Social Interaction

Video-Based Studies of Human Sociality

_____

Achieving mutual accessibility through the coordination of multiple perspectives in open, unstructured landscapes

Michael Sean Smith

Linköping University

Abstract

Whenever actors perceptually engage with the surrounding world in concert with others, they routinely attend to the degree to which their perceptions (whether visual, aural, tactile, etc.) do or do not overlap with their co-participants. In making a perception publicly accessible then, participants must not only attend to potential perceptual gaps, but have-at-hand a range of discursive and embodied practices for closing those and making what is perceived by one mutually accessible to others. In this paper, using data collected from a geological field-school, I investigate the embodied and mobile practices that participants use for coordinating perception via perspective in open, wilderness settings. I focus in particular on the visual practices that participants use for making what one "sees" in the landscape or activity "seeable" for others. These practices are in turn analyzed with regard to how they highlight the camera's role in documenting the embodied means by which these practices work. In the analysis of data, we will see the participants' perspective or line of sight, i.e., the axis of their gaze become a more explicit and salient feature for coordinating the interaction. Field geology provides a perspicuous setting for not just investigating how participants reconfigure themselves vis-a-vis local features in the landscape in order to perceive those features, but also for examining the relationship between the videographer's perspective as documented on camera and that of the participants.

1. Introduction

Video has become indispensable in the analysis of human conduct and interaction. Its proliferation, however, has come with many unexamined assumptions with regard to the video recording and what it does or does not capture. Careful consideration is often given to issues like camera placement, how the participants are framed, and what is included on- versus off-screen. Even where conditions are optimal, however, no video recording should be treated as a transparent, objective record of what happened (Goodwin, 2000a, p. 160; Mondada, 2009, p. 68). Any recording only captures a version of a given event or interaction and does so from a very restricted point of view—one that inadvertently flattens both the spaces which our participants inhabit and the viewpoints of those participants. While this may not be problematic when analyzing video-recorded interactions that occur in more domestic and built environments (e.g., homes, offices, etc. with tables, desks, seating, etc.), it may be a problem when analyzing interactions that occur in unstructured, open setting, where the participants are continuously mobile, such as is the case in the setting analyzed here: field geology. What is perceptually accessible in such settings cannot always be taken for granted even by the co-participants themselves, and in turn, extensive embodied, mobile work is often required in achieving a mutually shared perception of features in the landscape.

In this paper, I examine the work performed by interactants in mobilizing others to see either features in the landscape or aspects of their current activity which, though ostensibly "co-present", are nonetheless not accessible to all the interactants. This interactional work, though concerning "visuality", is similar to the work seen in studies on the use of other non-aural/visual senses in interaction (e.g., touch, taste, smell, etc.); it operates through a wide range of embodied and multimodal practices, all geared towards making the intended feature accessible and recognizable to others, often in the pursuit of some larger course of action. In the analysis of data presented here, we will see the participants' line of sight; that is, the axis of their gaze, become a much more explicit and salient feature in accomplishing this work. Such work requires a much more detailed analysis on the participants' part of the local and distal spaces in which they are situated—details that may be difficult to capture on video. This necessarily complicates how one effectively reconstructs this work via video.

1.1 Background

Studies in ethnomethodology and conversational analysis are generally concerned with the public ratification of action and perception in interaction—that is, how talk, gesture, and situated understandings are mutually recognized by co-participants. All situated action and perception are mediated in part from the perspective of experiencing, sensing actors who, when perceiving the surrounding world in concert with others, routinely display their awareness of the degree to which these perceptual engagements (whether via sound, smell, sight, or taste) may or may not overlap with their co-participants. In order to make those sensorial experiences publicly ratifiable, actors must not only attend to what is and is not accessible or perceivable for their co-participants and recognize potential gaps, but have a range of discursive and embodied practices for closing these gaps and making those percepts accessible to one another. Publicly ratifiable actions in multi-modal and multi-sensorial domains in turn must always be perceptually accountable to others' embodied points of view.

Recent work in multi-sensoriality in interaction has attested to this perceptual accountability in a range of different sensory modalities, from smell (Mondada, 2020), to taste (Mondada, 2018a; 2019), to touching objects (Mondada, 2016; Goodwin & Smith, 2020; Kreplak & Mondémé, 2014). When participants achieve or maintain a mutual sensory experience of some co-present object, they do so through a wide range of embodied, and multimodal practices. The ways in which these are organized as orderly not only relates to how participants experience an object but how that experience is made accountable to others as a recognizably sensorial practice. Thus, the sensorial experience of co-present objects is continuously made witnessable via its public manifestation in interaction with others (Mondada, 2019a). Moreover, when the activities that participants are engaged in are necessarily constituted through some sensorial practice (i.e., smelling and touching cheese in a shop, wine tasting, etc.), establishing an intersubjectivity around the sensory experience (whether via touch, taste, smell, etc.) becomes a continual focus for the participants' interactive work.

While we tend to assume visuality, as oppose to other senses, is relatively unproblematic (at least among sighted participants), there exist many contexts where participants being able to easily determine that they are seeing the "same thing" is not always so given and must instead be achieved via interactive work. In such contexts, we see interactants deploy a wide variety of "visual practices" (that is, the practices used for showing, looking, and seeing things in the current setting) for building and sustaining common foci of attention (Mondada, 2019, p. 64). These practices have been particularly studied in contexts like guided visits in museum, exhibitions, and botanical gardens, where attending to features in the co-present environment plays a constitutive role in the participants' activity (Broth & Lundström, 2012; Mondada, 2012; 2019b). Establishing a common focus of attention presents a recurrent practical problem for the participants, which in turn requires a continual reorganization of the interactional space (that is, the

arrangement of the participants' bodies, their orientations, and attended-to objects in the setting). In doing so, we observe participants deploying a number of embodied, mobile, and multimodal practices in making the co-present feature visually accessible to others. Similar to the work previously described on sensory modalities like taste, touch, smell, etc., the ways in which a given feature is visually located and recognized by a participant witnessable and thus provide for the public accountability of having recognized that feature. Bringing co-present features to the forefront of an interaction is in turn inseparable from the interactive, visual practices that participants use for accomplishing this.

Adequately capturing this type of interactional work on video (or reconstructing it afterwards) can be difficult. As new foci of attention emerge and participants rearrange themselves in space, these changes can be abrupt and unexpected, making it difficult for the videographer to follow where the action is at any given moment (Mondada, 2019b, p. 99). Additionally, when participants point out some feature in the landscape or otherwise make something visually accessible to a co-participant, they do so specifically with regard to the situated point-of-view of a co-present participant, and not to the co-present videographer. While the videographer may be an observer and treated as "actively present", they are not typically treated as a ratified addressee (*ibid*, p. 98). Moreover, when considering the activities analyzed here, the videographer is not a part of the participants' collaborative project. As such, the videographer's engagement in the activity is largely orthogonal with respect to the courses of action that participants are pursuing: They are not accountable for recognizing the presence of events or features within the activity (outside of their recording); nor do the participants actively make those accessible to the videographer.

This might not present an obstacle in more mundane and institutional settings, where the actors are sedentary, where the majority of actions are performed via talk, or where the co-present features being referred to are easily recognizable to the analyst. It does present a problem in activities and settings found in field geology. Here, participants engage in a collaborative, technical investigation, where they actively move through open wilderness settings, for the purpose of locating and documenting geological features about which the participants themselves may be uncertain. Moreover, these co-present features can range in size and distance, from the small and immediately local to the expansive and distant. Participants are furthermore continuously mobile in relation to these features, allowing for a radically expanding set of contextual configurations between the participants, co-present features, and the immediate and distant spaces in which these are situated (Goodwin, 2000b). The range of scales that participants operate within, the fluidity in which they shift from one to another, and the uncertainty that they often display in scrutinizing various features makes capturing their work difficult from the single point of view of the camera. What appears as notable for the participants is often unpredictable—not just for themselves, but especially for the videographer.

This type of setting makes it difficult to capture both the participants as well as the features at which they are looking. Some of these issues can be introduced with the vignette below. Here, we have three geology students working on a project on mapping and identifying folds over a large geographic area (*n.b:* S1, S2, etc. designate sedimentary layers, "foliations", that have been deformed, i.e., "folded", through metamorphic processes). The participants have just arrived at a rock outcrop, and as we join the action, Tyler describes his difficulties in seeing structures as he approaches:

Figure 1: Tyler, Drew, and Kyle at a rock outcrop

```
01  Tyler:  It's looks like we're seeing an S1 and an S2 though.
02          It looks like it has, but then when you actually
03          look at the rock, you don't see any internal structure.
04             (1.4)
05  Tyler:  But it's visible when you walk up, which is kind of
06          puzzling.
```

What Tyler sees in the rock changes as he walks towards it. As he explains it, whereas he thought he could see S1 and S2 foliations as they approached the rock, these "disappeared" when he is immediately in front of it. This is a common problem for geologists: Visible differences in rock types and structures are often better viewed from a distance, and a more experienced geologist will often make note of different structures from a distance before moving closer to analyze the rock in detail.

How the co-participants in this setting move through a landscape and operate on co-present features is relevant to how they interact and build action within that space. Early work in video studies of talk-in-interaction primarily focused on interaction within "physically constrained" locales (Haddington et al., 2013, p. 23) and relatively "stationary interaction" and tended to focus on how participants

organized their bodies relative to one another (e.g., "F-formations" in Kendon, 1990), rather than to "the material surround" (Broth & Lundström, 2013, p. 92). New domains of interaction emerge, however, when analyzing how participants incorporate the material surround as a problem, resource, and practical outcome of the interaction. The shifting physical and interactional configuration between participants, the terrain, and co-present features in turn provide new resources for use and re-use in the subsequent interaction.

Bodily arrangements in space in different configurations can alternatively integrate or disintegrate "...common space[s] of action and multiple dynamic rearrangements of the bodies" (Haddington et al., 2013, p. 21) which in turn provides a resource for building action towards co-present features. When participants move with reference and attention to that environment, the movement itself is made accessible as meaningful action: it can be parsed and decoded with regard to its facilitating social action, particularly with reference to the work-related pursuits of the participants, and organized as an accountable practice that is routinely oriented with reference to objects or locales in the environment. Charles Goodwin (2000b) referred to the arrangement of bodies, space, and actions as a *contextual configuration*; that is, the "particular, locally relevant array of semiotic fields that participants demonstrably orient to…[and]…which frame, make visible, and constitute the actions of the moment"…[while providing]…a systematic framework for investigating the public visibility of the body as a dynamically unfolding, interactively organized locus for the production and display of meaning and action" (*ibid.,* p. 1490).

As suggested in Tyler's observation, however, what he and his co-participants see is indexically tied to their movement through the landscape. For Ingold (2004, p. 330) any analysis of the body in motion begins with walking: It is through our movement on the ground "...against which things 'stand out' as foci of attention…[and emerges as]...a focus in itself." Similarly, Gibson (1979, p. 197) argues with regards to visual perception and movement that "...the forms of the objects we see are specified by transformations in the pattern of reflected light reaching our eyes as we move about in their vicinity. We perceive, in short, not from a fixed point but...[from]...a 'path of observation', a continuous itinerary of movement." For participants, features emerge from the landscape as "condensations or crystallizations" of mobile activity rather just being forms superimposed on a material substrate (Ingold, 2004, p. 333). Locomotion as such precedes and makes possible the participants' "perceptual work," *that is, seeing categorically-relevant, geological structure in the landscape.*

Tyler's noticing raises issues when attempting to capture on video what is relevant for the participants in these contexts. From the videographer's point of view—myself being interested in embodied action, orientation, gaze, gesture, and mobility—I prioritized capturing as much of participants as possible, preferably from the front, in their immediate surroundings, while simultaneously minimizing the degree to which features (and other participants) might obscure the

participants' actions. While this framing is advantageous for analyzing things like talk, gaze, orientation, embodied action, and mobility, it omits, as a consequence, the actual things that the participants are looking at and, by extension, the interaction between them and their material surround. In Figure 1, for instance, there is no record of the rock face, let alone what they see in the rock. Similarly, as they moved towards the outcrop, priority was given to recording all three participants forward facing in the frame as much as was possible, thereby omitting how they coordinated their movement towards the feature. Not having access to such features in the camera may preclude an analysis of how participants build and transform their contextual configurations in order to incorporate distant features in the landscapes.

In each of the extracts below, we observe one participant's line of sight of a distant feature becoming a salient feature for the interaction, as a point towards which some course of action is launched. Furthermore, in each of the extracts, we observe the analysis benefitting from being recorded from an angle that more or less approximates this participant's line of sight. These recordings come either from the videographer's camera, which just happened to be at the appropriate location (Ex. 1), or from a wearable camera mounted on either the participant's chest or head (Ex. 2 & 3). The first recording, Extract 1, was shot with a handheld camera positioned behind Kyle, facing the same direction towards the feature to which they were referring. Extracts 2 and 3 come from the body-mounted cameras that the participants were wearing, which by design only capture what is to the front of the participant. In each of the extracts, we will observe that while some features of the interaction are lost, we gain insight into how participants align their bodies and action with regard to the larger and more distant material surround.

2. Data and methods

This corpus is drawn from four video-documented ethnographic trips to field-based projects with field geologists. The study participants involved included late- to early-career geologists, graduate students, and undergraduate students in a geology capstone field course. Each of these visits was video recorded while the researcher(s) accompanied the participants in the field documenting how they move through the landscape, find locales of interest, locate and investigate geological objects, make drawings, measurements, or collect samples of geologically relevant phenomena. After the data were collected, the video and audio were transcribed using conversation analysis (Jefferson, 2004) with a focus

on the participants' use of embodied and multimodal action (Goodwin 2007, 2010; Mondada 2016, 2018b).

3. Analysis

The ways in which the participants move and interact within the landscape develop largely from their joint project and its relation to co-present features within that landscape. The resulting contextual configurations that emerge in turn provide for the observability of participants, specifically with regard to the sensibility of their conduct and their courses of action. Through these unfolding configurations, participants attend to how others reveal their engagement with features in the landscape, and those features in turn are incorporated into the participants' embodied work, even when quite distant or large in structure. The more distant or large a co-present feature is relative to the participants, the more important that line of sight becomes when coordinating action. This can be seen in Extract 1.

3.1 Positioning action in others' line of sight

Here we find the same three geology students on the same day measuring the strike and dip (the angle and direction) of the folds over which they are currently standing. Just prior to the transcript, Tyler relayed the measurements for S1 to Kyle, who then repeats them as he records them in his notebook (lines 13 & 14). He and Tyler subsequently reconfirm this in lines 16 and 18. While this is going on, Drew walks over to the far side of the exposed folds. As he reaches the other side (right at the end of Kyle's talk in line 23), Drew looks down the length of the fold and calls the others' attention to something he notices in the rock.

Extract 1 "look down there"

```
S1 - "ess-one"
S2 - "ess-two"

13    Kyle:   Two thirty two sixty si:x
14            °two thirty two sixty si:x.°
15               (3.5)
16    Kyle:   for ess-o:ne?
17               (.)
18   Tyler:   Yeah.
19               (0.8)
20    Kyle:   day two:: stop two::_
21    Drew:   I: just go::t one ess-two::,
22               (2.0)
23    Kyle:   phyl:::i::::te.
24               (0.5)*(1.5)
        dre          *begins to depict curve-->
```

```
25    Drew:  >Dude< th↑ese almo∆†st look#∆
      dre                     ∆--------∆ ((looks to Kyle))
      kyl                    †looks to Drew–>
      fig                              #fig. 2
```


Figure labels: Kyle, Tyler, Drew, *approximate angle of gesture*

2
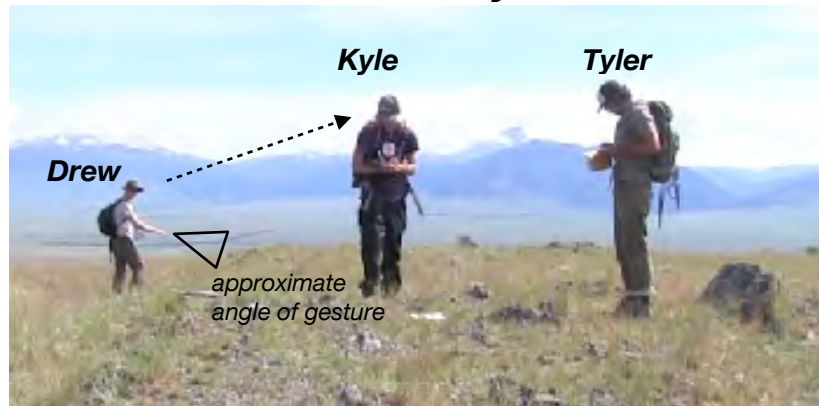
```
26         (1.0)*†
      dre  *-------*
      kyl  †--------†
27    Drew:  #‡°(( ) they're like)°‡#
      dre    ‡increases curve------‡
      fig    #fig. 3              #fig. 4
```


3 **4**

```
28    Kyle:  *I knO:W.=look *#down there.
      kyl.   *.............*points--->
      fig                 #fig-5
```


5

```
29    Kyle:  *=I see one*[%@=that's going%@#this way%
```

9

```
30    Drew:              *[ @I kno:w_ you %@#can ^see% this%@ cu::rve.@
      kyl   *,,,,,,,,,,* % ,,,,,,,,,,,,,%gesture¹-%,,,,,,%
      dre                @,,,,,,,,,,,,,@gesture¹-------@,,,,,,,,,@
      fig                            #fig-6
```



```
31             (1.0)
32    Kyle:  We should >wal::k this_<
33           and take some strike n' dips,
```

¹Both Drew and Kyle use the palm of their hand to depict the plane or
curve of the fold

While formulating his noticing (in lines 25 through 27), *>Dude< th↑ese almost look (1.0) °( ) they're like°"*, Drew positions the back of his flattened hand perpendicular to the length of the exposed fold to depict the angle of the folding. He begins to depict an angle just prior to his talk in line 25. In doing so, he looks toward Kyle while holding his hand at the angle (see Fig. 2) until having Kyle's gaze (toward the end of the 1.0s gap in line 26). Once having that, Drew rotates the angle of his gesture over his subsequent talk in line 27: *°( ) they're like°* (figs. 3 to 4). Kyle responds to Drew's noticing stating, *I knO:W.=look down there* before pointing down the length of the fold (Fig. 5). He subsequently drops his point before gesturally depicting the angle that he sees in line 29: *I see one that's going this way*. Drew simultaneously depicts the curving of the folding in line 30: *I kno:w_ you can ^see this cu::rve.* (Both in Fig. 6). The sequence ultimately leads to Kyle suggesting the next course of action: *We should >wal::k this_< and take some strikes n' dips.*

Both Drew and Kyle's depictions are *environmentally coupled gestures* (ECG) (Goodwin, 2007): Each simultaneously operates on a co-present feature and the concurrent talk incorporating both into a *semiotic complex* for the purpose of drawing out specific aspects of the co-present feature for their interlocutor. When we look at Drew's initial gesture (lines 25-27), in particular, however, we can note that it does more than just depict the folding at his feet. When increasing the angle of his hand (Figs. 2-4) Drew depicts not so much the angle of the folds in front of him, but the degree to which they change over the observable length of the

folding. Being able to effectively depict that change, however, particularly for a recipient who can assess it, necessitates a recipient who is in a position to see both (the gestural depiction and the length of folding) layered over one another. This is demanded in large part by the expansiveness of the folding and how it extends over the landscape toward the horizon. Kyle's line of sight, that is, the axis of his gaze and how it extends out and into the landscape *at a distance* becomes a much more central feature for understanding how the participants coordinate the interaction. As the environment being coupled through the talk and gesture begins to extend beyond the participants' immediate space, the contextual configurations being built through the interaction become more geometric with regard to how the participants position themselves vis-a-vis one another and co-present features.

The analysis in Extract 1 benefits from how the videographer and camera are positioned vis-a-vis the participants and the physical referent in the landscape. In being positioned behind Kyle while he faces down the length of the rock outcrop, we, as analysts, can observe how Drew's gesture is coordinated on the folding as it extended over the horizon and how this was built specifically with regard to Kyle's situated line of sight as he also looks down the length of the folding. Having this specific perspective onto the interaction, though entirely serendipitous at the time, made this coordination accessible in ways that may not have been possible if the camera were positioned elsewhere. If the camera were positioned to either side of the outcrop, for instance, the specific configurations being built might be much more difficult to reconstruct. In being positioned behind Kyle, however, we gain a somewhat parallel perspective onto the interaction and its configuration within the local environment specifically at a point in the interaction when his line of sight becomes integral in building that configuration.

3.2 Repositioning others in one's line of sight

Another way of capturing a participant's line of sight, or at least a proxy thereof, is through recordings made with wearable cameras. A wearable camera by design only captures what is immediately in front of the wearer. Extract 2 was recorded with a wearable camera mounted on one participant's head. Here, three students, Jevrem, Ronald, and Evan, are mapping the profile of the terrain including the large scarp (a steep slope formed through seismic faulting) that can be seen in the background (see Fig. 6 - 10). Jevrem is wearing the head-mounted camera. Evan is standing to Jevrem's right holding the clipboard, and Ron stands in the center of the frame. Mapping the profile consists of recording elevations at five-foot (approximately 1,5 meter) intervals using a survey tape until they reach a flag that Ronald had earlier placed on top of the scarp in the distance (not viewable on the camera). As we join the action, Ronald and Jevrem are about to record their first elevation when Jevrem notes that Ronald is currently not standing in line with the flag in the distance.

Extract 2 "a little more"

```
01    Evan:   Well go cross ways
02            (.)
03    Evan:   cuz we want your-
04    Jevrem: wait @you['re gonna want- you're-
      jev         @........................>
05    Evan:          [you're#gonna want your toes=
06            =to@be at [(.)   @#%(your xxxxxx)
      jev     ...@point up----@ %.............>
      fig                   #fig. 7
```



```
07    Jevrem:          [you're gonna %#want to be over here
      jev     .......................%point down------->
      fig                            #fig. 8
```



```
08    Ronald: (    my feet) where?
09    Jevrem: Ron you're gonna want to [be,%
      jev     ---------------------------%
10    Evan:                             [I just [mean-
11    Jevrem:                                   [*Move.
      jev                                        *moves point>
12    Evan:   your feet* don't have to do that
      jev     ........*holds point--------->
13    Evan:   I just mean mark it
14    Ronald: yeah yeah yeah
```

```
15  Jevrem:  Ron you're gonna wanna head- head ΩΔlike this #way
        ron                                       Δsteps back---->
        ron                                       Ωlooks to jev--->
        fig                                                      #fig. 9
```



```
16           (0.4)
17  Jevrem:  ¶yeah. a little Ω•more. >ld'l •Δmore.<
        eva  ¶looks up to hill; walks off camera behind jev-->>
        ron  ---------------------------Δ
        ron  --------------Ω•-----------• ((•looks to hill))
18           (0.6)§(1.0)
        ron        §steps back>§
19  Ronald:  (that-) that's not (.) ‡our flag #right there
        ron  ‡.....................‡points to hill----->
        jev  ---------------------------------------->
        fig                                       #fig. 10
```

```
20  Jevrem:   +naw::::: I see it +in #the [+background here‡( +  )]
       jev    +.................+point up-+,,,,,,,,,,,,,,,,,,+
       fig                               #fig. 11
```



```
                                                                11
21   Evan:                                 [+yeah you're-  ‡I +mean]=
       ron    ‡---------------------------------------->‡,,,,,,,>
22          =he's right.
       ron    ‡,,,,,,,,,‡
23          †you're about- †you're about right now
       ron    †turns back----†
24  Ronald:   right here
25    Evan:   [yeah
26  Jevrem:   [yeah
```

Here, the sequence begins with one participant, Jevrem, formulating his situated line of sight via talk and embodied action for the purpose of correcting his co-participant, Ronald's placement within the ongoing activity, vis-a-vis the flag in the distance. Shortly after launching his talk in line 4, *wait you're gonna want- you're-,* he first points towards the flag on the scarp in the distance (Fig. 7). He then subsequently redirects his points during Evan's turn in line 6 to the ground just to the left (Jevrem's left) of where Ronald is currently standing, as he relaunches his corrective in line 7: *you're gonna want to be over here* (Fig. 8).

Jevrem's turns occur in overlap with Ronald and Evan's talk (lines 1-14), however. In response, Jevrem makes numerous attempts at getting Ronald's attention, again in lines 9 and 11, before finally getting Ronald's attention in line 15: *Ron you're gonna wanna head- head like this way.* Here, Ronald looks towards Jevrem and begins to step backwards (Fig. 9). Jevrem acknowledges this while prompting Ronald more in 17: *yeah. a little more. >ld'l •more.<* As Ronald continues stepping back in line 17, he turns toward and points to a flag on the hill stating, *(that-) that's not (.) our flag right there* (fig. 10), presumably treating Jevrem's insistence as being based on him looking at the wrong flag. In response, Jevrem quickly points back up to the hill while stating in disagreement in line 20: *naw:::: I see it in the background here (   ).* Ronald eventually drops his point once Evan confirms that Ronald is in line with the flag in lines 21-23: *yeah you're- I mean- =he's right. you're about- you're about right now.* The contextual configuration that Jevrem attempts to maintain is essentially a straight

line projected over the topography. Where Jevrem stands marks the first point and the flag standing in the distance marks the last point. As such, Jevrem has a unique perspective for assessing Ronald's position vis-a-vis line. His embodied line of sight in turn provides Ronald the means for seeing the degree to which his position deviates from the line that Jevrem can see.

Jevrem's pointing throughout the sequence accomplishes more than just indexing locations in the landscape; moreover, they provide embodied demonstrations of his line of sight, whether that was towards the flag in the distance (Figs. 7 and 11) and how that reflects onto the place on the ground where Ronald should be standing (Figs. 8 to 10). In making his line of sight material and thus accessible for Ron, Jevrem provides his co-participant a concrete means for correcting his position vis-a-vis the line. Ronald initially aligns with this by moving back in lines 15 through 18, before subsequently objecting with his own point to the flag he took Jevrem to be using (see Fig. 10). In holding his point along what he sees as the line from him towards the flag in the distance, Jevrem effectively takes something that was only accessible to him and embodies it as a material and accessible structure in the interaction.

The ways in which the interactions unfold in both Extract 1 and 2 are motivated in large part by how they physically map onto the local landscape. In each, the participants are coordinating around a structure that extends quite a distance from their immediate vicinity and moreover does so along a relatively straight axis (this being by design in Ex. 2). As a result, the participants' ability to perceive phenomena within the parameters of the structure or activity makes line of sight, that is, the axis of one's gaze and the direction that it projects into space, a much integral feature in coordinating the interaction. In Extract 2, we see this become particularly salient because it provides the participants, first a means for discovering and displaying a discrepancy in the participants' positioning vis-a-vis the flag in the distance, and, second a means for remedying that discrepancy.

The camera mounted on Jevrem's head provides a unique viewpoint for analyzing the interaction, as we effectively see the world from Jevrem's location—something not typically possible with the videographer. This in turn provides a proxy of sorts for analyzing the interaction from his unique point of reference, which incidentally becomes integral for him in how he launches his course of action as he and the others coordinate the interaction. This point of view is at the same time extremely limited and necessarily omits relevant interactional detail (that is, what is occurring around Jevrem that is not directly in front of him). For instance, we can note Evan's positioning throughout the latter part of the interaction (lines 15-26): Just at the beginning of line 17, as Jevrem continues prompting Ronald to move back, Evan looks up to the hill and then moves off-camera to the right of Jevrem. The sound of Evan's next turn-at-talk (lines 21-23) suggests that Evan moved closer to Jevrem, positioning his line of sight in line with Jevrem and the flag in the distance, so as to check Ronald's positioning, which he subsequently assesses in his turn-at-talk in lines 21-23:

yeah you're- I mean he's right. you're about- you're about right now. While this interactive work would support the arguments being made here, very little of it is actually preserved on the video. We will, however, see a very similar form of repositioning in another's line of sight in the following extract.

3.3 Positioning oneself in other's line of sight

In the previous extract, we observe a participant using his own line of sight in order to move his interlocutor into line with a flag that he can see in the distance. Moving oneself into others' lines of sight also provides participants a means for coordinating the interaction, particularly when getting an interlocutor to see something in the landscape or confirming that multiple participants are indeed seeing the same thing. In Extract 3, three students, Trent, Jevrem, and Tracy, are mapping metamorphic folds running through an expansive and uneven landscape (the same project as in Extract 1). Just prior to the extract, Jevrem and Trent (who is wearing a chest-mounted camera) had been trying to locate a marble outcrop that they had visited the day before, when Tracy joins the conversation. As the transcript begins, Trent (not seen in Fig. 12 due to his wearing the camera) walks towards Jevrem and points to a hill in the distance, stating *see:_ we were sitting there yesterday* in lines 1 to 3 (Fig. 13).

Extract 3 - "down over there"


```
        fig      #fig. 12
01      Trent:   #see:_
        tre      >>walks forward-->
02               (1.6)
03      Trent:   we were sitting @down there#@
        tre                      @...........@points--->
        fig                                  #fig. 13
```

```
04                   *@(0.4)@(1.4)*
         tre         *------------*
         tre          @,,,,,@
05    Jevrem:    sitting down there,
06     Trent:    yeah†
         tra            †gazes towards Trent's point->
07                     (.)
08    Jevrem:    or is that †them
         tra        †----------†
09                    (1.0)
10     Tracy:    yeah:: +that's the marble down over there#+
                        +...............................+
         fig                                        #fig. 14
```


14

```
11               +   (0.7)+(0.3)+
         tra     +points-+,,,,,+
12    Jevrem:    that's the marble
13                  (0.3)
14     Trent:    where:::,
15                  (0.3)
16               over there.
17               $(0.3)$    (0.3)    $Δ# (2.4)
         tre     $.....$points--,,,$
         tra                        Δbegins walking-->
         fig                          #fig. 15
18    Jevrem:    yeah:::,
19                #(0.4)
                 ∂pivots to point-->
         fig      #fig. 16
```


15

16

```
20      Tracy:    the marble's ∂#over [there
21      Jevrem:                     [the marble's right there∂
        tra    ∂-----------∂holds point-------------------∂,,,>>
        fig                    #fig. 17
```



```
22      Trent:    [m:mh
23      Jevrem:   [yeah, so dude I think we- we weren't
24                  (0.2)
25      Jevrem:   .hhh
26                far enough over
27      Jevrem:   nhhhhh
28                  (2.4)
29      Jevrem:   or wo- one ridge shy
30       Trent:   okay, so push back another contour line
31                  (1.0)
32      Jevrem:   yeah
```

Both Jevrem and Tracy attend to Trent's talk. Tracy looks up towards Trent, and Jevrem turns and points towards where Trent had previously pointed while recycling Trent's *sitting down there* in line 5, which Trent confirms in line 6. In line 8, Tracy turns towards the place originally located by Trent, points, and responds with agreement in line 10: *Yeah, that's the marble down over there* (Fig. 14). Trent's response in line 14-16*,* however, rather than accepting Tracy's confirmation, reopens the question of where the marble is: *where:::, (0.3) over there.* Tracy's subsequent response is a "complex multimodal Gestalt" (Mondada, 2014, p. 98) constructed out of multiple components including her talk and point in line 20 (Fig. 17). Prior to her producing these, however, she first repositions herself, effectively placing herself between Trent and the feature in the landscape to which she is pointing. Shortly after his question, Tracy begins walking towards Trent in line 17 (Fig. 15). Just as she steps a few feet in front of Trent, she then pivots towards the hill in the distance and points until she stands a few feet directly in front of him. Turning towards the location and relaunches her talk and point from line 10, *The marble's over there*.

Tracy's repositioning in Trent's line of sight plays an essential role in building her course of action. In moving closer and repositioning herself in between him and the marble in the distance, along the same axis that Trent is oriented, Tracy is able to better approximate for herself and for him where in the landscape she is pointing to. This reduces the deictic ambiguity of her point and effectively constitutes a *reformulation* of her prior pointing and talk in line 10. Her mobile action is in a sense preparatory to a better recipient-designed response, but more so its preparatory function is accomplished by her (re)positioning herself for

perception (Goodwin, 1997). What is remarkable about this instance, however, is that the perceptual benefit is not just hers but Trent's as well. In standing in a position that has a better fit or correspondence with Trent's vantage point, Tracy in essence points through both his and her light of sight, and thus more precisely occupies his perspective in doing so. This in turn allows Trent to reciprocally look through her gesture towards the place she is pointing.

Karl Bühler described what he called the *origo,* the "I-here-how" of subjective experience or primitive point of reference from which speakers perceive and describe their surrounding world (Bubandt, 1997). The *origo* has traditionally been treated as solely being a primitive faculty of subjectivity and representing the "…the individualization in discourse of the temporal, spatial and social dimensions of life that posit the individual and make its speech intelligible by locating its fix-point" (Bubandt, 1997, p. 142). As can be seen in the data, in moving into her co-participant's line of sight, the origo for Tracy's utterance and point is interactionally distributed between her and her recipient. Far from being a primitive faculty of subjective experience, we see *origo* and deictic reference emerge through the interactive work of the participants (Mondada, 2005, p. 76). Tracy's mobility is crucial in reducing the perspectival gap between them, not because their prior points or referents were incorrect but because they were not in the position to mutually recognize that those were correct in the first place.

Because particular spaces can be demarcated via movement; we can see participants interpreting their mobile actions in a manner similar to how they do with other actions in talk. The mobile action deployed by Tracy in the preceding example can be seen as analogous to remedial practices we see in talk; she moves for the purpose of re-launching a prior point for locating the "marble". Altogether, we observe the participants deploying and interpreting embodied and mobile actions with a sensitivity to their placement with regard to their "occasioning" in an ongoing course of action.

4. Conclusion

Throughout the extracts, we see participants' mutual perception of features in the landscape indexically being built through the visual practices they use for showing, looking, and seeing with one another. When making accessible some distant feature in the landscape, we observe participants alternatively, gesturally depicting aspects of the feature within their recipient's line of sight (Ex. 1), embodying their own line of sight toward a given feature in repositioning their recipient vis-a-vis that distant feature (Ex. 2), or repositioning their own line of sight within their recipient's for the purpose of pointing towards a distant feature in the landscape (Ex. 3). The participants' coordinated movement, specifically in relation to one participant's embodied perspective or line of sight towards the distant referent in the landscape becomes integral for achieving the mutual accessibility in interaction. In doing so, the participants transform the contextual

configurations between themselves, the ongoing activity, and the landscape in order to make features within that landscape mutually accessible and thus actionable. These practices are interpreted as meaningful, not only due to the placement within the participants' local contextual configuration, but also because of how they facilitate the participants' larger projects. As Goodwin (1997, p. 121) argues, "[t]he larger activity...provides a motivational framework that leads those involved in the activity to make particular perceptual distinctions in the first place (i.e., it establishes a texture of relevancies, a focus for perception)." How the material surround is made relevant and actionable in the interaction; how meaningful places and locales can be delineated within a given space; and how the larger environment perceived as categorically-relevant phenomena, ultimately emerges from within the larger activity in which the participants are engaged.

Studies in multisensoriality have expanded our understanding of multimodality in interaction and have shown how that sensing is more than just private experience and that the methodical practices by which experience is made publicly accessible are mutually elaborated through interaction. Indeed, perception in this line of analysis is not solely private, but rather "…produced for the self and for the other, visually displayed, recognized as such" (Mondada, 2016, p. 359), and as a result is intersubjectively grounded in both the phenomena being perceived and the interactive practices by which that perception is publicly manifested. While the turn toward multisensoriality was initiated largely as a counter to the preoccupation with visual and aural practices in multimodal analyses, many of its insights can be reapplied back to the visual practices that participants use for making co-present phenomena accessible—specifically where that accessibility may not be immediately available, as was the case in the data presented here.

The videographer is faced with many of the same problems as is the co-participants they are studying: they have "…to recognize new emergent action, identify the objects being made relevant…and anticipate the emergence of new interactional spaces" (Mondada, 2019b, p. 99). This can be difficult even in routine settings, but it is especially difficult in the data analyzed here. The participants' courses of action, collaborative work, and motives for moving through or indeed just being in the landscape are often quite technical and difficult to anticipate as an uninitiated observer. When the participants "saw" something relevant for their work, that was often inaccessible for the videographer, even if they are looking at the same thing. More fundamentally, however, much of the interactive work we saw throughout the extracts was being built specifically toward or from one participants' perspective or line of sight. Having a camera as close as possible to that participant's line of sight was the most advantageous way of accessing that perspective. This, however, was not always possible with the videographer holding a handheld camera. Conversely, wearable cameras, though they essentially occupied to the position of the participant, omit other interactive work.

The analysis shows co-present features progressively being revealed via the participants embodied and mobile movement through the landscape. As such features cannot be treated as immediately obvious and ready to be seen; rather, they emerge as the outcome of the perceptual work from participants, who in turn use their bodies, orientation, movement, gesture, and talk so to reveal how they perceive the world in order to make that accessible to their interlocutors. This is largely a product of the both the setting and projects being pursued by the participants: What is visually accessible to others as relevant for their ongoing work cannot be presupposed as available and must instead be achieved through the participants' concerted work. Moreover, participants' embodied perspective or line-of-sight, that is, the axis of gaze and how it projects out toward some distant feature becomes a crucial resource for how the participants would reposition and reorganize themselves vis-a-vis the attended feature. This appeared throughout the extracts to be a largely due to the distance of the features from the participants' immediate vicinity. Without immediate access to the co-present feature, participants instead would have to first demonstrate or determine in which direction one was looking. In this way, visual practices are shown here to be like perception via other modalities: a "multisensorial experience that is methodically organized by the participants engaging in accountable sensorial practices" (Mondada, 2019a, p. 57). As such, similar to other senses (taste, smell, and touch) we gain a glimpse at the interactive work needed in order for what one participant "sees" to be rendered as "seeable" for others. In pushing at the boundaries through which the sensorial is made accessible, and thus available for subsequent use and re-use in interaction, we further explore the margins where sensorial experience is made public, accountable, and intelligible.

Acknowledgements

References

Broth, Mattias & Lundström, Fredrick (2013). A walk on the pier: Establishing relevant places in mobile instruction. In P. Haddington, L. Mondada, & M. Nevile (Eds.), *Interaction and Mobility: Language and the Body in Motion* (pp. 91-122). Berlin/Boston: Walter de Gruyter, GmbH.

Bubandt, Nils Ole (1997). Speaking of Places: Spatial Poesis and Localized Identity. In J. J. Fox (Ed.), *The Poetic Power of Place. Comparative Perspectives on Austronesian Ideas of Locality* (pp. 132-162). Canberra: Australian National University Press.

Goodwin, Charles (1997). The Blackness of Black: Color Categories as Situated Practice. In L. B. Resnick, R. Säljö, C. Pontecorvo, & B. Burge (Eds.), *Discourse, Tools and Reasoning: Essays on Situated Cognition* (pp. 111-140). Berlin, Heidelberg: Springer Berlin Heidelberg.

Goodwin, Charles (2000a). Practices of Seeing: Visual Analysis: An Ethnomethodological Approach. In T. van Leeuwen & C. Jewitt (Eds.), *Handbook of Visual Analysis* (pp. 157-182). London: Sage Publications.

Goodwin, Charles (2000b). Action and embodiment within situated human interaction. *Journal of Pragmatics, 32*(10), 1489-1522.

Goodwin, Charles (2007). Environmentally Coupled Gestures. In S. D. Duncan, J. Cassell & E. T. Levy (Eds.), *Gesture and the Dynamic Dimension of Language* (pp. 195-212). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Goodwin, Charles (2010). Things and their Embodied Environments. In L. Malafouris & C. Renfrew (Eds.), *The Cognitive Life of Things* (pp. 103-120). Cambridge, UK: McDonald Institute Monographs

Goodwin, Charles & Smith, Michael Sean (2020). Calibrating professional perception through touch in geological fieldwork. In A. Cekaite & L. Mondada (Eds.), *Touch in Social Interaction: Touch, Language and Body* (pp. 269-287). London: Routledge

Gibson, James (1979). *The Ecological Approach to Visual Perception.* Boston: Houghton Mifflin.

Haddington, Pentti; Mondada, Lorenza & Nevile, Maurice (2013). Being Mobile: Interaction on the Move. In P. Haddington, L. Mondada & M. Nevile (Eds.), *Interaction and Mobility: Language and the Body in Motion* (pp. 3-64). Berlin/Boston: Walter de Gruyter, GmbH.

Ingold, Tim (2004). Culture on the Ground: The World Perceived Through the Feet. *Journal of Material Culture, 9*(3), 315-340.

Jefferson, Gail. (2004). Glossary of transcript symbols with an Introduction. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the first generation* (pp. 13-23). Philadelphia: John Benjamins.

Kendon, Adam (1990). *Conducting interaction: Patterns of behavior in focused encounters.* Cambridge: Cambridge University Press.

Kreplak, Yaël & Mondémé, Chloe (2014). Artworks as touchable objects: Guiding perception in a museum tour for blind people. In M. Nevile, P. Haddington, T. Keisanen & M. Rauniomaa (Eds.), *Interacting with objects: language, materiality, and social activity* (pp. 295-319). Amsterdam, The Netherlands: John Benjamins Publishing Company.

Mondada, Lorenza (2005). La constitution de l'origo déictique comme travail interactionnel des participants: une approche praxéologique de la spatialité. *Intellectica. Revue de l'Association pour la Recherche Cognitive*, *41*(2), 75-100.

Mondada, Lorenza (2009). Video recording practices and the reflexive constitution of the interactional order: some systematic uses of the split-screen technique. *Human Studies, 32*(1): 67-99.

Mondada, Lorenza (2012). Garden lessons: Embodied action and joint attention in extended sequences. In H. Nasu & F. Chaput Waksler (Eds.), *Interaction and Everyday life: Phenomenological and ethnomethodological essays in honor of George Psathas* (pp. 293-311). Lanham: Lexington Books.

Mondada, Lorenza (2014). Bodies in action: Multimodal analysis of walking and talking. *Language and Dialogue, 4*(3), 357-403.

Mondada, Lorenza (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics, 20*(3), 336-366.

Mondada, Lorenza (2018a). The multimodal interactional organization of tasting: Practices of tasting cheese in gourmet shops, *Discourse Studies, 20*(6), 743-769

Mondada, Lorenza (2018b). Visual practices: video studies, multimodality and multisensoriality. In D. Favareau (Ed.) *Co-Operative Engagements in Intertwined Semiosis: Essays in Honour of Charles Goodwin* (pp. 304-325). Tartu Semiotics Library, 19. Tartu: Tartu University Press

Mondada, Lorenza (2019a), Rethinking bodies and objects in social interaction: a multimodal and multisensorial approach to tasting. In U. T. Kissmann & J van Loon (Eds.) *Discussing New Materialism: Methodological Implications for the Study of Materialities*, (pp. 109-134). Wiesbaden: Springer

Mondada, Lorenza (2019b). Practices for Showing, Looking, and Videorecording: The Interactional Establishment of a Common Focus of Attention. In E. Reber & C. Gerhardt (Eds.), *Embodied Activities in Face-*

to-face and Mediated Settings: Social Encounters in Time and Space (pp. 63-104). Springer International Publishing.

Mondada, Lorenza (2020). Audible Sniffs: Smelling-in-Interaction. *Research on Language & Social Interaction*, *53*(1), 140-163.