



Understanding as an embodied, situated and sequential achievement in interaction

Lorenza Mondada^{a,b,*}

^a FRIAS (Freiburg Institute for Advanced Studies), Germany

^b ICAR Research Lab (CNRS & University of Lyon), France

ARTICLE INFO

Article history:

Received 8 June 2010

Accepted 26 August 2010

Keywords:

Understanding

Ethnomethodology

Social interaction

Conversation analysis

Accountability

Sequentiality

Time

Multimodality

Embodiment

ABSTRACT

This paper aims at outlining the sequential, situated and embodied dimensions of understanding in interaction. This perspective on understanding, originating in Ryle and Wittgenstein, further developed by ethnomethodology and conversation analysis, focuses on the orderly unfolding of sequences of actions in time. It also focuses on understanding as a collective achievement, publicly displayed and interactively oriented to within the production and the monitoring of action. Its accountability is built through a plurality of displays, claiming and demonstrating understanding, thanks to the mobilization of linguistic and embodied resources at specific sequential positions. The paper discusses evidences of understanding as they are publicly displayed by the participants in interaction: It both offers a discussion of the ethnomethodological and conversation analytic literature, and an analysis an empirical case, focusing on the interplay of embodied and sequential features in the production and monitoring of understanding.

© 2010 Elsevier B.V. All rights reserved.

1. A praxeological, sequential and temporal approach to understanding

Understanding is a fundamental aspect of everyday life and a key issue for social sciences and cognitive sciences alike, defining the basic conditions for mutual communication, joint action and social co-existence. In a long tradition of studies within both fields, understanding has mainly been considered as a cognitive, private, individual phenomenon. It has often been treated in terms of propositional contents more than in terms of actual actions. Within more recent praxeological perspectives – principally inspired by Ryle and Wittgenstein – such as ethnomethodology, conversation analysis, and other models of situated action, understanding has been treated as an indexical, collective, publicly accountable achievement in time. As Garfinkel puts it, “the appropriate image of a common understanding is therefore an operation rather than a common intersection of overlapping sets” (1967:30).

This special issue pursues this empirical, praxeological, interactional approach to understanding by stressing its situated, contingent, embodied, and intersubjective dimensions. In this sense, it contributes to a rich set of studies which question, reformulate and respecify classical cognitive topics through the study of social practices within ethnomethodology (Coulter, 1991; Lynch, 2006), discursive psychology (Edwards, 1996; Potter, 1998), the study of situated and distributed cognition (Lave, 1988; Hutchins, 1995) and conversation analysis (Schegloff and Sacks, 1973; Moerman and Sacks, 1988; Schegloff, 1991, 1992) (see te Molder and Potter, 2005; Van Dijk, 2006 for edited volumes presenting a diversity of approaches).

* Correspondence address: ICAR Research Lab, ENS Lyon, 15, parvis René Descartes, BP 7000, F-69342 Lyon Cedex 07.

E-mail address: Lorenza.Mondada@univ-lyon2.fr.

This paper offers a conclusive comment for this special issue by focusing on the situated, sequential, and embodied organization of understanding in interaction. It draws both on the papers of the special issue, that is, the ethnomethodological and conversation analytic literature, and a specific empirical case, pursuing the central focus of this special issue: the interplay of embodied and sequential features in the production and monitoring of understanding.

Both ethnomethodology and conversation analysis have stressed the temporal and sequential nature of understanding within action and interaction. As shown by Koschmann in the introduction to this issue, the notion that understanding is displayed in the orderliness of the temporal unfolding of practical actions is already present in Garfinkel's early writings: "Understanding refers to the modes of treatment that A affords B that operate, as far as A sees it, under constant confirmation of A's anticipations of treatment from B" (1952:367). This conception of understanding as related to what a second participant *does* in response to a first participant resonates with Wittgenstein's grammar of understanding: "Try not to think of understanding as a 'mental process' at all. – For that is the expression which confuses you. But ask yourself: in what sort of case, in what kind of circumstances, do we say, 'Now I know how to go on' when, that is, the formula has occurred to me?": 1953:§154). In both cases, understanding is not treated as a mental process but is related to the next action achieved by the co-participant and demonstrating her understanding.

This "nextness" is further specified by conversation analysis in terms of sequentiality. Understanding is naturally displayed by the fact that, and the manner in which, participants "go on" within the conversation: "Understanding matters as a natural phenomenon in that conversational sequencing is built in such a way as to require that participants must continually, there and then – without recourse to follow up tests, mutual examination or memoirs, surprise quizzes and other ways of checking out on understanding – demonstrate to one other that they understood or failed to understand the talk that they are party to" (Moerman and Sacks, 1988:85).

This sequential notion of understanding (cf. Schegloff, 1992:1300 for a development and an example) is a fundamental aspect of what is achieved within an adjacency pair, in Schegloff and Sacks's words: "by an adjacently positioned second, a speaker can show that he understood what a prior aimed at, and that he is going along with that. Also, by virtue of the occurrence of an adjacently produced second, the doer of a first can see that what he intended was indeed understood, and that it was or was not accepted. Also of course, a second can assert his failure to understand, or disagreement, and inspection of a second by a first can allow the first speaker to see that while the second thought he understood, indeed he misunderstood. It is then through the use of adjacent positioning that appreciations, failures, corrections etcetera can be themselves understandably attempted" (1973:297–298). This display constitutes a "proof procedure" (Sacks et al., 1974:728–729) since it is both a resource for the participants building the mutual accountability of their conducts and an object of analysis for the observer (see Lynch, this issue, for a discussion of this point).

2. Claiming vs. demonstrating understanding

In this final comment, I am interested in the evidences of understanding publicly displayed by the participants in interaction. As shown by the literature derived from the early praxeological insights into understanding, the orderliness and sequentiality of understanding as a practice embedded in the very continuation of the ongoing action opens up the possibility of studying how understanding is achieved and displayed within actual courses of action. A variety of manifestations of understanding have been pointed up both in this special issue and in the literature, which I discussed before to analyze a fragment of data.

In his *Lectures*, Sacks distinguishes between *claiming* and *demonstrating* understanding, on the basis of the following (invented) example:

(Sacks, 1992:II:141)

- | | | |
|----|----|-----------------------------|
| 1 | A: | where are you staying |
| 2 | B: | Pacific Palisades |
| 3a | A: | oh at the west side of town |
| | | vs |
| 3b | A: | oh Pacific Palisades |

Whereas in 3a, by re-describing the location given by B, A displays that he recognizes the place referred to, in 3b, by merely repeating it, A does not. In the former case, he *demonstrates* understanding, in the latter case he just *claims* it. In this sense, the repetition is equivocal, whereas the reformulation of the location is not – even when it provides for the evidence of a misunderstanding (such as in 3c: "oh in the center of town", which can then be repaired in the next turn). Thus, Sacks answers the question of "how understanding is shown" (1992:II:140), by pointing to the fact that participants make available different forms of understanding by performing some kind of operation on the previous turn. In these cases, participants "do showing understanding," as they can do "questioning" or "answering" in conversation (1992:I:141). Moreover, Sacks shows that there are specific sequential environments within talk for "doing understanding" – what he calls "understanding positions" (1992:II:426). The completion of a story is an example of such a position, where recipients can exhibit their possible understanding, for example by using a proverbial expression.

The task of the analyst is to identify what for the participants counts as claimed vs. demonstrated understanding, as sufficient or insufficient proof of understanding (see Hindmarsh et al., 2011). Overt demonstrations of understanding are not the most usual case (Fele, 1992:428; Heritage, 2007:275) and understanding is generally achieved in a tacit way. Recognition is the default, understanding mostly working as a taken-for-granted matter, embedded in the next action, but not in a specialized practice for displaying it: “The understandings are displayed en passant for the most part [...], as by-products of bits of talk designed in the first instance to do some action such as agreeing, answering, assessing, responding, requesting, and so on” (Schegloff, 1992:1300).

This opens up the question of the variety of manifestations of understanding and their situated character – as both contextually and sequentially tied to the ongoing action and situationally interpreted by the co-participants.

Formulations of understanding are evidently the most explicit way of claiming understanding. They can, however, be used for doing other jobs than showing understanding: stating that one understands (or not) can be in service of introducing a complaint (Depperman and Schmitt, 2008), for asking or offering help (cf. Lindwall and Lymer, 2011), for closing or delaying closure of a task, etc. Interestingly, these actions are especially done with negative formulations (“I don’t understand”), which can display resistance to some ongoing activity. Similar practices have been described for “I don’t remember” (Lynch and Bogen, 2005), or “I don’t know” (Metzger and Beach, 1996). Moreover, claims of understanding can be treated differently by recipients orienting to their situated production: in Macbeth’s last example (2011), the mere fact that Eddie says “I understand really well” is not enough to characterize what he does: the timed position of when he does it (in overlap with teacher’s turn pre-completion) displays his understanding as much as what he says.

Moreover, as clearly described by Coulter (1979), using Ryle’s (1949:149ff) distinctions, the verb “to understand” is not a process-verb like “to play” but an achievement-verb like “to win.” Therefore, its occurrence refers to the result – the success – of a process but *not* to the process itself: it does not describe a “temporally-extended course of action” (Coulter, 1979:37). As Ryle notes, “there can be achievements which are prefaced by no task performances. We sometimes find things without searching, secure appointments without applying and arrive at true conclusions without having weighed the evidence” (Ryle, 1949:150). Thus, formulations claiming to understand or to have understood do not per se imply that a process of understanding has taken place.

Even if explicit formulations can be used, most often talk proceeds with tacit claims of recognition, embedded in the prospective and progressive continuation of the ongoing action. Thus, understanding concerns crucially two fundamental principles of conversation: the preference for progressivity and the preference for intersubjectivity (Schegloff, 1992, 2007:15; Heritage, 2007). Progressivity is the product of the relevance of the basic feature of sequentiality: “nextness” or adjacency. Its suspension occasions participants’ inspection of the “understanding-so-far,” identification of a possible difficulty and clearing up problems of hearing or understanding before going on with the next action.

These principles are closely related to the two preferences organizing reference to persons (Sacks and Schegloff, 1979): as Heritage (2007:260) shows, the preference for recognitional reference is a special case of the preference for securing intersubjectivity and the preference for minimized reference a special case of the preference for progressivity. These preferences are concurrently relevant: if proper names satisfy both of them, absence of recognition and displays of misunderstanding can generate solicitations of signs of understanding (in try-marked reference for example), maximizations of the reference and/or repair practices, and thus a prevalence in securing intersubjectivity over progressivity.

Exploring the way in which multi-unit turns are methodically and interactionally achieved, moment by moment, by the participants, Schegloff (1982) shows that at transition-relevance points, recipients can produce continuers (such as tokens of interest, acknowledgment tokens, “uh huh’s,” nods) or alternatively initiate repairs. The completion of a TCU is an “understanding position” (Sacks, 1992:II:426), where it is structurally relevant to display understanding: “uh huh” exhibits the recipient’s understanding of the unit so far, passing an opportunity to initiate a turn and, more specifically, to initiate repair. In this sense, the very notion of “continuer” directly relates to the organization of progressivity (“the ‘in-progress’ character of the talk”: 1982:82).

As Schegloff says, “uh huh’s”, etc. as continuers do not merely claim an understanding without displaying anything of the understanding they claim. The production of talk in a possible turn position which is nothing other than ‘uh huh’ claims not only ‘I understand the state of the talk’ but embodies the understanding of that extended talk by another is going on by declining to produce a fuller turn in that position” (1982:81). Other displays of understanding than “uh huh” can be mobilized and can be more or less embodied or emotively implemented, such as “yeah” or “yes” (Jefferson, 1983); change-of-state tokens like “oh” (Heritage, 1984), displays of surprise (Wilkinson and Kitzinger, 2006), nods (Goodwin, 1980; Stivers, 2008), assessments (Goodwin and Goodwin, 1987), etc.

Continuers decline to perform a specific alternative class of actions, next-turn repair initiation (Schegloff, 1992). Repair devices show that talk provides for systematic opportunities to re-establish mutual understanding at every moment: although the preference for self-correction (Schegloff et al., 1977) invites withdrawal from other-repair initiation until the current turn containing the repairable is completed, the transition-relevance point after the repairable is a potential position where it could be initiated. In particular, other-initiated repairs are the mechanism through which intersubjectivity can be reprimed after any possible trouble: “when such ‘problematic understandings’ occur, and whatever their apparent ‘source’, speakers of the ‘misunderstood’ talk can undertake to ‘repair’ the misunderstanding, and this can thus constitute ‘third position repair’ – repair after an interlocutor’s response (second position) has revealed trouble in understanding an earlier turn (the ‘repairable’ in first position)” (Schegloff, 1992:1301). Third position repairs reveal that speakers constantly

pay attention to possible troubles in understanding, at all the sequential positions, within the turn as well as within the sequence.

Thus, understanding is constantly actively managed by the participants along with the emergent, incremental, sequential organization of turns moment by moment; in the way in which they respond, they allow the speaker to go on with a continuer, or repair the previous bit of talk. Moreover, understanding is constantly displayed in a multimodal way: participants manifest their current understanding in their gesture, gaze, facial expression, body position, etc. This management of understanding through visual embodied resources, as well as through linguistic ones, provides for a “transparency of understanding” (LeBaron and Koschmann, 2003) as it is publicly available, mutually intelligible, that is, accountable (see also Schmitt, 2010).

Whereas gesture and other multimodal conducts are often observed by focusing on speakers – the speaker being the one who gesticulates during his talk – multimodal conducts of recipients have been less studied. As magistrally shown by Goodwin (1981), the gaze of the recipient does not only display her attention and monitoring of the ongoing utterance, but also reflexively shapes the emergent format of the speaker’s turn, who adjusts his emergent construction to the temporality of gaze and head movements, thanks to delaying devices like re-starts, expansions and insertions. As demonstrated by Hindmarsh and colleagues (2011), the way in which participants dispose their bodies, lean over a common object of attention, construct mutual focus, bodily align with the recipient-design action, displays their understanding of the ongoing activity. Interestingly, multimodal displays and even negotiations of understanding are not limited to linguistic and visual features, but include also touch, as extensively described by Nishizaka (2011).

3. Understanding in institutional and instructional settings

In institutional settings, such as those to which this special issue is devoted (see also Depperman et al., 2010), ordinary as well as professional understandings are at stake. Institutional and professional activities can imply particular modes of understanding, related to specific tasks and practices, implying specific forms of listening, interpreting, inferring, diagnosing, etc.), but also specific forms of perception, both visual (Hindmarsh et al., 2011; Koschmann, 2011) and tactile (Nishizaka, 2011), which constitute a “professional understanding” similar to the “professional vision” described by Goodwin (1994) for archeologists and police officers.

For example, Peräkylä (2005, 2008) describes the sequential organization of “interpretations” in psychoanalysis, by identifying a three-turn sequence in which the psychoanalyst produces an interpretation of the patient’s talk, the patient displays understanding of and agreement with it and, within the third turn, the psychoanalyst elaborates and reorients the patient’s reception and/or continues her interpretation. Understanding is central in all of these positions, and concerns the professional understanding of the patient’s case, the patient’s understanding of the psychoanalyst’s interpretation and, in third position, the psychoanalyst’s understanding of the patient’s understanding of her interpretation.

Studying health appraisal interviews, Beach and Dixon (2001) describe the initiation by the doctor of a similar sequential pattern, a three-part formulation circle constituted by the interviewer’s formulated understandings, the patient’s confirmation, and a topic shift achieved by the interviewer. Similarly to the psychoanalyst’s interpretation, which is an interactional implementation of the professional appraisal of the patient’s case, the doctor’s formulation (often prefaced by “from what I’m understanding”) displays a focused understanding of selected and often delicate details of patient’s biopsychosocial history, in service of an empathic form of medical interview.

More generally, medical encounters, court hearings, counseling sessions, classroom interactions, and many other institutional settings are characterized by claims and displays of understanding (cf. Depperman, 2010) that are specifically related to specific membership categories, to epistemic asymmetries, and to specific rights and obligations (Stivers et al., in press). For example, Lindwall and Lymer (2011) show the relevance of different rights and obligations among students and teachers, and their interactional implications, visible in a specific distribution of verbal and pronominal forms referring to understanding.

More specifically, as shown by the majority of the papers in this special issue, in instructional and educational settings understanding is a key issue, both in participants’ terms and in the terms of the disciplines theorizing them – which often regard understanding in terms of outcome, of normative evaluation, of “correct understanding” (Macbeth, this issue, 2000:64; Lindwall and Lymer, 2011). Instruction in educational settings can be seen as an activity devoted to the production of new understandings, and to the checking of these understandings, whereby teachers orient to the fact that students can *do* the task but this is not yet a demonstration that they have *understood* it (Lindwall and Lymer, 2011). These vernacular distinctions show how understanding is deeply related to the normative, disciplinary, institutional features of the setting. These orientations are accountably constructed within the situated, interactive, temporal and sequential achievement of understanding: as Macbeth puts it, “sequential organization is itself the on-going achievement of common understanding” (this issue). This is typically displayed in the IRE sequence (cf. Mehan, 1979; McHoul, 1990; Macbeth, 2004), where the initial question can be revised in order to adjust to the answers and displays of the students, and to afford recipients’ new opportunities to produce revised responses orienting to the revision (Zemel and Koschmann, 2011).

Instructed action in a variety of settings – well documented across medical (Hindmarsh et al., 2011; Koschmann, 2011; Nishizaka, 2011) and educational (Macbeth, 2011; Lindwall and Lymer, 2011; Zemel and Koschmann, 2011) contexts in this issue – offers a perspicuous setting for the documentation of understanding as it is interactively built across sequences. Instructing action and instructed action constitute a paired action: as shown by Garfinkel, rather than unilaterally

determining instructed action, instructions acquire a sense through the situated task of following them. More particularly, instructions are always incomplete, and “following” them supposes a competent interpretation – an understanding of the rule – which retrospectively and reflexively configures their meaning (Garfinkel, 2002:ch. 6; Garfinkel, 1967:29; Amerine and Bilmes, 1988). Instructions can be more or less explicit, but are always indexical, acquiring their sense in the situated and interactive achievement of the instructed action (cf. Mondada, *in press-a,b*). Similarly, the understanding of the surgical procedure analyzed by Koschmann (2011) is not given a priori, but emerges as the procedure unfolds, relying on the expectations of the participants and on their prospective/retrospective definition and recognition of a progressively discovered object.

The language game of instruction is precisely the context in which Wittgenstein discusses the idea of “following a rule” and makes explicit his praxeological conception of understanding: “Let us now examine the following kind of language-game: when A gives an order B has to write down [a] series of signs according to a certain formation rule [...] At first perhaps we guide his hand in writing out the series 0 to 9; but then the possibility of getting him to understand will depend on his going on to write it down independently” (1953:§143).

4. Embodied understanding and monitoring of understanding: an empirical case

Audio and video recordings document the way in which the instructor guides the instructed and in which the latter finally grasps the instruction and understands it. As shown by Hindmarsh and colleagues’ fine-grained multimodal analysis of trainees following dentists’ instructions and displaying their attention, interest and understanding in their bodily postures (this issue) or Koschmann et al.’ analysis (this issue) of the way in which the resident’s understanding of the course of the surgery is embedded in the prospective production of a situatedly appropriate action within the procedure, it is difficult to investigate situated understandings without taking into account the detail of the embodied conducts of the participants and their mutual monitoring. In order to give a further example of the way in which understanding is progressively enabled, displayed and monitored by participants in the linguistic and embodied details of their emergent action, and to show how these details can be exploited as evidences of the situated understanding and understanding of understanding, I will briefly focus on an empirical case of instructions.

When a customer has bought a new car from a garage showroom and comes to pick it up, the salesman delivers it with a set of instructions, given to the customer just before she drives the car home. The excerpt analyzed below is taken from a series of video-recordings of the car deliveries made in a garage on the periphery of a large town in France. In this excerpt, the dealer, Jan, is explaining to the customer, Marie, the technological functionalities of the dashboard. This activity involves both instructing actions followed by instructed actions, as well as instructor’s monitoring the instructor’s understanding of the instructions.

The excerpt was video-recorded in the garage, with a camera placed on the dashboard, in front of the participants who were sitting side by side (Fig. 1a and b).

We join the action as Jan engages in the explanation of a series of controls activated by repeatedly pushing a button (circled in Fig. 1b) on the steering wheel: at each pressure of the button, a small screen situated in front of them, just below the camera (circled in Fig. 1a), displays specific information about the remaining gasoline, the kilometers driven, the mean consumption, and the consumption at that moment. In the following excerpt, the first item is explained. In order to capture various embodied evidences of participants’ action and action understanding, the transcript documents various details in concomitant parallel lines: the original talk (in bold), its translation, and a series of embodied conducts, such as Jan’s hand gesture (jaH) and gaze (jaG), Marie’s hand gesture (maH) and gaze (maG), synchronized with the ongoing talk with various reference points (*, ●, +, †) (see the conventions explained at the end of this paper).

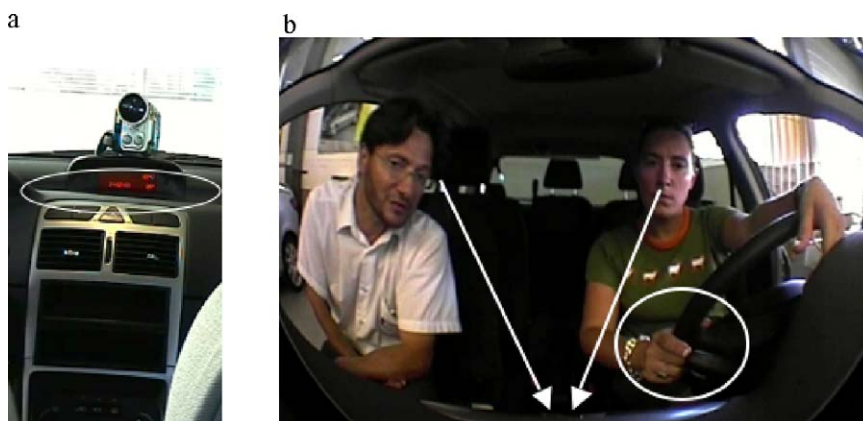


Fig. 1. Position of the camera and side-by-side arrangement of the participants looking at the screen in front of them.

Excerpt 1 (EMIC1607L_7.04)

1 Jan: **.h ici.* +(.) les +positions# ici *c'est,* (0.4)* la +consom#mation.
 .h here. (.) the positions here are, (0.4) the consumption.
 jaH *.....*points to and shows button-*pushes*.....*
 jaG *looks at the button*looks at the dashboard-->
 maG +.....+looks at the button-----+looks dsb-->>
 fig #fig.2 fig.3#



Figure 2



Figure 3

2 c't-à-dire qu'en gros* on va +savoir qu'est-ce qu'il vous# reste
 that means roughly that we'll know how many remaining kilometers
 jaG *.....*turns to and looks at Mar-->
 maH †.....†grasps wheel-->
 fig fig.4#



Figure 4

3 comme kilomètres à faire pour ne pas tomber
 you can do in order not to run

4 en panne d'es+sen#ce. •
 out of gas.

jaG -->*
 maG +raises eyebrows-->
 fig #fig.5



Figure 5



Detail of figure 5

Jan's instructing turn (1–4) is produced while he shows the relevant position on the steering wheel to be looked at and Marie not only looks at it but grasps the wheel and activates the explained functionality. Thus, the sequence involves not only explaining a tool but actually inviting the other to use it.

The turn is organized as a description, where the initial part (“les positions ici c’est”, 1) projects a list of functionalities of the button, followed by an NP naming the first position (“la consommation”, 1), then expanded by an explanation introduced by “c’t-à-dire” (2). The instruction is achieved by pointing at the button and by pushing it (1).

Marie’s understanding of the instruction is finally expressed in line 5 by a change-of-state token (“AH,” cf. Heritage, 1984) and an assessment (cf. Mondada, 2009), but well before this, it is embodied in various actions: line 1, she looks first at the button indicated by the deictic reference “ici” and by Jan’s pointing, and immediately after the pause she looks at the dashboard, where the result of Jan’s action, verbalized by him (“la consommation”, 1), is displayed on the little screen in front of them. We can see that Jan looks at the dashboard in an anticipated way (when his pointing has reached its maximal extension and before he pushes the button), projecting the future display, whereas Marie looks at it later: her gazing at the dashboard embodies a first understanding of the relation between the pushed button and the screen display.

During the subsequent expansion of the explanation, immediately initiated by Jan in line 2, Marie grasps the wheel, further embodying her understanding of the instruction, which is not merely aimed at her abstract comprehension of the system but at her active and tactile understanding of how it works.

At this point, Jan has already turned his head towards her (just after the beginning of his expansion, 2) and looks at her: he continues to sustain this posture and gaze until his turn pre-completion (4), that is, until the very moment at which Marie’s facial expression displays her final understanding, manifested by her raised eyebrows (4). As the eyebrows are maximally raised, she utters the change-of-state “AH.” As soon as this facial display emerges, projecting the final demonstration of understanding, Jan looks at the dashboard again and, in overlap with her assessment, initiates the explanation of the second item.

Jan’s turn completion reveals a perfect coordination between various activities constituting the instructing action, the instructed action and the understanding of the instruction: both participants orient towards turn completion, projectable thanks to its syntactic and pragmatic emergent features; Jan’s turn and Marie’s actions reflexively achieve the intelligibility of the tool being focused on (the turn accounts for what Marie is seeing on the display and she sees it, having literally grasped how the button works).

Marie’s change-of-state token in line 5 constitutes the final expression of her change of state, which has been achieved within the previous moments, within an embodied process deployed within the time of the utterance, the gesture and the gaze as they are interactively and sequentially organized. Her process of discovery resonates with Wittgenstein’s description: “A writes [a] series of numbers down; B watches him and tries to find a law for the sequence of numbers. If he succeeds he exclaims: ‘Now I can go on!’—So this capacity, this understanding, is something that makes its appearance in a moment” (Wittgenstein, 1953:§151).

The way in which the instruction goes on after this first item reveals how systematic this embodied arrangement of multimodal resources and actions is:

Excerpt 2 (continuation of excerpt 1)

```

5 Mar: AH d'accord. c'est su[per ça,
      OH okay.      that's gr[eat that,
6 Jan: [après la deuxième pression, allez-y,
      [then the second push, go on,
maG ---> continues to look at the screen in front of her-->
7 (+1.1)
maH †pushes†
8 Jan: c'est l'nombre de kilomètres qu'vous allez faire à *partir
      it's the number of kilometers that you can drive from
jaG *looks at Mar-->
9 de <maintenant>.
  <now on>.
9 (1.2)
10 Jan: +donc c'est un deuxième compteur+ journa•lier.*
      thus it's a second daily counter
jaG ----->*,,,,,*
maG +slow nod-----+
11 (0.8)

```

In line 6, Jan mentions the second item of the list projected at the beginning of the fragment (1). He then adds a directive, formulated with an imperative (“allez-y”). This directive is the first uttered within this episode: it exploits the format installed by the previous sequence, as well as the bodily and gestural arrangement created by it, since Marie is still grasping the wheel. Whereas in the first fragment Jan activated the button, in this second fragment he orders her to activate it. The directive is granted by Marie’s next action (7). Once her action has occasioned the display of a new piece of information on

Finally, the last item is introduced:

Excerpt 5 (continuation of excerpt 4)

```

25 Jan:  et ensuite, la dernière h,
          and then, the last one h,
          maH                               †pushes-->
26      (0.3)†
          maH                               -->†
27 Jan:  c'est la moyenne kilomètre heure,
          it's the average kilometre per hour
28      (0.5)
29 Jan:  *qu'vous allez faire entre un point,* et un* autre point.*
          that you will do between one point, and another point.
          jaG                               *r r r r*
          maG                               †multiple nods-----†

```

Jan mentions the last item (25), Marie pushes the button (26), and he briefly describes what the number appearing on the screen refers to (27). A pause occurs in line 28, where she does not respond, and he produces an increment (Ford et al., 2002), in the form of a relative phrase (29). During the increment, he looks at her and is responded to by her multiple nods.

The iterative format of this instructional episode shows how instructions develop and stabilize in a systematic format, where instructing actions reflexively adjust to the emergent understanding visible in the instructed actions. Linguistically formatted responses at the end of the sequence represent only one kind of display: the fact that the instructor systematically gazes at the instructed as soon as he explains the output of her action shows that he actively monitors her embodied responses and does not only rely on her verbal ones (cf. Hindmarsh et al., 2011). Responses reflexively achieve the closing of the sequence (cf. Lindwall and Lymer, 2011), whereas absence of responses are oriented to by expanding or incrementing it, adjusting it to the recipient's displays of understanding. The temporality of turn and sequence formats is adjusted to the emergent understanding, being accelerated or slowed down as responsive conducts are produced or delayed.

5. Conclusions

In this paper, I have tried to outline the sequential, situated and embodied dimensions of understanding in interaction. This perspective on understanding, originating in Ryle and Wittgenstein, further developed by ethnomethodology and conversation analysis, focuses on the orderly unfolding of sequences of actions in time. It also focuses on understanding as a collective achievement, publicly displayed and interactively oriented to within the production and the monitoring of action. Its accountability is built through a plurality of displays, claiming and demonstrating understanding thanks to the mobilization of linguistic and embodied resources at specific sequential positions. In this context, the papers of this special issue provide for extensive demonstrations of the way in which the multimodal dimension of understanding can be empirically and analytically considered. This special issue also shows the importance of institutional contexts for the study of the situated nature of understanding and their embeddedness within social, organizational and institutional constraints as well as within specific material and spatial environments.

Transcript conventions

Talk

Data were transcribed according to conventions developed by Gail Jefferson and commonly used in Conversation Analysis.

[overlapping talk
=	latching
(.)	micro pause
(0.6)	timed pause
:	extension of the sound or the syllable it follows
.	stopping fall in tone
,	continuing intonation
?	rising inflection
mine	emphasis
°uh°	quieter fragment than its surrounding talk
.h	aspiration

h	out breath
((sniff))	described phenomena
< >	delimitation of described phenomena
()	string of talk for which no audio could be achieved

An indicative translation is provided line per line, in order to help reading the original.

Embodied conducts:

An indicative translation is provided line per line.

Multimodal details have been transcribed according to the following conventions (Mondada, 2007):

* *	delimit descriptions of the dealer/Jan's actions (jaH).
• •	delimit description of dealer/Jan's gaze (jaG).
† †	delimit descriptions of the customer/Marie's actions (maH).
+ +	delimit descriptions of the customer/Marie's gaze (gaH).
* - - - >	action described continues across subsequent lines.
* - - - >>	action described continues until and after excerpt's end.
- - - - >*	action described continues until the same symbol is reached.
>> - -	action described begins before the excerpt's beginning.
....	action's preparation.
.....	action's retraction.
jan	participant doing the action is identified in small characters when he is not the current speaker or when the gesture is done during a pause
fig	figure; screen shot
#	indicates the exact moment at which the screen shot has been recorded

References

- Amerine, R., Bilmes, J., 1988. Following instructions. *Human Studies* 11, 327–339.
- Beach, W.A., Dixon, C., 2001. Revealing moments: formulating understandings of adverse experiences in a health appraisal interview. *Social Science & Medicine* 52, 25–45.
- Coulter, J., 1979. *The Social Construction of Mind*. Macmillan, London.
- Coulter, J., 1991. Cognition in an ethnomethodological mode. In: Button, G. (Ed.), *Ethnomethodology and the Human Sciences*. Cambridge University Press.
- Depperman, A., 2010. Zur Einführung. In: Depperman, A., et al. (Eds.), *Verstehen in professionellen Handlungsfeldern*. Narr, Tübingen, pp. 7–26.
- Depperman, A., Schmitt, R., 2008. *Verstehendokumentation: Zur Phänomenologie von Verstehen in der Interaktion*. Deutsche Sprache 36, 220–245.
- Depperman, A., Reitemeier, U., Schmitt, R., Spranz-Fogasy, T., 2010. *Verstehen in professionellen Handlungsfeldern*. Narr, Tübingen.
- Edwards, D., 1996. *Discourse and Cognition*. Sage, London.
- Fele, G., 1992. *La comprensione nell'interazione*. *Rassegna Italiana di Sociologia* 33 (3), 425–438.
- Ford, C.E., Fox, B.A., Thompson, S.A., 2002. Constituency and the grammar of turn increments. In: Ford, C.E., Fox, B.A., Thompson, S.A. (Eds.), *The Language of Turn and Sequence*. Oxford University Press, Oxford, pp. 14–38.
- Garfinkel, H., 1952. *The perception of the other: a study in social order*. PhD dissertation, Harvard University.
- Garfinkel, H., 1967. *Studies in Ethnomethodology*. Prentice-Hall, Englewood Cliffs, N.J..
- Garfinkel, H., 2002. *Ethnomethodology's Program: Working out Durkheim's Aphorism*. Rowman and Littlefield, Lanham.
- Goodwin, C., 1981. *Conversational Organization: Interaction Between Speakers and Hearers*. Academic Press, New York.
- Goodwin, C., 1994. Professional vision. *American Anthropologist* 96 (3), 606–633.
- Goodwin, M.H., 1980. Processes of mutual monitoring implicated for the production of description sequences. *Sociological Inquiry* 50 (3–4), 303–317.
- Goodwin, C., Goodwin, M.H., 1987. Concurrent operations on talk: notes on the interactive organization of assessments. *Pragmatics* 1 (1), 1–55.
- Heritage, J., 1984. A change-of-state token and aspects of its sequential placement. In: Atkinson, J.M., Heritage, J. (Eds.), *Structures of Social Action*. Cambridge University Press, Cambridge, pp. 299–345.
- Heritage, J., 2007. Intersubjectivity and progressivity in person (and place) reference. In: Enfield, N.J., Levinson, S. (Eds.), *Person Reference in Interaction: Linguistic, Cultural, and Social Perspectives*. Cambridge University Press, Cambridge, pp. 255–280.
- Hindmarsh, J., Reynolds, P., Dunne, S., 2011. Exhibiting understanding: The body in apprentices. *Pragmatics* 43, 489–503.
- Hutchins, E., 1995. *Cognition in the Wild*. MIT Press, Cambridge.
- Jefferson, G., 1983. Notes on a systematic deployment of the acknowledgement tokens 'yeah' and 'mm hm'. *Tilburg Papers in Language and Literature* 30, 1–18.
- Koschmann, T., 2011. Understanding understanding in action. *Pragmatics* 43, 435–437.
- Lave, J., 1988. *Cognition in Practice: Mind, Mathematics, and Culture in Everyday life*. Cambridge University Press, Cambridge.
- LeBaron, C., Koschmann, T., 2003. Gesture and the transparency of understanding. In: Glenn, P.J., LeBaron, C.D., Mandelbaum, J. (Eds.), *Studies in Language and Social Interaction. In Honor of Robert Hopper*. L. Erlbaum, Mahway, NJ, pp. 119–130.
- Lindwall, O., Lymer, G., 2011. Uses of "understand" in science education. *Pragmatics* 43, 452–474.
- Lynch, M., 2006. Cognitive activities without cognition? Ethnomethodological investigations of selected 'cognitive' topics. *Discourse Studies* 8 (1), 95–104.
- Lynch, M., Bogen, D., 2005. "My memory has been shredded": a non-cognitivist investigation of "mental" phenomena. In: Molder, H.t., Potter, J. (Eds.), *Conversation and Cognition*. Cambridge University Press, Cambridge, pp. 226–240.
- Macbeth, D., 2000. Classrooms as installations: classroom instruction in the early grades. In: Hestor, S., Lee, J. (Eds.), *The Local Education Order: Ethnomethodological studies of Knowledge in Action*. John Benjamins, Amsterdam, pp. 21–71.

- Macbeth, D., 2004. The relevance of repair for classroom correction. *Language in Society* 33, 703–736.
- Macbeth, D., 2011. Understanding understanding as an instructional matter. *Pragmatics* 43, 438–451.
- McHoul, A., 1990. The organization of repair in classroom talk. *Language in Society* 19, 349–377.
- Mehan, H., 1979. *Learning Lessons. Social Organization in the Classroom*. Harvard University Press, Cambridge.
- Metzger, T.R., Beach, W.A., 1996. Preserving alternative versions: interactional techniques for organizing courtroom cross-examinations. *Communication Research* 23, 749–765.
- Moerman, M., Sacks, H., 1971 [1988]. On 'understanding' in the analysis of natural conversation. In: Moerman, M. (Ed.), *Talking Culture*. University of Pennsylvania Press, Philadelphia, PA, pp. 180–186.
- Mondada, L., 2007. Multimodal resources for turn-taking: pointing and the emergence of possible next speakers. *Discourse Studies* 9 (2), 195–226.
- Mondada, L., 2009. The embodied and negotiated production of assessments in instructed actions. *Research on Language and Social Interaction* 42 (4), 329–361.
- Mondada, L. Coordinating mobile action in real time: the timed organization of directives in video games. In: Haddington, P., Mondada, L., Nevile, M. (Eds.), *Being mobile: Movement as social action*. De Gruyter, Berlin, in press-a.
- Mondada, L. The organization of concurrent courses of action in surgical demonstrations. In: Goodwin, C., LeBaron, C., Streeck, J. (Eds.), *Embodied Interaction*. CUP, Cambridge, in press-b.
- Nishizaka, A., 2011. Touch without vision: Referential practice in a non-technological environment. *Pragmatics* 43, 504–520.
- Peräkylä, A., 2005. Patients' responses to interpretations. A dialogue between conversation analysis and psychoanalytic theory. *Communication & Medicine* 2 (2), 163–176.
- Peräkylä, A., 2008. Conversation analysis and psychoanalysis: interpretation affect and intersubjectivity. In: Peräkylä, A., Antaki, C., Vehviläinen, S., Leudar, I. (Eds.), *Conversation Analysis and Psychotherapy*. Cambridge University Press, Cambridge.
- Potter, J., 1998. Cognition as context (whose cognition?). *Research on Language and Social Interaction* 31 (1), 29–44.
- Ryle, G., 1949. *The Concept of Mind*. Hutchinson, London.
- Sacks, H., 1992. *Lectures on Conversation [1964–72]*, 2 Vols.. Basil Blackwell, Oxford.
- Sacks, H., Schegloff, E.A., 1979. Two preferences in the organization of reference to persons and their interaction. In: Psathas, G. (Ed.), *Everyday Language: Studies in Ethnomethodology*. Irvington Publishers, New York, pp. 15–21.
- Sacks, H., Schegloff, E.A., Jefferson, G., 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50, 696–735.
- Schegloff, E.A., 1982. Discourse as an interactional achievement: some uses of 'uh huh' and other things that come between sentences. In: Tannen, D. (Ed.), *Analyzing Discourse: Text and Talk*. Georgetown University Roundtable on Languages and Linguistics. Georgetown University Press, Washington, D.C., pp. 71–93.
- Schegloff, E.A., 1991. Conversation analysis and socially shared cognition. In: Resnick, L., Levine, J., Teasley, S. (Eds.), *Perspectives on Socially Shared Cognition*. American Psychological Association, Washington, DC, pp. 150–171.
- Schegloff, E.A., 1992. Repair after next turn: the last structurally provided for place for the defence of intersubjectivity in conversation. *American Journal of Sociology* 95 (5), 1295–1345.
- Schegloff, E.A., 2007. *Sequence Organization in Interaction: A Primer in Conversation Analysis*, vol. 1. Cambridge University Press, Cambridge.
- Schegloff, E.A., Sacks, H., 1973. Opening up closings. *Semiotica* 8, 289–327.
- Schegloff, E.A., Jefferson, G., Sacks, H., 1977. The preference for self-correction in the organization of repair in conversation. *Language* 53, 361–382.
- Schmitt, R., 2010. Verfahren der Verstehensdokumentation am Filmset. In: Depperman, A., et al. (Eds.), *Verstehen in professionellen Handlungsfeldern*. Narr, Tübingen, pp. 209–362.
- Stivers, T., 2008. Stance, alignment and affiliation during storytelling: when nodding is a token of affiliation. *Research on Language and Social Interaction* 41 (1), 31–57.
- Stivers, T., Mondada, L., Steensig, J. (Eds.). *Knowledge and Morality in Conversation. Rights, Responsibilities and Accountability*. Cambridge University Press, in press.
- te Molder, H., Potter, J. (Eds.), 2005. *Conversation and Cognition*. Cambridge University Press, Cambridge.
- Van Dijk, T. (Ed.), 2006. *Discourse, Interaction and Cognition*. Special issue of *Discourse Studies*, vol. 8 (1).
- Wilkinson, S., Kitzinger, C., 2006. Surprise as an interactional achievement: reaction tokens in conversation. *Social Psychology Quarterly* 69 (2), 150–182.
- Wittgenstein, L., 1953. *Philosophical Investigations*. Blackwell, Oxford.
- Zemel, A., Koschmann, T., 2011. Pursuing a question: Reinitiating IRE sequences as a method of instruction. *Pragmatics* 43, 475–488.

Lorenza Mondada is Professor at the Department for Linguistics, University of Lyon 2, and Director of the ICAR Research Lab (CNRS). Her research deals with the organization of grammatical and multimodal practices in interaction. Current research is carried out on video-recordings from various institutional and professional settings, and on ordinary conversations, focusing on the ways in which participants sequentially and multimodally organize their (often multiple) courses of action. Recent publications include: 'Working with Video', *Visual Studies* 18(1), 2003, 58–72; 'Ways of Doing: "Being Plurilingual" in International Work Meetings'. In: R. Gardner, J. Wagner (Eds.), *Second Language Conversations*. London, 2004; 'Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers'. *Discourse Studies* 9 (2), 2007, 195–226; 'Operating together through videoconference'. In: Hester, S., Francis, D. (Eds.), *Orders of Ordinary Action*. Aldershot, 2007, 51–67; *The Embodied and Negotiated Production of Assessments in Instructed Actions*, *Research on Language and Social Interaction*, 42 (4), 2009, 329–361; *Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space*. *Journal of Pragmatics* 41, 2009, 1977–1997.