundant: Lilah’s gestural phrase from 1.6 and 1.11 that Ted *does* repeat *already* demonstrates simultaneity – the two displacements occur during the same time interval (the hands move at the same time and stop at the same time). Her gestural imagery of *higher* conveyed in 1.12 is also

¹ While Macbeth [2011] documented iconic speech as a semiotic resource used by a teacher in a science class (in the example, a science teacher asks a student “how fast” dye is spreading in hot water, and speaks the words very rapidly), the role iconic speech plays in multimodal dialogic interactions in learning settings is underexamined.

already demonstrated in the gesture phrase Ted repeats, since the right displacement is shown to be greater. Either way, the omission of these phrases is a visible display that Ted considers their imagery irrelevant to the core of the idea.

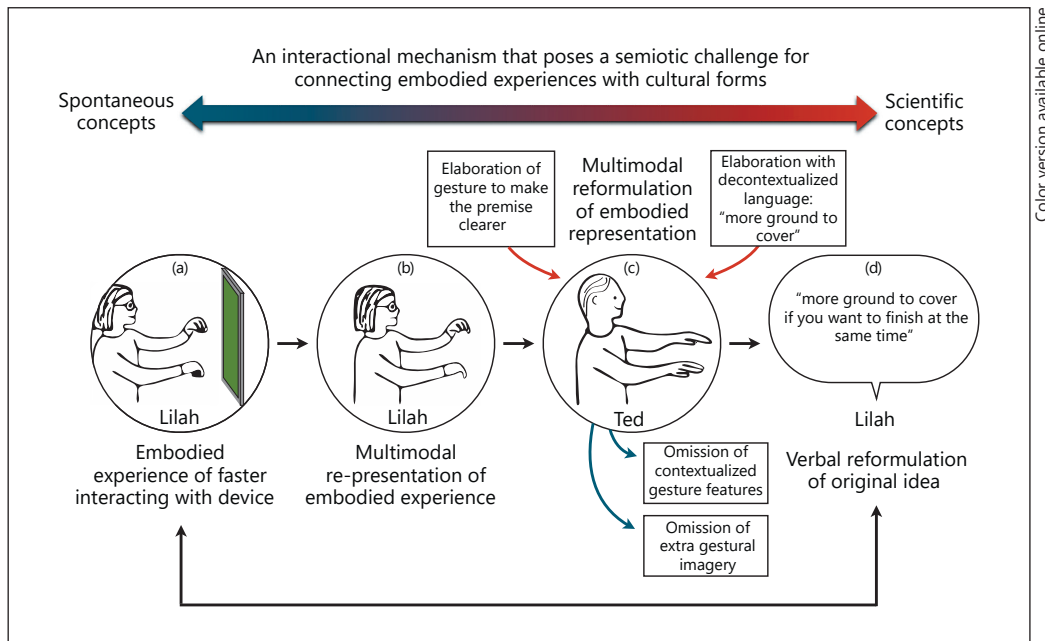
In another significant reformulation of Lilah's gesture, Ted holds his hands out flat when he moves them, instead of reproducing Lilah's curled, prehensile hand shapes. In Lilah's gesture phrases, her hands visibly resemble and act as hands operating the device. Her movements in 1.6 and 1.11 recreate part of the *kinetic-kinesthetic* experience [Streeck, 2013] of achieving a green screen with the device. Ted's flattened hand shapes, on the other hand, suggest he is *bounding regions of space to show extent* instead of *pantomiming the operation of the remotes*. Streeck has discussed *bounding* and *pantomiming* as distinct practices for evoking aspects of the world in people's gestural repertoires [Streeck, 2008]. By omitting some of the detail provided in Lilah's rendering, Ted is able to make *the extent of the distances traveled* become the figural component of Lilah's gesture phrase and downplay the embodied gestalt of operating the device. This *decontextualizes* (c.f., decontextualized language [De Temple & Beals, 1991]) *the gesturally produced imagery* of distance, making it more generic as a bounded region of space.

In addition, Ted's *verbal* description reformulates Lilah's to be less situated in the specific circumstances of the manual action of manipulating the device by decontextualizing her *talk*. This presents a more standardized, generalized, culturally appropriate definition of faster as *greater distance traveled during the same amount of time*. Rather than talk about speed as related to vertical distance, he forgoes repeating Lilah's use of the word "height" and instead provides a description that is directionless and could apply to situations of spatial displacement beyond the MIT-P by using the phrase "ground to cover." As an idiom, to "cover ground" or to "go a given extent or considerable distance" [Ammer, 2013, p. 97] is likely a familiar description of the general idea of distance traveled accessible to Lilah and relatable to her experiences. Ted's ownership line "this is my word" is a preface that marks what comes next as a novel addition to Lilah's original contribution, alerting the children that they should see this as added content to the original substance of Lilah's idea.

In 2.18-19, Lilah *demonstrates* her understanding of Ted's reformulation by skillfully incorporating part of Ted's reformulation back into her original idea. In weaving together her own idea with Ted's, she displays her emerging appreciation of how the two strategies, "twice as high" and "going up faster," are connected: She agrees with Ted that the right remote has to move faster because "it's more ground to cover" to "finish at the same time."

How Multimodal Revoicing Interactions Create Semiotic Challenges

The fine details of this interaction occasioned a semiotic challenge for linking a spontaneous awareness with a scientific awareness of "faster." Lilah was able to organize aspects of her embodied experience of operating the MIT-P with a more standardized, generalized way of describing that experience, as evidenced by her skillful incorporation of Ted's reformulation (summarized in Fig. 5). The initial experience of faster and Lilah's first attempt to represent it (a spontaneous awareness) are what provided the raw materials for the development of this more refined version of faster (a scientific awareness). At the same time, when Lilah adopted Ted's revised descrip-



Color version available online

Fig. 5. Lilah uses her embodied experience with the device (a) to create a multimodal representation of her idea of “faster” (b). The reformulation of Lilah’s idea is accomplished by the tutor’s selective reproduction and elaboration of her gestural imagery, accompanied by the substitution of decontextualized language (c). Lilah uses this reformulation to give a revised, more refined description of “faster” that is closer to the culturally specified definition (d).

tion in her last turn (Fig. 5d), we can see how she now uses the more scientific concept to (re-)organize her perception of her past embodied experiences. Thus, it appears that Lilah’s spontaneous awareness grew upwards providing the material for the scientific awareness, and the scientific awareness grew downwards to organize the spontaneous awareness [Vygotsky, 1986].

Leading up to the episode examined above, Lilah had a variety of kinetic, haptic, and kinesthetic experiences as she worked with the device (Fig. 5a). She likely experienced many potentially relevant sensations – perhaps the feeling that it was taking more effort to move her right arm so that it would glide up above her left arm by a certain amount, or the feeling of having to stall her left arm to keep it slower underneath the right arm. When Dejan asked her to explain why the remote that goes higher is always faster too, Lilah was suddenly faced with finding a way to relate and convey these tacit [Polanyi, 1958] kinetic-kinesthetic experiences to her audience – that is, to make them publicly available. Lilah’s response involved efforts to capture, preserve, and *re-present* the embodied sensations of faster and higher that she had experienced so that she could communicate them. Lilah transformed her sensations into an explanatory, public re-*re-resentation* of experience, using gesture to invoke and render dynamic imagery for the idea of simultaneity and also to pantomime working with the device (Fig. 5b).

A significant and pervasive problem that newcomers in any community face is identifying which aspects of their experiences are relevant and valuable for appreciating the world in culturally or discipline-specific ways. Many studies have examined how newcomers' perceptions of their physical surroundings are instructed or disciplined by more experienced practitioners so they can begin to experience these settings in culturally or disciplinarily specific ways [Alac & Hutchins, 2004; Goodwin, 1994, 2018; Lindwall & Lymer, 2008; Stevens & Hall, 1998]. However, discovering *what is there to be learned* is never trivial [Zemel & Koschmann, 2014]. Here, Lilah had an opportunity to recognize how Ted had transformed her utterance and to adopt this transformation.

Lilah's public re-presentation of private sensory experience, composed of a variety of different semiotic resources (e.g., hand shapes, trajectories, words), created a *substrate* [Goodwin, 2018] for Ted to act on. Ted's next action was carefully calibrated and responsive to Lilah's public display of what she was attending to. He used the public substrate of her re-presented experience to build his response by decomposing, reusing, and transforming the semiotic resources Lilah had made available.

Ted's multimodal revoicing turn reformulated Lilah's explanation by adding gesture phrases, selectively repeating some gesture phrases and omitting others, modifying hand shapes, and elaborating verbal content (Fig. 5c). His recognizable repetition of one of Lilah's gesture phrases served as a resource for Lilah to recognize the continuity between her first explanation and Ted's more generalized verbal description "more ground to cover." Aspects of Lilah's gesture specific to the manual operation of the device were omitted, further generalizing the idea to situations outside the immediate one with the MIT-P. Finally, omitting some of the extra imagery rendered in Lilah's gesture provided an opportunity to recognize the imagery Lilah included that was most pertinent to the idea of faster: traveling a greater distance in the same amount of time. Overall, Ted's reformulation created a culturally refined "scientific" version of Lilah's original explanation: the right hand can be said to be going faster *because* it has more ground to cover than the left hand in the same amount of time. Ted's reformulation provided Lilah with multimodal resources to recognize how her original explanation connected with the culturally appropriate definition of faster (i.e., greater distance traveled during the same interval of time).

Goodwin and others have demonstrated the myriad ways more experienced practitioners separate signal from noise to highlight consequential features of the physical environment (e.g., a patch of soil or a vat of chemical reaction), inscriptions produced by others (e.g., a map or graph), or cultural tools. Here, I have uncovered a slightly different form of highlighting: Instead of highlighting relevant features in the environment or in an inscription, Ted highlighted the relevant, consequential features in Lilah's *own semiotic rendering of the world* – her first efforts at representing her experience. Lilah's initial representation had already extracted and generalized some aspects of her rich embodied experiences and perceptions. In response, Ted further refined this representation so that it more closely resembles how scientists and mathematicians perceive what makes something faster. Ted was able to shape and instruct Lilah's perception of her embodied representational performance by calling attention to aspects of it that are relevant to a more culturally appropriate definition of faster and downgrading the relevance of other features. This multimodal revoicing interaction posed a semiotic challenge, inviting Lilah to reconsider her initial experience and her semiotic representation of that experience in a way more akin to how cultural members would.

Du Bois's [2014] *dialogic syntax* provides a more specific mechanism for how learning opportunities arise within co-operative action like multimodal revoicing sequences. He proposed that when speakers selectively reproduce aspects of prior utterances to build their own, the perceivable parallels that emerge create opportunities to draw inferences and co-construct new meanings. One mechanism for co-constructing new meanings from dialogic syntax is *substitution*. The perceived parallels between sequences of dialogic utterances render departures like substitutions especially noticeable. In turn, noticing these substitutions creates possibilities for perceiving new relationships – e.g., structural and functional equivalences – between the original and substituted components. For example, in Lilah and Ted's revoicing interaction, the emergent parallel structure between the multimodal utterances makes the points of departure noticeable. The substitution of "lift higher" with "more ground to cover" makes these different descriptions now recognizable as potentially interchangeable, *categorically equivalent* ways of describing the experienced event. The multimodal parallel structures of Lilah's and Ted's utterances also make the differing gestural imagery of their respective turns perceivable as interchangeable: Ted's decontextualized rendition can be appreciated as functionally equivalent to Lilah's. While Du Bois focused on linguistic forms, Arnold [2012] argued for extending dialogic syntax to embodied action to understand how gesture-mimicking contributes to learning. My analysis of Lilah and Ted's revoicing interaction extends Arnold's argument by illustrating a new form of dialogicity in multimodal semiotic resources that provides a semiotic mechanism for learning.

Overall, this complex instance of multimodal revoicing constitutes a clear example of dialogic learning [Koschmann, 1999]. The explanation of the connection between "going up faster" and "twice as high" – that *faster* means more ground to cover in the same amount of time – is a situated, locally emergent product of Lilah and Ted's multimodal utterances coming into contact with and interanimating one another. The precise mechanism by which this new understanding was co-constructed is only recoverable with careful attention to the multimodal methods and resources Ted and Lilah used to interactionally achieve a temporarily shared perspective on reality in this moment.

Concluding Remarks

In this paper, I have demonstrated how multimodal revoicing interactions can provide semiotic challenges that create opportunities for learners to link spontaneous and scientific concepts in co-operative action. Through emergent multimodal dialogic syntax, learners can recognize what aspects of their experiences are considered relevant to the conventional, culturally specified ways of organizing and describing such experiences. In the case I examined, multimodal reformulation of a learner's idea provided her with opportunities to make links between her embodied experience of *faster* and a more standardized, generalized concept of *faster*. The multimodal revoicing turn was accomplished by selectively repeating and elaborating certain aspects of the learner's original gesture and by elaborating her accompanying verbal description. Following Wertsch and Stone, I extended Vygotsky's theory of development by contributing a new semiotic mechanism that provides an arena for the bidirectional linkage of spontaneous and scientific concepts.

By adopting an approach to social interaction inspired by EMCA and the cooperative action framework, I was able to reveal the precise methods and resources participants employed to create a temporarily shared reality. In particular, this approach made it possible to uncover the fine details of *ordinary communicational practices* for building and maintaining intersubjectivity that were able to generate a semiotic challenge. My analysis contributes to broadening our current understanding of the ways in which learning is a dialogic process, where new meanings arise moment by moment through the local, situated, embodied interactional work of participants. Additionally, by helping to further refine our understanding of the pedagogical practice of revoicing, this analysis helps bring into relief more of the nuanced and “seen-but-unnoticed” [Garfinkel, 1996] practices of instructional work.

In the interactional sequence I examined, I discovered a new form of multimodal revoicing that involves the reformulation of a learner’s multimodally expressed idea through the selective reproduction and modification of gesture phrases and handshapes alongside the elaboration of speech. This investigation also contributes the discovery of a new form of dialogicity in gesture, building on Arnold’s work on the dialogic syntax of gestures [Arnold, 2012]. The rich and thorough record of this event that I developed will make it possible to identify, investigate, and compare similar interactional processes in other settings and situations.

A limitation of this approach is that it cannot disclose the extended trajectory of a student’s ontogenetic development surrounding the appropriation of a particular cultural form (e.g., the definition of faster). However, it is able to robustly characterize the critical interactional mechanisms of moment-by-moment meaning-making that contribute to this trajectory over time [Saxe, 2012]. In this way, microanalytic investigations can complement existing longitudinal accounts of how students connect their spontaneous, embodied experiences with scientific concepts [e.g., Saxe et al., 2015]. Analytic attention to the fine details of how intersubjectivity is procedurally accomplished in familiar interactional sequences like revoicing holds potential for better understanding processes of learning and development.

This study also adds to a growing body of work [Arzarello et al., 2008, Shein, 2012] that documents the beneficial role that “conversations of gesture” have in developing mathematical and scientific ideas. Attending to, reproducing, and modifying learners’ gestures during unfolding dialog can create opportunities to enrich learning and may be a productive practice for instructors to adopt intentionally as part of teaching practices that strive to be responsive to the substance of students’ ideas in the moment. While the dialogic nature of talk has been thoroughly examined in meaning-making processes in instructional interactions [Alexander, 2006; Mercer, Dawes, & Staarman, 2009; Scott, Mortimer, & Aguiar, 2006], more investigations and attention are needed to establish the different roles that *dialogic multimodal semiotic resources* play during instruction to create ethnographically adequate [McDermott et al., 1978] accounts of these sequences. In particular, as STEM learning environments are increasingly being designed to deliberately incorporate the body into explorations of STEM phenomena [e.g., Abrahamson & Lindgren, 2014], more careful examinations of the interactional mechanisms that allow learners and instructors to negotiate, make sense of, and connect embodied experiences to scientific concepts are urgently needed.

Acknowledgments

I would like to thank Rogers Hall, Geoffrey Saxe, Niral Shah, and Dor Abrahamson for helpful conversations about this project. I am especially grateful for the thorough and constructive reviews from Timothy Koschmann and an anonymous reviewer that vastly improved this paper. I would also like to thank Charles Goodwin's Co-Operative Action Lab at UCLA and Eve Sweetser's Gesture and Multimodality Group at UC Berkeley for invaluable opportunities to learn and be inspired. Finally, I thank Benedikt W. Harrer for generous technical support, editing, and many illuminating discussions.

This material is based on work supported in part by the National Science Foundation under Award No. 1321042 and in part by the Institute of Education Sciences under Award No. R305B090026. Preliminary findings appeared in a paper presented at the annual meeting of the American Educational Research Association in Chicago, April 16–20, 2015, and a paper presented at the annual meeting of the International Society for Gesture Studies in Paris, France, July 18–22, 2016.

Appendix 1

Adapted Jeffersonian transcript conventions [Jefferson, 2004].

Symbol	Sample	Meaning
<i>Brackets</i>	[so he was you're it	Denotes overlapping talk or bodily activity
<i>Equals sign</i>	Does=	Latching: one utterance follows the other unusually quickly
<i>Number in parentheses</i>	(2.5)	Timed pause in seconds (measured to a tenth of a second)
<i>Period in parentheses</i>	(.)	Micropause less than a tenth of a second
<i>Pound signs</i>	#only#	Exaggerated voice quality
<i>Period or down arrow</i>	. or ↓	Falling pitch
<i>Question mark or up arrow</i>	? or ↑	Rising pitch
<i>Comma</i>	,	Slight falling pitch
<i>Hyphen</i>	if this-	Cut off, abrupt stop, unfinished words
<i>Greater than/less than</i>	>if this is<	Speech delivered more quickly than surrounding talk
<i>Less than/greater than</i>	<if this is>	Speech delivered more slowly than surrounding talk
<i>Degree signs</i>	°So°	Quiet speech or whisper
<i>Capital letters</i>	TAKE IT	Loud speech or shouting
<i>Underline</i>	Take it <u>now</u>	Denotes emphasis (voiced stress)
<i>Colons</i>	So::o	Indicate elongated word or syllable
<i>Repeated letter h</i>	hhh	Audible exhalation
<i>Text in parentheses</i>	(and)	Indicates analyst uncertainty, approximation
<i>Text in double parentheses</i>	((RH flick))	Annotation of bodily activity, RH = right hand; LH = left hand

Appendix 2

Verbal transcript of the multimodal revoicing episode.

Dejan: Do you think there is some connection with (1.2) how this one is, (0.5) the, (0.4) right one is double the left (0.5) and speed?

Gideon: Um (5.2)

-> Lilah: Maybe because::Sse (0.5) like if- (1.2) because if you're gonna do it like at the same time that one would have to go faster 't'like end at the same time that one would have to lift 'higher' (and) 'faster.'

Ted: Mhm that's really cool. Gi- Gideon, did that make sense what Lilah just said?=
Gideon: =Yeah

Ted: Could you kindly repeat that?

Gideon: Uhm (5.3) [°um°

Ted: [Hhh hhh hi hhh it's always- always difficult to repeat what someone else said. (2.7) I- I- I'll tell you what, shall I tell you what I understood from Lilah, or do you want to have- (0.4) give it a shot?

Gideon: Uhee=you tell.

-> Ted: Okay so if I understood correctly what Lilah is saying is: (0.8) well (0.4) the one on the right (0.7) has to be (1.0) double as high >as the one on the left, < if they're moving up at the same time, >the #:only# way it will get there=if it is- if it-< it has kind of more- it has to, (0.5) it's got- this is my word, it has more ground to cover. (1.0)

Lilah: Ye::ah=it's more ground to cover if you wanna finish at the same time.

References

- Abrahamson, D., Gutiérrez, J., Charoenying, T., Negrete, A., & Bumbacher, E. (2012). Fostering hooks and shifts: Tutorial tactics for guided mathematical discovery. *Technology, Knowledge and Learning, 17*, 61–86. doi:10.1007/s10758-012-9192-7
- Abrahamson, D., Lee, R.G., Negrete, A.G., & Gutiérrez, J.F. (2014). Coordinating visualizations of polysemous action: Values added for grounding proportion. *ZDM, 46*, 79–93. doi:10.1007/s10758-012-9192-7
- Abrahamson, D., & Lindgren, R. (2014). Embodiment and embodied design. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 358–376). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139519526.022
- Alac, M., & Hutchins, E. (2004). I see what you are saying: Action as cognition in fMRI brain mapping practice. *Journal of Cognition and Culture, 4*, 629–661. doi:10.1163/1568537042484977
- Alexander, R.J. (2006). *Towards dialogic teaching*. New York, NY: Dialogos.
- Alibali, M.W., Nathan, M.J., Wolfgram, M.S., Church, R.B., Jacobs, S.A., Johnson Martinez, C., & Knuth, E.J. (2014). How teachers link ideas in mathematics instruction using speech and gesture: A corpus analysis. *Cognition and Instruction, 32*, 65–100. doi:10.1080/07370008.2013.858161
- Ammer, C. (2013). *The American heritage dictionary of idioms* (2nd ed.). Boston, MA: Houghton Mifflin Harcourt.
- Arnold, L. (2012). Dialogic embodied action: Using gesture to organize sequence and participation in instructional interaction. *Research on Language and Social Interaction, 45*, 269–296. doi:10.1080/08351813.2012.699256
- Arzarello, F., Paola, D., Robutti, O., & Sabena, C. (2008). Gestures as semiotic resources in the mathematics classroom. *Educational Studies in Mathematics, 70*, 97–109. doi:10.1007/s10649-008-9163-z
- Bahktin, M. (1981). *The dialogic imagination* (M. Holquist, Ed.). Austin, TX: University of Texas Press.
- Becvar, L.A., Hollan, J., & Hutchins, E. (2005). Hands as molecules: Representational gestures used for developing theory in a scientific laboratory. *Semiotica, 1*, 89–112. doi:10.1515/semi.2005.2005.156.89
- Crowder, E.M. (1996). Gestures at work in sense-making science talk. *The Journal of the Learning Sciences, 5*, 173–208. doi:10.1207/s15327809jls0503_2
- De Fornel, M. (1992). The return gesture: Some remarks on context, inference, and iconic gesture. In P. Auer, A. Di Luzio, & A. Di Luzio (Eds.), *Contextualization of language* (pp. 159–176). Amsterdam: Benjamins. doi:10.1075/pbns.22.11for

- De Temple, J.M., & Beals, D.E. (1991). Family talk: Sources of support for the development of decontextualized language skills. *Journal of Research in Childhood Education*, 6, 11–19. doi:10.1080/02568549109594818
- Du Bois, J.W. (2014). Towards a dialogic syntax. *Cognitive Linguistics*, 25, 359–410. doi:10.1515/cog-2014-0024
- Eckert, A., & Nilsson, P. (2017). Introducing a symbolic interactionist approach on teaching mathematics: The case of revoicing as an interactional strategy in the teaching of probability. *Journal of Mathematics Teacher Education*, 20, 31–48. doi:10.1007/s10857-015-9313-z
- Enyedy, N. (2005). Inventing mapping: Creating cultural forms to solve collective problems. *Cognition and Instruction*, 23, 427–466. doi:10.1207/s1532690xci2304_1
- Erickson, F. (2011). Reflections on practice, teaching/learning, video, and theorizing. In T. Koschmann (Ed.), *Theories of learning and studies of instructional practice* (pp. 385–402). New York, NY: Springer. doi:10.1007/978-1-4419-7582-9_23
- Flood, V.J., Amar, G., Nemirovsky, R., Harrer, B.W., Bruce, M.R.M., & Wittmann, M.C. (2015). Paying attention to gesture when students talk chemistry: Interactional resources for responsive teaching. *Journal of Chemical Education*, 92, 11–22. doi:10.1021/ed400477b
- Flood, V.J., Harrer, B.W., & Abrahamson, D. (2016). The interactional work of configuring a mathematical object in a technology-enabled embodied learning environment. In C.-K. Looi, J. Polman, U. Cress, & P. Reimann (Eds.), *Proceedings of the International Conference of the Learning Sciences. Vol. 1* (pp. 122–129). Singapore: ISLS.
- Forman, E.A., & Ansell, E. (2002). Orchestrating the multiple voices and inscriptions of a mathematics classroom. *Journal of the Learning Sciences*, 11, 251–274. doi:10.1080/10508406.2002.9672139
- Forman, E.A., & Larreamendy-Joerns, J. (1998). Making explicit the implicit: Classroom explanations and conversational implicatures. *Mind, Culture & Activity*, 5, 105–113. doi:10.1207/s15327884mca0502_4
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood, NJ: Prentice-Hall.
- Garfinkel, H. (1996). Ethnomethodology's program. *Social Psychology Quarterly*, 59, 5–21. doi:10.2307/2787116
- Garfinkel, H. (2002). *Ethnomethodology's program: Working out Durkheim's aphorism*. New York, NY: Rowman & Littlefield.
- Gibson, J.J. (1979). *The ecological approach to perception*. Boston, MA: Houghton Mifflin.
- Goffman, E. (2013). *Encounters: Two studies in the sociology of interaction*. Mansfield Centre, CT: Martino Publishing.
- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96, 606–633. doi:10.1525/aa.1994.96.3.02a00100
- Goodwin, C. (1997). The blackness of black: Color categories as situated practice. In L.B. Resnick, R. Saljo, C. Pontecorvo, & B. Burge (Eds.), *Discourse, tools and reasoning: Essays on situated cognition* (pp. 111–140). New York, NY: Springer. doi:10.1007/978-3-662-03362-3_6
- Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32, 1489–1522. doi:10.1016/S0378-2166(99)00096-X
- Goodwin, C. (2007). Environmentally coupled gestures. In S. Duncan, J. Cassell, & E. Levy (Eds.), *Gesture and the dynamic dimension of language: Essays in honor of David McNeill* (pp. 195–212). Philadelphia, PA: Benjamins. doi:10.1075/gs.1.18goo
- Goodwin, C. (2013). The co-operative, transformative organization of human action and knowledge. *Journal of Pragmatics*, 46, 8–23. doi:10.1016/j.pragma.2012.09.003
- Goodwin, C. (2018). *Co-operative action*. Cambridge, MA: Cambridge University Press.
- Goodwin, M.H., & Alim, H.S. (2010). “Whatever (neck roll, eye roll, teeth suck)”: The situated coproduction of social categories and identities through stancetaking and transmodal stylization. *Journal of Linguistic Anthropology*, 20, 179–194. doi:10.1111/j.1548-1395.2010.01056.x
- Grice, P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.
- Heritage, J. (1984). *Garfinkel and ethnomethodology*. New York, NY: Polity Press.
- Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G.H. Lerner (Ed.), *Conversation analysis: Studies from the first generation* (pp. 13–31). Amsterdam: Benjamins. doi:10.1075/pbns.125.02jef
- Kastens, K., Agrawal, S., & Liben, L. (2008). Research methodologies in science education: The role of gestures in geoscience teaching and learning. *Journal of Geoscience Education*, 56, 362–368.
- Keevallik, L. (2010). Bodily quoting in dance correction. *Research on Language and Social Interaction*, 43, 401–426. doi:10.1080/08351813.2010.518065
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511807572
- Kimbara, I. (2006). On gestural mimicry. *Gesture*, 6, 39–61. doi:10.1075/gest.6.1.03kim
- Koschmann, T. (1999). Toward a dialogic theory of learning: Bakhtin's contribution to understanding learning in settings of collaboration. In C. Hoadley & J. Roschelle (Eds.), *Proceedings of the 1999 conference on computer support for collaborative learning* (pp. 308–313). Palo Alto, CA: ISLS. doi:10.3115/1150240.1150278

- Koschmann, T., & LeBaron, C. (2002). Learner articulation as interactional achievement: Studying the conversation of gesture. *Cognition and Instruction*, 20, 249–282. doi:10.1207/S1532690XCI2002_4
- Koschmann, T., LeBaron, C., Goodwin, C., & Feltovich, P. (2011). “Can you see the cystic artery yet?” A simple matter of trust. *Journal of Pragmatics*, 43, 521–541. doi:10.1016/j.pragma.2009.09.009
- Koschmann, T., LeBaron, C., Goodwin, C., Zemel, A., & Dunnington, G. (2007). Formulating the triangle of doom. *Gesture*, 7, 97–118. doi:10.1075/gest.7.1.06kos
- Koschmann, T., & Zemel, A. (2014). Instructed objects. In M. Nevile, P. Haddington, T. Heinemann, & M. Rauniomaa (Eds.), *Interacting with objects: Language, materiality, and social activity* (pp. 357–377). Philadelphia, PA: Benjamins. doi:10.1075/z.186.16kos
- Kress, G., Jewitt, C., Ogborn, J., & Tsatsarelis, C. (2001). *Multimodal teaching and learning: The rhetorics of the science classroom*. London: University of London.
- Krussel, L., & Edwards, B. (2004). The teacher’s discourse moves: A framework for analyzing discourse in mathematics classrooms. *School Science and Mathematics*, 104, 307–312. doi:10.1111/j.1949–8594.2004.tb18249.x
- Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods, Instruments, & Computers*, 41, 841–849. doi:10.3758/BRM.41.3.841
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, MA: Cambridge University Press. doi:10.1017/CBO9780511815355
- LeBaron, C., & Streeck, J. (2000). Gestures, knowledge, and the world. In D. McNeill (Ed.), *Language and gesture* (pp. 118–138). Cambridge, MA: Cambridge University Press. doi:10.1017/CBO9780511620850.008
- Lerner, G.H. (2002). Turn-sharing: The choral production of talk-in-interaction. In C.E. Ford, B. Fox, & S. Thompson (Eds.), *Language of turn and sequence* (pp. 225–236). New York, NY: Oxford University Press.
- Lindwall, O., & Lymer, G. (2008). The dark matter of lab work: Illuminating the negotiation of disciplined perception in mechanics. *Journal of the Learning Sciences*, 17, 180–224. doi:10.1080/10508400801986082
- Luff, P., & Heath, C. (2015). Transcribing embodied action. In D. Tannen, H.E. Hamilton, & D. Schiffrin (Eds.), *The handbook of discourse analysis. Vol. 2* (pp. 367–390). Hoboken, NJ: Wiley & Sons. doi:10.1002/9781118584194.ch17
- Macbeth, D. (2011). Understanding understanding as an instructional matter. *Journal of Pragmatics*, 43, 438–451. doi:10.1016/j.pragma.2008.12.006
- Majlesi, A.R. (2015). Matching gestures: Teachers’ repetitions of students’ gestures in second language learning classrooms. *Journal of Pragmatics*, 76, 30–45. doi:10.1016/j.pragma.2014.11.006
- McDermott, R., Gospodinoff, K., & Aron, J. (1978). Criteria for an ethnographically adequate description of concerted activities and their contexts. *Semiotica*, 24, 245–275. doi:10.1515/semi.1978.24.3–4.245
- McNeill, D. (1992). *Hand and mind*. Chicago, IL: University of Chicago Press.
- McNeill, D. (2000). Catchments and contexts: Non-modular factors in speech and gesture production. In D. McNeill (Ed.), *Language and gesture* (pp. 312–328). New York, NY: Cambridge University Press. doi:10.1017/CBO9780511620850.019
- Mercer, N., Dawes, L., & Staarman, J.K. (2009). Dialogic teaching in the primary science classroom. *Language and Education*, 23, 353–369. doi:10.1080/09500780902954273
- Mondada, L. (2012). The conversation analytic approach to data collection. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 32–56). Boston, MA: Blackwell Publishing Ltd. doi:10.1002/9781118325001.ch3
- Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics*, 20, 336–366. doi:10.1111/josl.1_12177
- Nam, O.K., Ju, M.K., Rasmussen, C., Marrongelle, K., & Park, J.H. (2008). Utilization of revoicing based on learners’ thinking in an inquiry-oriented differential equations class. *The SNU Journal of Education Research*, 17, 111–134.
- Nathan, M., & Alibali, M.W. (2011). How gesture use enables intersubjectivity in the classroom. In G. Stam & M. Ishino (Eds.), *Integrating gestures: The interdisciplinary nature of gesture* (pp. 257–267). Amsterdam: Benjamins. doi:10.1075/gs.4.23nat
- Nemirovsky, R., & Ferrara, F. (2009). Mathematical imagination and embodied cognition. *Educational Studies in Mathematics*, 70, 159–174. doi:10.1007/s10649–008–9150–4
- Nemirovsky, R., Kelton, M.L., & Rhodehamel, B. (2012). Gesture and imagination: On the constitution and uses of phantasms. *Gesture*, 12, 130–165. doi:10.1075/gest.12.2.02nem
- Newman, D., Griffin, P., & Cole, M. (1989). *The construction zone: Working for cognitive change in school*. New York, NY: Cambridge University Press.
- O’Connor, M.C., & Michaels, S. (1993). Aligning academic task and participation status through revoicing: Analysis of a classroom discourse strategy. *Anthropology and Education*, 24, 314–318. doi:10.1525/aeq.1993.24.4.04x0063k

- O'Connor, M.C., & Michaels, S. (1996). Shifting participant frameworks: Orchestrating thinking practices in group discussion. In D. Hicks (Ed.), *Discourse, learning and schooling* (pp. 63–103). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511720390.003
- Park, Y. (2013). The roles of third-turn repeats in two L2 classroom interactional contexts. *Applied Linguistics*, 35, 145–167. doi:10.1093/applin/amt006
- Parsons, C.S. (2017). Socializing intellectual talk: A case study of instructor follow-up statements in classroom discourse. *Journal of the Scholarship of Teaching and Learning*, 17, 63–77. doi:10.14434/v17i1.20514
- Polanyi, M. (1958). *Personal knowledge. Towards a post-critical philosophy*. New York, NY: Harper Torch Books.
- Pozzer-Ardenghi, L., & Roth, W.-M. (2007). On performing concepts during science lectures. *Science Education*, 91, 36–74. doi:10.1002/sce.20172
- Radford, L. (2003). Gestures, speech, and the sprouting of signs: A semiotic-cultural approach to students' types of generalization. *Mathematical Thinking and Learning*, 5, 37–41. doi:10.1207/S15327833MTL0501_02
- Radford, L. (2009). Why do gestures matter? Sensuous cognition and the palpability of mathematical meanings. *Educational Studies in Mathematics*, 70, 111–126. doi:10.1007/s10649-008-9127-3
- Rasmussen, C., Stephan, M., & Allen, K. (2004). Classroom mathematical practices and gesturing. *The Journal of Mathematical Behavior*, 23, 301–323. doi:10.1016/j.jmathb.2004.06.003
- Rogoff, B. (1990). *Apprenticeship in thinking*. New York, NY: Oxford University Press.
- Rommetveit, R. (1979). On “meanings” of situations and social control of such meaning in human communication. Paper presented at the Symposium on the Situation in Psychological Theory and Research, Stockholm.
- Roth, W.-M., & Lawless, D. (2002). Scientific investigations, metaphorical gestures, and the emergence of abstract scientific concepts. *Learning and Instruction*, 12, 285–304. doi:10.1016/S0959-4752(01)00023-8
- Roth, W.-M., & Thom, J.S. (2009). The emergence of 3D geometry from children's (teacher-guided) classification tasks. *Journal of the Learning Sciences*, 18, 37–41. doi:10.1080/10508400802581692
- Roth, W., & Welzel, M. (2001). From activity to gestures and scientific language. *Journal of Research in Science Teaching*, 38, 103–136. doi:10.1002/1098-2736(200101)38:1 < 103::AID-TEA6 > 3.0.CO;2-G
- Ruiz-Primo, M.A., & Furtak, E.M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, 44, 57–84. doi:10.1002/tea.20163
- Sacks, H., Schegloff, E.A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, 50, 696–735. doi:10.1353/lan.1974.0010
- Saxe, G.B. (2008). Reflections on J.V. Wertsch's “From social interaction to higher psychological processes.” *Human Development*, 51, 80–89. doi:10.1159/000113157
- Saxe, G.B. (2012). *The cultural development of mathematical ideas*. New York, NY: Cambridge University Press. doi:10.1017/CBO9781139045360
- Saxe, G.B., De Kirby, K., Kang, B., Le, M., & Schneider, A. (2015). Studying cognition through time in a classroom community: The interplay between “everyday” and “scientific concepts.” *Human Development*, 58, 5–44. doi:10.1159/000371560
- Schegloff, E. (1991). Conversation analysis and socially shared cognition. In L.B. Resnick, J.M. Levine, & S.D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 150–170). Washington, DC: APA. doi:10.1037/10096-007
- Schegloff, E.A. (1996). Confirming allusions: Toward an empirical account of action. *American Journal of Sociology*, 102, 161–216. doi:10.1086/230911
- Scott, P.H., Mortimer, E.F., & Aguiar, O.G. (2006). The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school. *Science Education*, 90, 605–631. doi:10.1002/sce.20131
- Shein, P. (2012). Seeing with two eyes: A teacher's use of gestures in questioning and revoicing to engage English language learners in the repair of mathematical errors. *Journal for Research in Mathematics Education*, 43, 182–222. doi:10.5951/jresmetheduc.43.2.0182
- Shintel, H., Nusbaum, H.C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language*, 55, 167–177. doi:10.1016/j.jml.2006.03.002
- Stevens, R., & Hall, R. (1998). Disciplined perception: Learning to see in technoscience. In M. Lampert & M.L. Blunk (Eds.), *Talking mathematics in school: Studies of teaching and learning* (pp. 107–149). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511571251.007
- Stone, C.A. (1993). What is missing in the metaphor of scaffolding? In E.A. Forman, N. Minnick, & C.A. Stone (Eds.), *Contexts for learning: Sociocultural dynamics in children's learning*. New York, NY: Oxford University Press.
- Streeck, J. (2008). Depicting by gesture. *Gesture*, 8, 285–301. doi:10.1075/gest.8.3.02str

- Streeck, J. (2009). *Gesturecraft: The manufacture of meaning*. Amsterdam: Benjamins. doi:10.1075/gs.2
- Streeck, J. (2013). Interaction and the living body. *Journal of Pragmatics*, 46, 69–90. doi:10.1016/j.pragma.2012.10.010
- Streeck, J., Goodwin, C., & LeBaron, C. (2011). Embodied interaction in the material world: An introduction. In J. Streeck, C. Goodwin, & C. LeBaron (Eds.), *Embodied interaction: Language and body in the material world* (pp. 1–26). New York, NY: Cambridge University Press.
- Strom, D., Kemeny, V., Lehrer, R., & Forman, E. (2001). Visualizing the emergent structure of children’s mathematical argument, 25, 733–773. doi:10.1016/S0364-0213(01)00052-0
- Vygotsky, L.S. (1986). *Thought and language*. (A. Kozulin, Ed.). Cambridge, MA: MIT Press.
- Wardak, D. (2017). Encapsulating concepts in gestures and drawings during educational design team meetings. *Mind, Culture, and Activity*, 24, 47–66. doi:10.1080/10749039.2016.1169549
- Wertsch, J. (1980). The significance of dialogue in Vygotsky’s account of social, egocentric, and inner speech. *Contemporary Educational Psychology*, 5, 150–162. doi:10.1016/0361-476X(80)90036-3
- Wertsch, J.V. (1985). *Vygotsky and the social formation of mind*. Cambridge, MA: Harvard University Press.
- Wittmann, M., Flood, V., & Black, K. (2013). Algebraic manipulation as motion within a landscape. *Educational Studies in Mathematics*, 82, 169–181. doi:10.1007/s10649-012-9428-4
- Yifat, R., & Zadunaisky-Ehrlich, S. (2008). Teachers’ talk in preschools during circle time: The case of revoicing. *Journal of Research in Childhood Education*, 23, 211–226. doi:10.1080/02568540809594656
- Zemel, A., & Koschmann, T. (2014). “Put your fingers right in here”: Learnability and instructed experience. *Discourse Studies*, 16, 163–183. doi:10.1177/1461445613515359