Conference Proceedings

Department of Medical Education

8-2001

Dissecting Common Ground: Examining an Instance of Reference Repair

Timothy Koschmann
Southern Illinois University Carbondale

Curtis LeBaron

Charles Goodwin
UCLA

Paul Feltovich
Southern Illinois University Carbondale

Dissecting Common Ground:
Examining an Instance of Reference Repair

Timothy Koschmann
(**tkoschmann@siumed.edu**)
Medical Education, 801 N. Rutledge
Springfield, IL 62711 USA

Curtis LeBaron
(**lebaron@stripe.colorado.edu**)
Communication Department, Campus Box 270
Boulder, CO USA

Charles Goodwin
(**cgoodwin@humnet.ucla.edu**)
Applied Linguistics, 3300 Rolfe Hall
Los Angeles, CA 90095 USA

Paul Feltovich (pfeltovich@siumed.edu)
Medical Education, 801 N. Rutledge
Springfield, IL 62711 USA

Abstract

How participants to a joint activity come to develop a shared or mutual understanding of what they are perceiving has long been a problematic issue for philosophers, sociologists, and linguists. We examine the abstract model proposed by Clark and Marshall (1981) whereby speakers and hearers construct mutual knowledge and by which discrepancies in definite reference are repaired. We focus in particular on forms of demonstrative reference that depend upon physical co-presence. We examine an attested example of reference repair in the operating room of a teaching hospital. It involves learning to recognize pertinent structures within endoscopic surgeries, that is surgeries in which internal spaces are rendered visible by inserting a fiber-optic lens into the body of the patient. Clark and Marshall provide a useful vocabulary for discussing referential practices in this applied setting. We are left with some questions about how to interpret certain features of their model, however. We conclude that further theoretical framing is required before we develop a full appreciation of how reference and reference repair is accomplished in day-to-day interaction.

How participants to a joint activity come to develop a shared or mutual understanding of what they are perceiving has long been a problematic issue for philosophers, sociologists, and linguists (cf., Heritage, 19884; Lewis, 1969; Sperber & Wilson, 1986; Stalnaker, 1978). One means of building "common ground" (Clark, 1996), of course, is through demonstrative reference. Even here, however, potential problems abound. When one issues the utterance "It's right here," how is it that one assures oneself that what is presented as *here* is the same as what is taken as *here* by the listener? Further, how do we detect when discrepancies have arisen and how are these discrepancies to be reconciled? We begin this paper by reviewing the pragmatic model of reference repair proposed by Clark and Marshall (1981). We then examine an instance of reference repair in an applied setting to evaluate the usefulness of this model in understanding actual referential practice.

Clark and Marshall's Model of Reference Repair

Clark and Marshall (1981) proposed an abstract model for the repair of direct references based on their proposal for how mutual knowledge is constructed. This proposal can be expressed succinctly by the following formula:

Evidence + Assumptions + Induction Schema = Mutual Knowledge[1]

where *evidence* is the grounds for the speaker and hearer's belief that both understand some matter in the same way, *assumptions* are the things taken for granted when accepting these grounds as warrants, and *induction schema* is a recursive formulation of Lewis' (1969) iterative definition of common knowledge. By this formula, evidence and assumptions are interrelated in that weaker bases of mutuality must be compensated by increasing levels of assumption. Clark and Marshall's taxonomy of evidence is broken into three categories: community membership, physical co-presence, and linguistic co-presence. [2] These evidence types, along with their associated assumptions are listed in Table I.

Mutual understanding proceeds on the assumption that speakers and listeners are each members of many different cultural communities (e.g.,

---

[1] In later writing (see Clark, 1996), *mutual knowledge* was expanded to *common ground*, a broader notion that subsumed mutual belief, mutual knowledge, mutual assumptions, and mutual awareness.
[2] Clark and Marshall (1981) listed indirect co-presence as a fourth category of evidence. For ease of presentation, we have condensed the categories into three.

Table 1: Bases of Common Ground
(adapted from Clark & Marshall, 1981)

Evidence	Associated Assumptions
1. Community membership	co-membership, universality of knowledge
2. Physical co-presence	
a. Immediate	simutaneity, attention, rationality
b. Potential	assumptions of 2a. + locatability
c. Prior	assumptions of 2a. + recallibility
d. Indirect potential	assumptions of 2b. + associativity
e. Indirect prior	assumptions of 2c. + associativity
3. Linguistic co-presence	
a. Potential	assumptions of 2b. + understandability
b. Prior	assumptions of 2c. + understandability
c. Indirect potential	assumptions of 3a. + associativity
d. Indirect prior	assumptions of 3b. + associativity

African Americans, soccer fans, Presbyterians, pipe fitters, speakers of French) and that membership in these communities imparts special forms of shared vocabulary and knowledge. Reference based purely on community membership assumes that the speaker and hearer hold one or more of these cultural communities in common (i.e., co-membership) and that the object of reference is known to all members of these shared communities (i.e., universality of knowledge). Clark and Marshall theorized that mutual knowledge based on community membership has an extended scope and can be carried from one conversation to another.

A second form of evidence is based on physical co-presence. When speaker and hearer are aware of an object present to both at the moment of reference (sometimes referred to as "triple co-presence"), the situation is labeled *immediate co-presence*. Although this is the strongest form of co-presence for Clark and Marshall, it too has certain assumptions. The speaker assumes that the listener is not only oriented to the object, but is also attending to it (attention) and that both are attending to it at the same time (similtaneity). It also assumes that the listener possesses the faculties to appreciate the meaning of the utterance (rationality). If only the speaker is focusing on the object, but it is available to the hearer (i.e., locatability), *potential physical co-presence* is established. If the hearer does not happen to be attending to the object of reference, but is known to have attended to it previously and can be counted upon to remember it (recallibility), then *prior physical co-presence* can be established. Attributes of components of physically co-present objects can be referred to indirectly provided the hearer recognizes (via community co-membership) the semantic links connecting the attribute or component of the object to the object (assumption of associativity).

The third category of co-presence is *linguistic*. It allows for reference to objects that have been previously introduced into the conversation. Such forms of co-presence are only prior or potential, depending on whether the object is introduced earlier or later in the stream of talk. Both types depend upon a form of assumption Clark and Marshall refer to as "understandibility." As with physical co-presence, more complex forms of linguistic co-presence are possible through association. Unlike community co-membership which is sustained over long periods, Clark and Marshall considered physical and linguistic co-presence to have relatively brief temporal extent.

As evidence for their model, Clark and Marshall direct attention to the way that speakers repair definite references. They described two forms of reference repair: *horizontal* and *vertical*. Horizontal repairs involve enhancing reference by providing additional information without altering the set of underlying assumptions. Vertical repair, on the other hand, involves advancing to a level of co-presence with fewer assumptions. For example, moving from an indirect form of co-presence to a direct form or moving from potential to immediate co-presence or shifting from linguistic to physical co-presence. Because community co-membership has assumptions that are entirely different from those underlying physical and linguistic co-presence, it allows only for horizontal forms of repair.

The model of reference repair presented by Clark and Marshall was largely linguistic. Clark (1996) later elaborated on the notion of common ground. He made a conceptual distinction between *communal common ground*, something that rests largely on community co-membership, and *personal common ground*, with a correspondence to what has been previously described as physical and linguistic co-presence. He expanded his treatment of personal common ground to include "joint perceptual experiences" and "joint actions" (p. 112), that is gesticulation, observed actions, and other

features of the social setting in addition to talk. As we turn to an instance of actual reference repair in an applied setting, we see the importance of taking a broader and more situated view of referential practice. In particular, we begin to see some of the complexities embedded in certain features of Clark and Marshall's model, such as the assumption of locatability.

Analyzing Reference in an Endoscopic Surgery

The setting within which we have chosen to study referential practice is the operating room (OR) of a busy teaching hospital. Within this context, there are multiple forms of work being performed simultaneously. On the one hand, there is a cycle of activity surrounding the performance of a particular surgical procedure itself within which each of the members of surgical team plays a specific role. At the same time, there is instructional work to be done as well. In the fragment to be analyzed here, one participant ("Attending") is a highly-experienced surgeon, ultimately responsible for the safe and successful outcome of the surgery. A second ("Resident") is a surgeon in the final year of his surgical residency, who had by his own estimate participated in 80 to 90 surgeries of the type to be described here (by comparison, the attending surgeon reported that he has performed 1200-1300 of these surgeries over the course of his career). The remaining participant ("Clerk") is a third-year medical student enrolled in a clerkship rotation. This was his first surgical experience. Attending, therefore, is providing guidance and supervision to the resident and both Attending and Resident are responsible for providing instruction to the medical student.

The surgical procedure in which they are engaged is a laparoscopic cholecystectomy, that is the removal of the gall bladder with the aid of an endoscopic camera. Such surgeries were of interest to us because of the manifold challenges to perception and coordination that they pose to participants. Surgeons are called upon to translate what they see on a 2-D TV monitor into a model of what is happening within the not directly inspectable belly of the patient. The image seen on the screen is a magnified view that facilitates precise manipulation on the part of the surgeon, but can be disorienting for newcomers. The orientation of the view on the screen is arbitrary, though the convention is to orient the lens in such a way that the projected image most closely resembles what would be seen in an open surgery (that is a ventral view in which up is anterior and down is posterior). Since participants on opposite sides of the operating table observe different monitors, however, the person assisting the surgeon from the opposite side of the table receives an inverted view.
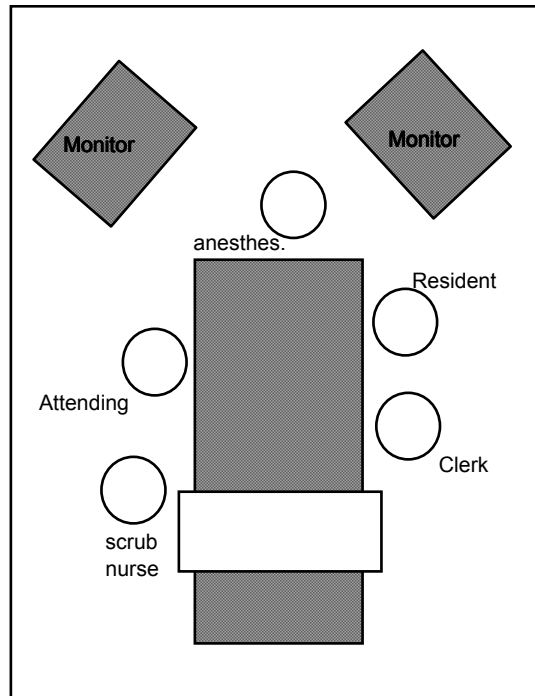

Figure 1. Arrangement of the operating room.

Looking within the endoscopic space is a team effort with different members responsible for operating the camera, "retracting" obstructing organs, and conducting the surgery. This requires substantial coordination in that a view of the workspace adequate to carry out the procedure can only be achieved if all members of the team correctly anticipate the needs of the surgeon. Although the participants work in close proximity to one another, many of the normal resources for effecting mutual orientation are not available to them. Their hands, for example, are occupied much of the time and, as a consequence, cannot be employed for gesture. Further, because they do not attend to the same monitors and because the monitors are located at a distance from where they work, it is difficult for them to use each others' gaze as a cue for orientation as is often done in more typical face-to-face interaction (Goodwin, 1986; Streeck, 1993, 1994).

Attending assists the resident from the left side of the table (see Figure 1). Clerk, standing to the left of Resident on the right side of the table, controls the rod lens of the endoscopic camera. The surgery is considered routine. It consists of isolating the small duct (the cystic duct) through which the gallbladder empties into the common bile duct and the vessel (the cystic artery) that supplies the gallbladder with blood, ligating both with surgical clips, and severing them. The gall bladder is then gently teased from the liver and extracted through one of the "ports" in the abdominal

wall. The greatest technical challenge is correctly identifying the cystic duct and cystic artery, as serious post-surgical complications may arise were clips to be applied to the wrong structures.

Repairing Reference in the OR

Space restrictions prevent us from presenting here a full analysis of the interaction. A more detailed analysis of the fragment can be found elsewhere (Koschmann, Goodwin, LeBaron, & Feltovich, in prep). A transcript can be found in Appendix A. [3] It begins (lines 1-9) with Attending describing the surgical procedure to Clerk. At the same time and throughout the course of this interaction, the resident was performing a blunt dissection to expose the cystic duct and the cystic artery using the tool in his right hand (a "black grasper"). This dissection was performed by burrowing the tip of the grasper into a bundle of connective tissue binding the bottom edge of the gall bladder to the common bile duct and then gently spreading apart the jaws of the instrument. Attending and Clerk observed his progress on their respective monitors.

The expression *cystic artery* is introduced here for the first time (line 2). In terms of Clark and Marshall's model, Attending's use of this expression is authorized by Clerk and Attending's membership in some common community. Attending displays by his choice of language a set of presuppositions about what would be understandable to a third-year medical student. Resident's first demonstrative reference to the cystic artery (line 5) specifies a region in which the cystic artery can be found, though it may not necessarily be visible at the moment in which he makes the reference. In Clark and Marshall's terminology, therefore, these references signal potential physical co-presence. This raises interesting questions about what the assumption of locatability means in this particular situation, however. If it means that the cystic artery is simply available to Clerk's viewing, Resident's utterance would suggest that he believed the cystic artery to be locatable at the moment of reference. If one has never seen a cystic artery on an endoscopic display, however, is it still locatable there?

Clerk's query in line 10 makes visible his orientation to unfolding process. The cystic artery may or may not be visible at that point in time, but his use of the adverb *yet* expresses a confidence that it will eventually be made manifest to all. Attending's reply in line 13 ratifies this view. Like Goodwin's (1999) archeologists excavating through sedimented strata of soil, surgeons must dissect through various layers of anatomical structure. They speak of *planes of dissection*, meaning the surfaces available to sight at specific junctures within a procedure. In an endoscopic surgery, however, the cystic artery will never be physically co-present in the same way that it would in an open surgery since its presence is mediated through a video viewing system. Attending's deictic particle *here*, therefore, anchors not to the conventional origo of the speaker's corporal location, but rather to a virtual origo located in the shared media space.

Resident eventually provides six separate demonstrations of the cystic artery before receiving a tentative sign of recognition on the part of Clerk (line 19).[4] Learning to locate pertinent structures on the video display is an important aspect of "professional vision" (Goodwin, 1994). Resident's *there* (line 18) was coordinated with a point to a white stripe within the bundle of connective tissue being viewed. Although gesture is often characterized by linguists as supplementing speech, Hindmarsh and Heath (2000) described instances in which "The deictic term segments the gesture, displaying just the moment at which it is sequentially relevant" such that "the talk reflexively works on behalf of the gesture" (p. 15).

Resident's repeated efforts to demonstrate the cystic artery, could be described in terms of Clark and Marshall's model of reference repair as an attempt to eliminate the assumption of locatability. That is, he was striving to promote his shared knowledge with Clerk from potential to immediate co-presence. But what does it really mean to be "locatable." The whole idea of "professional vision" is to acquire the ability to see as presumably more-skilled others can see. If locatability assumes not only that the listener can see (in the sense of having adequate vision, an unblocked view, etc.) what is visible to the speaker, but must also be able to see in the same ways as the speaker (i.e., share the speaker's "professional vision"), then it becomes a very complex kind of assumption, in many ways just as complex as the thing it sets out to explain, namely mutual understanding.

As the fragment continues, Attending raises some concerns about Resident's identification of the cystic artery. On paper, Attending's "That may be right" (line 27) might be construed as a tentative positive appraisal. Resident's reply (line 29), however, treats it as an incomplete utterance, as in "That may be right [hepatic]." Resident's efforts to achieve mutual understanding with Clerk, therefore, have revealed a potential discrepancy in understanding among Resident and Attending. The fragment concludes with Resident

---

[3] The transcription conventions used here are described in Atkinson and Heritage (1984).

[4] Resident's demonstrative reference in line 3 ("Right there") is heard to be referring to the cystic duct, a topic of discussion prior to the transcribed segment, rather than the cystic artery.

and Attending resolving to search further for the cystic artery.

Discussion

Here in a nutshell we see the problem of mutual knowledge. Resident takes some pains to demonstrate to Clerk what he (Resident) believes to be the cystic artery. After some prompting, Clerk declares that he now sees it. Other than his avowal, however, we have no evidence that he indeed sees what Resident has taken such trouble to display. In demonstrating for Clerk what he has taken to be the cystic artery, however, Resident has inadvertently made visible a discrepancy in his presumed common ground with Attending (or, at the very least, a difference in their levels of confidence that the indicated structure is in fact the cystic artery). Clark (1996) defined *grounding* as establishing a claim "as a part of common ground well enough for current purposes" (p. 221). For the purposes of Clerk's instruction, the exchange would seem to have provided ample grounding for his understanding. However, for the purposes of conducting a safe surgery, the concerns raised by Attending might suggest that more grounding is required.

Clark and Marshall provide a useful vocabulary for discussing referential practices in this applied setting. Their model of reference repair, however, hinges upon a calculus of assumption maintenance and herein lies the rub. The conceptual difficulties of mutual knowledge that their model was meant to address have not been completely dispelled, but, instead, arise in new forms when we look more carefully at the underlying assumptions. As we have seen, the assumption of locatability can be quite complex when examined *in situ*. We are in full accord with Clark's shift from a treatment of reference as a simple matter of linguistic interpretation to a more situated model that encompasses "joint actions" and "joint perceptual experiences" and we think that this will lead to a richer understanding of concepts like locatability. For one thing, it would help to illuminate how participants' own unfolding activities contribute to the determinant sense of what is seeable at any given moment. Furthermore, we have much to learn about the interactions between different kinds of bases of shared understanding. Professional vision, for example, draws upon the associated assumptions of both community membership and physical co-presence.

In a situation in which the establishment of common ground is essential, we see just how elusive shared understanding can be to achieve. Our analysis of the fragment of interaction in the OR would suggest that we have a way to go before fully appreciating how these factors enter into our day-to-day practices of reference and reference repair.

References

Atkinson, J.M. & Heritage, J. (1984). Transcription notation. In J.M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. ix–xvi). NY: Cambridge University Press.

Clark, H. (1996). *Using language*. NY: Cambridge University Press.

Clark, H. & Marshall, C. (1981). Definite reference and mutual knowledge. In A.K. Joshi, B.L. Webber, & I.A. Sag (Eds.), *Elements of discourse understanding* (pp. 10-63). NY: Cambridge University Press.

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

Goodwin, C. (1994). Professional vision. *American Anthropologist*, *96*, 606-633.

Goodwin, C. (1999). Pointing as situated activity. Paper presented at meeting of Society for Social Studies of Science.

Heritage, J. (1984). *Garfinkel and Ethnomethodology*. Cambridge, U.K.: Polity Press.

Hindmarsh, J. & Heath, C. (2000). Embodied reference: A study of deixis in workplace interaction. *Journal of Pragmatics*, *32*, 1855-18788.

Koschmann, T., Goodwin, C., LeBaron,, C., & Feltovich, P. (in prep). A mediated conversation on the cystic artery: Reflexivity and the 'inscrutability of reference'. *Human Computer Interaction.*

Lewis, D. (1969). *Convention*. Cambridge, MA: Harvard University Press.

Sperber, D. & Wilson, D. (1986). *Relevance*. Cambridge, MA: Harvard University Press.

Stalnaker, R.C. (1978). Assertion. In P. Cole (Ed.), *Syntax and semantics: Pragmatics, Vol 9* (pp. 315-332). NY: Academic Press.

Streek, J. (1993). Gesture as communication I: Its coordination with gaze and speech. *Communication Monographs, 60*, 275–299.

Streek, J. (1994). Gesture as communication II: The audience as author. *Research on Language and Social Interaction, 27*, 239–267.