Challenges of multimodality: Language and the body in social interaction[1]

Lorenza Mondada
University of Helsinki, Finland and
University of Basel, Switzerland

This article reflects on recent challenges emerging from the study of language and the body in social interaction. There is a general interest in language and the body across disciplines that has invited a reconceptualization of the broader issues relative to action, cognition, culture, knowledge, social relations and identities, spatiality and temporality. The study of social interaction focuses on how multimodal resources – including language and bodily movements – are holistically and situatedly used in building human action. This article discusses some consequences and challenges of putting the body at the center of attention: it repositions language as one among other modalities, and invites us to consider the involvement of entire bodies in social interaction, overcoming a logo-centric vision of communication, as well as a visuo-centric vision of embodiment. These issues are developed through a series of conversation analytic studies, firstly of classic topics in linguistics like deixis, then of more recent topics, such as mobility and sensoriality.

Cet article offre une réflexion sur des enjeux récents émergeant de l'étude du langage et du corps dans l'interaction sociale. Un intérêt général pour les rapports entre langage et corps est apparu à travers les disciplines des sciences sociales et cognitives, qui a invité à re-conceptualiser des enjeux fondamentaux touchant à l'action, la cognition, la culture, la connaissance, les relations sociales et les identités, la spatialité et la temporalité. L'étude de l'interaction sociale y répond en se focalisant sur la manière dont les ressources multimodales – incluant le langage et le corps – sont mobilisées de manière holiste et située dans l'organisation de l'action humaine. L'article discute d'un certain nombre d'enjeux et conséquences issus de la remise du corps au centre de l'attention : cela repositionne le langage comme une modalité parmi d'autres, et invite à se pencher sur l'engagement du corps entier des participants dans l'interaction sociale, dépassant ainsi une vision logo-centrique de la communication, ainsi qu'une vision visio-centrique de la corporéité. Ces enjeux sont développés dans une série d'études en analyse conversationnelle portant sur d'abord sur des thèmes classiques en linguistique, comme la deixis, puis sur des problématiques plus récentes, comme la mobilité et la sensorialité. [French]

KEYWORDS: Social interaction, multimodality, language, body, conversation analysis, temporality, sequentiality, mobility, sensoriality

1. INTRODUCTION

The role of language and the body in social interaction has been revisited in recent research, with proposals for new ways to conceptualize how communicative resources are mobilized by speakers and other participants to build their actions within social interaction, and ultimately how human action is organized. The objective of this article is to highlight some conceptual and analytical issues raised by the analysis of multimodality and the use of video for the study of human action in interaction.

As a starting point, in this introduction (section 1) I briefly sketch the issues introduced by the recent attention paid to embodiment in linguistics and beyond, and, by multimodal analyses within an interactional perspective, with a special focus on the contribution of Conversation Analysis. Next, the challenges of multimodality are developed by showing how they concern not only gesture, gaze and talk but much more complex arrangements involving the entire body and multiple bodies in interaction (section 2). Two fundamental issues are then discussed: mobility (section 3) and sensoriality (section 4). Taken together, they invite us to adopt an integrated holistic approach to multimodality (section 5 – Conclusion).

1.1 The embodied turn in social sciences

Since the 1960s, several disciplines, both inside and outside linguistics, have been interested in language not as an abstract system but rather as a social communicative discursive practice (Leeds-Hurwitz 1987). Pragmatics, ethnography of communication, interactional sociolinguistics, conversation analysis, discourse analysis, critical discourse analysis and nexus analysis have all introduced an interest in language as related to its actual use in communicative settings, within socio-cultural practices, and situated in specific social contexts. This has also coincided, within neighboring social sciences, with an increased interest in language and discourse, viewed not only as a way of accessing social life, but also as a form of social action.

Most recently, this interest in language as used in actual contexts of communication has included consideration of the embodied way in which people communicate and gather together, as well as the ecology of the activities they engage in, and their material and spatial environment. This introduces both new challenges for disciplines traditionally focused almost exclusively on language, and new opportunities for theoretical re-conceptualizations of corporeal aspects of agency, critical of disembodied Cartesianism. This has fueled new contributions in the social sciences on human action, communication, affect and experience, gender and identity,

space and culture (Crossley 2001; Shilling 2003; Thrift 2007). In cognitive sciences too, embodiment has been used, e.g. by Andy Clark (1997), to propose an alternative agenda, 'putting brain, body and world together again' as featured in the subtitle of his book, along with a critique of the computational representational abstract conception of the mind, as opposed to a vision that recognizes the primacy of sensory, perceptual and motor experiences both for the mind and the brain.

Within this embodied turn, attempts to integrate previous studies of language and new studies of the body have occasioned various approaches to the *multimodality* of social action and social communication. *Multimodality* is a term used in very different ways within several epistemological and disciplinary fields (such as computer sciences, logistics and transport, discourse analysis, and studies of interaction). Within a perspective inspired by computer sciences, multimodality might refer to diverse interfaces, channels and mediums of communication. Within a perspective inspired by semiotics, it refers to the materiality of the representation characterizing various medias as well as to different types of signs providing and affording diverse semiotic effects – such as texts, fixed images, moving images and multimedia messages (Kress and van Leeuwen 2001). Within perspectives inspired by gesture studies and the study of social interaction, such as conversation analysis, the term is used to refer to the various resources mobilized by participants for organizing their action – such as gesture, gaze, facial expressions, body postures, body movements, and also prosody, lexis and grammar. The plurality of 'modalities' referred to in this term treats *multi*modality as constitutive and primary. This encourages a view of modalities as constitutively intertwined, and language as integrated within this plurality as one among other resources, without any *a priori* hierarchy (Mondada 2014a).

1.2 The contribution of Conversation Analysis

Among a huge variety of approaches to human action and social interaction, Conversation Analysis (hereafter CA) constitutes an original perspective representing a form of interdisciplinarity in itself – since it emerged in sociology (Sacks 1992; Sacks, Schegloff and Jefferson 1974) in close relation with ethnomethodology (Garfinkel 1967), and later developed in a variety of disciplines, such as anthropology, communication, education, and most notably linguistics, where it has given impulse to a new sub-field, interactional linguistics (Ochs, Schegloff and Thompson 1996).

The fingerprint of CA (see Sidnell and Stivers 2014 for a comprehensive view) consists of its focus on action: it studies situated action as it happens in both its systematic organization and its diversity within various social settings. Action is understood as being organized not by individual speakers but within social interaction. Consequently, careful attention is paid to how co-participants manage actions in a sequential way, moment by moment, through the way an

action recognizably responds to another one or initiates a new sequence, through the publicly accountable arrangement of resources (Schegloff 2007).

By showing how the mobilization of linguistic resources can be accounted for in the way: (a) *turns* are formatted (making identifiable their emergent and projectable elaboration and thus making it possible for co-participants to recognize possible completion and transition relevance points for managing turn-taking); (b) *actions* are constituted (making actions recognizable and permitting co-participants to produce relevant next actions); and (c) *sequences* of actions are organized, CA has offered a renewed approach to linguistic resources within phonetics and prosody, syntax, lexis and semantics, as well as pragmatics (Couper-Kuhlen and Selting 1996; Ford, Fox and Thompson 2002; Hakulinen and Selting 2005).

More recently, CA has been interested in considering not only verbal resources but also embodied ones, within a more global and holistic approach of multimodality, comprising language, gesture, gaze, head movements, facial expressions, body posture, body movements, and embodied manipulations of material objects (Goodwin 2000; Streeck, Goodwin and LeBaron 2011). This has recently been referred to as an 'embodied turn' (Nevile 2015), fostered by the increasingly systematic use of video methodologies.

Video – film in earlier times – has been a powerful technology that has been embraced by different disciplines in the social sciences, within what can be called a 'visual turn' (Mondada 2013). For CA, it has represented a way to strengthen its focus on naturally occurring interactions, recorded in their original social context, without being orchestrated by the researcher. This focus comes from the conception of action as situated, indexically organized, and specifically shaped by, as well as shaping, the social and material context in which it happens. Consequently, the documentation of naturally occurring actions as they happen in their setting – and not as they could be elicited by researchers in various manners – responds to the fundamental principle of indexicality of action.

This documentation mainly relied on audio recordings during the first decades of CA history, with the exception of some notable early uses of film (chief among these were the films produced by Charles and Marjorie Goodwin as early as the 1970s, which circulated among the first scholars of CA – see Goodwin 1981 – and other early studies, see Heath 1986; Sacks and Schegloff 2002; Schegloff 1984). Since the start of the 21st century, there has been an increase in video studies (Deppermann 2013; Heath and Luff 2000; Streeck, Goodwin and LeBaron 2011).

However, CA was far from the first approach to make use of video – a crucial precursor was the *Natural History of an Interview* (McQuown 1971; see Leeds-Hurwitz 1987), a project in which an interdisciplinary team of scholars engaged in the detailed study of a filmed session in which Gregory Bateson converses with his patient, Doris. This project gave birth to a series of important studies of embodiment, such a kinesics (Birdwhistell 1970) and context analysis (Scheflen 1972). Other precursors are identifiable even earlier,

since the birth of cinema, which was immediately used for scientific purposes in the social sciences, for instance by Etienne-Jules Marey and Muybridge to study movement and by Alfred Cort Haddon in his Torres Strait expedition as early as 1898 for the study of dance and other rituals (Mondada 2013).

What is distinctive about CA use of video is the careful and precise attention to temporally and sequentially organized details of actions that account for how co-participants orient to each other's multimodal conduct, and assemble it in meaningful ways, moment by moment. This has consequences for how videos are produced by CA researchers (Broth, Laurier and Mondada 2014; Heath, Hindmarsh and Luff 2010; Mondada 2006), which differs from other, more individualistic ways of focusing on single speakers, or more discontinuous ways of documenting a setting rather than the temporal continuity of actions. This also has consequences for how video materials are exploited within the analysis, and, crucially, how transcription shapes the conditions and possibilities for analysis (see Ochs 1979).

1.3 Some consequences

Together, the interactional turn and the visual turn have important consequences for all disciplines, and in particular for linguistics and for CA.

For linguistics, multimodality opens up new avenues for the analysis of well-studied as well as new phenomena. Although linguistics recognized early on the relevance of gesture for the study of some topics – such as deixis – this has been limited to very specific areas. Throughout its history, linguistics has been heavily based on the idea of the autonomy of language and has largely limited communication to language: this has produced a *logocentric* view of language. In contrast to this view, it is possible to see language and talk as fundamentally embodied: producing talk involves visible breathing and articulating movements not only of the face and the mouth, but of the entire body; moreover, these articulatory movements are indissociable from other bodily conduct (as gesture studies asserted very early on, both talk and gesture originate from the same process; McNeill 1992; Kendon 1990). Under this view, no aspect of language escapes a multimodal perspective: multimodality allows interactional studies to potentially revisit all of the fields of linguistics – not only deixis, but syntax, semantics, even phonetics and prosody – as well as writing and textuality (Komter 2006; Mondada and Svinhufvud 2016).

In addition to these wider consequences for linguistics as a whole, multimodality affects key aspects of CA in particular. From its very beginning, CA has been represented as a discipline that was not interested in language *per se,* but in how action is formatted in such a way that it is publicly recognizable and thus responded to by the co-participants (Garfinkel 1967; Sacks 1992). Although the early history of CA focused primarily on the study of talk and linguistic resources, this conception of the *accountability of action* is not *a priori* restricted to verbal means. Quite the opposite, this

accountability can be achieved in a situated and indexical way, thanks to multiple possible *resources* that are made locally available and relevant by the ecology of the activity and that are used and oriented to in an endogenous way by the participants. This emic and indexical conception of 'resources' for action (Mondada 2014a) entails that:

a. there is no principled priority of one type of resource over the others (e.g. of language over embodiment);
b. potentially every detail can be turned into a resource for social interaction, depending on the local ecology and the activity, though some resources might be more conventionalized than others, such as grammar or routinized/ritualized embodied actions; and, finally,
c. some ecologies and types of activities might favor verbal resources along with gestures and body movements, whereas other ecologies and activities might favor distinctive and specific embodied resources over talk (including actions achieved without a word; cf. Ivarsson and Greiffenhagen 2015).

Even if there is no principled restriction to talk in CA, the very fact that many of the conceptual and analytical advances of CA have been based on audio data and verbal resources has given a specific flavor to the conceptualization of key praxeological aspects, such as temporality and sequentiality. The challenge of multimodality is that it transforms our vision of the temporal and sequential organization of action: whereas studies focused on talk might emphasize a linear and successive conception of how action unfolds moment by moment (with no gap, no overlap; Sacks, Schegloff and Jefferson 1974), studies focused on multimodally accountable action highlight forms of sequentiality that integrate and intertwine multiple simultaneous sequentialities and temporalities, within which complex forms of projective and responsive actions are organized. As the analyses in this article will show, these temporal arrangements, which achieve the intersubjective organization of interaction, are the fundamental principles shaping human action.

The remainder of this paper discusses some of these challenges and advances made possible by a multimodal interactional approach to video data, first by showing how it can revisit classic topics, such as deixis in relation to language and gesture (section 2), and then by discussing some more recent issues, such as mobility, in relation to moving bodies (section 3), and sensoriality, in relation to sensing bodies (section 4).

2. FROM GESTURE TO THE ENTIRE BODY: EMBODIED SEQUENTIALITIES

Deixis is a phenomenon in which the importance of gesture was recognized very early on. Deictic expressions are commonly treated as co-occurring with gestures, especially when pointing. One can then wonder which specific contribution an interactional perspective, and more specifically multimodal CA, can offer in this area (Goodwin 2003; Hindmarsh and Heath 2000; Mondada 2014b).

Let us look at Excerpt 1a, a fragment in French from a meeting in which agronomists are working at a table covered by maps. In each turn, a deictic term (either a locative demonstrative or a demonstrative pronoun) is used, and co-occurs with a pointing gesture, indicated with ((ptg)) in the excerpt (see the Appendix for transcription conventions).

Excerpt 1a

```
1  VIV   .hh ↑oui. parce que: i'*m'semble: eh i- ici ((ptg)) c'étai::t
         .hh yes. because it seems to me uh he- here ((ptg)) it was::
2        qui- c'que ça voulait représen[ter, c'était
         w- what this was meant to represen[t, [it was
3  LAU                                   [c'est des am*andes ça? ((ptg))
                                         [are these almonds? ((ptg))
4  VIV   oh ça ((ptg)) c'était des: amandes, [c'était aussi ((cont.))
         oh these ((ptg)) were almonds, [it was also ((cont.))
5  LAU                                       [ouais
                                             [yeah
```

While saying *ici* 'here', Viviane points towards a specific location on the folder containing maps (line 1); in turn, when Laurence makes a request for confirmation (line 3), she points at the referent of the demonstrative pronoun, as does Viviane in her answer too (line 4). This simplified transcript highlights the co-occurrence of the deictic term and the gesture ((ptg)).

But the situation appears quite different if we adopt another mode of transcription, shown in Excerpt 1b, which identifies, locates and describes gestures and other embodied conducts on separate time lines that run parallel to talk, showing how they are coordinated with it but not coinciding with its units, as was suggested by Excerpt 1a. Moreover, by showing exactly when a gesture starts and ends (in the transcript, this is designated by a series of .... showing the emergence of the gesture; a series of ,,,, show its withdrawal), the entire trajectory is temporally described in detail and related to the temporality of talk (for a full explanation of the conventions, see the Appendix).

The transcription in Excerpt 1b shows that much more than just referring is going on. First, pointing does not merely co-occur with the deictic element: Viviane's gesture starts already at turn-beginning (Figures 1.1–1.4). Her rising hand begins to move precisely with her in-breath at the beginning of line 1 ('.hh'), being almost a bodily manifestation of it. Likewise, when Laurence, in overlap (line 3), initiates a repair sequence, she starts to point at turn-beginning (Figure 1.6) in such a way that she is fully pointing with her finger when the demonstrative *ça* 'this' is uttered (Figure 1.7). Moreover, we notice that Viviane continues to point during the repair sequence and she resumes her own turn, while Laurence continues to point until the completion of the sequence she has initiated. This indicates that rather than being a mere referential device, pointing is also a way of displaying speakership, of holding the turn and even exhibiting specific rights over the sequence (Mondada 2007).

Excerpt 1b

```
1  VIV    #+.hh   #+↑oui. parce que:# i'*m' #sem+**ble: eh i- **ici# c'étai::t
           .hh yes. because it seems to me uh he- here      it was
          +raises H+....comes in w pen.......+points w pen---------->>
   lau                                *opens folder-------*holds folder->
   lau    >>writes----------------------------**stops w----**
   fig    #fig1.1 #fig1.2          #fig1.3 #fig1.4              #fig1.5
```

VIVIANE PIERRE-ALAIN


1.1


1.2

LAURENCE

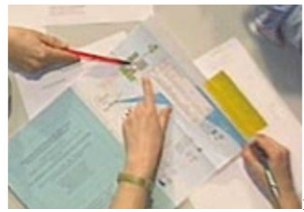

1.3


1.4


1.5

```
2          qui- c'que ça voulait représen[*ter,# c'était
           w- what this was meant to represen[t, [it was
3  LAU                                    [*c'est des am*andes ça?#
                                          [are these almonds?
   lau                                     *............*points w finger-->
   fig                                          #fig1.6          #fig1.7
```


1.6


1.7

```
4  VIV    oh ça c'était des: aman*des,* [c'était aussi l'idée que ((cont.))
          oh these were almonds, [it was also the idea that ((cont.))
5  LAU                                  [ouais
                                        [yeah
   lau                          -->*,,,,*
```

Second, the coordination between the gesture and the deictic *ici* 'here' is an interactional accomplishment: Viviane progressively approaches a location on the map that is not yet accessible, as it is being folded by another participant, Laurence, who is also taking notes during the turn (Figures 1.1–1.4); the turn format used by Viviane is stretched and delayed in such a way that it invites and allows Laurence to open the folder, and stop writing (Figure 1.5) exactly when *ici* 'here' is produced.

Third, the trajectory of the gesture involves a specific way of holding an object, the pen: Viviane's pointing gesture begins with her hand rising, holding the pen in a 'writing position' (Figures 1.1–1.2); then, she advances by rotating the tool (Figure 1.3), in such a way that the top of the pen prolongs her arm and finger (Figure 1.4) pointing towards the folder (Figure 1.4–1.5) in a way that does *not* project writing (here the pen becomes a pointer rather than a writing tool).

So even very classic instances of pointing reveal, within an interactional multimodal perspective, that much more than co-speech gestures are involved. Pointing does not only concern a gesture of the hand: it involves the body leaning over the table, the stretched arm, the hand, and specific fingers as well as the manipulation of objects. These relevant detailed patterns, together with specific linguistic constructions and body postures in given sequential environments constitute *complex multimodal Gestalts* (Mondada 2014a). These Gestalts are synchronized not only for doing reference, but for the interactional organization of turns, sequences and collective activities. Multimodal Gestalts arranged in space and time build emerging and changing positionings between the participants, whose relations, actions, and the rights and obligations related to them, are negotiated not only in discursive but also in embodied ways: an action can be aligned or disaligned verbally, but also bodily, disclosing subtle socio-interactional dynamics.

As we shall observe in the next excerpt – Excerpt 2 – taken from a walk in a garden, pointing involves the entire bodies of the participants. Here, the gardener, Luc, is explaining problems related to the use of pesticides when he spots a butterfly (line 2), which he points out to his co-participants. At the beginning of this short excerpt, the participants' bodies are arranged around Luc in a circular configuration, looking at him (Figure 2.1) while he is engaged in his explanation (lines 1–2). When he spots the butterfly, he interrupts his ongoing talk (*peu-* 'bit-', line 2) and instead utters an imperative (*r'gardez* 'look', line 2) while beginning to point at the object referred to. We notice that Luc does not just use a co-speech gesture co-occurring with the verbal reference. Rather, he moves his entire body, stepping forward (Figures 2.2–2.4) in such a way that his extended finger prolongs his extended body. Moreover, the way he utters the name of the insect (line 2) is characterized by a stretched article, an in-breath, the self-repair of the article, followed by the name and a locative deictic (*le: .hh le papillon bleu là* 'the: .hh the blue butterfly there', line 2). This way of formatting his turn extends it during the time it takes to the co-participants to change their bodily and spatial positions: they progressively look at him and pivot their bodies, in such a way that by the end of Luc's turn-constructional unit (end of line 2) they are all re-oriented towards the referent (Figure 2.5). The last to move is Jean (Figure 2.4). The cameraperson moves too, panning towards the left. This shows that the camera constitutes another participant adjusting and responding to the ongoing action, interpreting it in real time (Mondada 2006).

Excerpt 2

```
1  LUC   y a des limites, quoi:, là là on est# on est un
         there are limits, right, there there we are we are a
   fig                                        #fig2.1
```

JEAN YAN ELISE LUC

2.1

```
2        pξeu- *r'gardez +%le: #ξ* .hβh+h ξ #le papi%β©llon% +bleu là:*#+
         bit-  look at     the:    .hhh     the blue butterfly there
          ξ.....................ξpoints---ξpoints and walks twd insect-->
   luc         *one step fwd---*another step fwd---------------------*
   eli           %looks--------------------%pivots%
   yan                        +looks----------------+pivots-----+
   jea                       βturns H back---βpivots-->
   cam                                        ©pans to right--->>
   fig                  #fig2.2     #fig2.3              fig2.4#
```

2.2

2.3

2.4

```
3        c'st* unβ argus. >voyez<?#
         it's an argus. see?
           ->*points in a static position-->>
   jea         ->β
   fig                        #fig2.5
```

2.5

The progressivity of Luc's turn is *reflexively* organized with respect to the co-participants' responses: he adjusts to them as they respond to it. As soon as they all look at the butterfly, he gives its name and ends his utterance with an accelerated >*voyez?*< '>see?<' (line 3) – both forms expecting that the participants are now in a relevant position to grasp it, within the newly established interactional space (Mondada 2009).

As with Excerpt 1, Excerpt 2 demonstrates several important contributions of the multimodal perspective. First, pointing is not just a gesture of the hand but might involve the entire body. More generally, bodies' movements and positions are highly relevant for the organization of talk and action, including their arrangement within the interactional space. These movements draw trajectories that are deployed in time, and which are synchronized with the incremental structure of talk.

Second, actions, and responses to actions, can be implemented verbally but they might also be silently embodied. Indeed, the progressivity of Luc's turn and action is reflexively adjusted to the co-participants' silent embodied responses in real time: his invitation to look is stretched and delayed while the co-participants are still turning and repositioning themselves in space; and his delivery of the butterfly's name orients to their new position.

This shows that:

a. temporality and sequentiality are the fundamental principles governing social interaction, including when it is formatted through embodied resources; and
b. sequentiality might not be organized turn-at-talk after turn-at-talk, strictly successively, but rather in parallel flows of action, as emergent embodied conduct responds to a previous action and unfolds simultaneously with it (see Goodwin 1979 for an early example).

Contrary to interactions analyzed within an exclusive focus on talk, this implies a plurality of temporalities and sequentialities progressing at the same time within the local emergent design of complex multimodal Gestalts (Mondada 2014a, 2014b). Sequentiality is a less linear phenomenon than it appears just on the basis of talk; it relies on subtle ways of arranging and adjusting prior and next actions in real time, among differing, and yet coordinated, simultaneous multimodal conduct reflexively intertwined together.

The precise transcription of the timing of these phenomena is fundamental to catch the exact position, within a multimodal Gestalt, of emerging resources – this is crucial for the interpretation of actions as either initiating a sequence or responding to previous actions. This constitutes both a challenge and a new opportunity for sequential analysis: an action can be initiated before the turn is actually uttered; an action can be responded to very early, and even responded to *while* the previous action is being produced, either in the form of a turn or an embodied movement. Likewise, actions can be mutually adjusted in real time, as they progress – within complex micro-sequentialities.

More broadly, these phenomena show the importance of temporality for the understanding of language and the body in interaction: syntax and body movements are finely organized as emergent multimodal Gestalts, temporally coordinated within and between speakers. These Gestalts are both systematic (it is possible to describe their multimodal praxeological grammar) and deeply embedded in the specific ecology of the activity – since they are adjusted to its material and spatial environment, including the way it materializes the socio-institutional context.

3. WALKING BODIES: MOBILITY

The study of interacting bodies is an opportunity to radically overhaul the traditional view that reduces social interaction to talking heads or talking torsos – inviting a revised focus, from limited gesture and head movements to extended body postures and movements. In this respect, the study of walking-in-interaction (Ryave and Schenkein 1974) is promising because it deals with the entire bodies of the participants. Moreover, it introduces an important switch from rather sedentary activities (like dinner conversations or meetings) to mobile ones, within what has been termed the 'mobility turn' in the social sciences in general (Cresswell and Merriman 2011; Urry 2007), and in CA in particular (Haddington, Mondada and Nevile 2013; McIlvenny, Broth and Haddington 2009).

The analysis of walking-in-interaction offers important hints for the multimodal study of the organization of actions:

- In walking, multiple bodies are coordinated together, constituting 'mobile withs' (Goffman 1971) – which build complex participation frameworks (Goffman 1981; Goodwin and Goodwin 2004) in multi-party interactions.
- This has consequences for conceptualizing participation frameworks as situated assemblages of bodies within space, in what has been described as an F-formation (Kendon 1990), contextual configuration (Goodwin 2000), and interactional space (Mondada 2009), designed by their positions and changes as they move.
- Bodies constitute a holistic repository of multiple resources (gestures and gaze in mutual attention and monitoring, body positions and trajectories, steps and moves), mobilizable for actions and responses, possibly within different parts of the body (body-torque; Schegloff 1998).
- Talking and walking constitute a specific form of multiactivity (Haddington et al. 2014) that allows an exploration of the locally ordered distribution of resources within and across complex actions (Mondada 2014c).
- Walking itself is organized as micro-practices: stepping, walking, walking backwards, accelerating, slowing down, etc., which have specific sequential trajectories (Mondada 2014d).

Excerpt 3a shows how interacting bodies mobilize a wide range of multimodal resources within complex but systematic patterns. We join a walk in the countryside at the border between Switzerland and France, near Basel, in which Yves, a local inhabitant, guides four visitors from the city. The excerpt is video-recorded by three camerawomen, two following the group and one walking ahead (the latter shot is used for this analysis). The excerpt shows how a simple sequence, a question asked by Sophie and answered by Yves, mobilizes a multiplicity of multimodal details when asked on the move (Mondada in press).

Excerpt 3a

```
1  SOP   mais donc y avait& un pro#jet d'un% d'un train euh& entre les deux%
         but so there was a project of a, a train ehm between the two
   sop   >>looks at YVE----->>
   yve   >>looks at SOP and the group-->>
   lis                  &looks at SOP--------------------&
   mar                                  %looks back and walks sideways--%
   fig                           #fig3.1
```

MARIE SOPHIE ANN LISA YVES

3.1

```
2          pays? avant la guerre. °déjà?°
           countries? before the war. °already?°
```

Sophie asks a question (lines 1–2) about a topic Yves had mentioned earlier: a pre-war (First World War) train service between the countries, connecting Basel in Switzerland and Ferrette in France. At this point, the group is walking together side-by-side, as a 'with'. However, this mobile formation is not homogeneous (Figure 3.1): Marie is walking ahead; Ann and Lise are walking a bit behind; Sophie and Yves are walking along two opposite sides of the footpath and look at each other. This disposition shows that within the group, two participants are mutually bodily oriented – Sophie and Yves – while others are positioned as overhearers. Actually, Lisa looks at Sophie after she has started her question, and Marie looks back (walking sideways like a crab) a bit later: both orient to Sophie as initiating a new action. In a nutshell, the dynamic interactional space drawn by walking bodies achieves and exhibits agency, with speakership and recipiency characterizing this first pair part.

Let us now look at how the question is responded to, in Excerpt 3b. While progressing along the footpath, Yves starts his answer with the prepositional phrase (*avant la première guerre* 'before the First World War', line 3, cf. supra

Excerpt 3b

```
3  YVE   .h avant la *première guerre, eu gh* &c'q- $©c'était# pré&vu
         .h before the  first war,     eh gh  what w- it was planned
   yve   >>walks fwd-*slows down------------*advances on the side--->
   lis                                      &looks at YVE-------&
   ann                                             $looks at YVE--->
   cam   >>follows the group, behind----------------©moves aside-->>
   fig                                                    #fig3.2
4        de faire un* &$otr#ain, (0.3) +%bâ:le,o+ (0.3)o %ferrette.$&#
         to build a train,       (0.3)   basel,   (0.3)   ferrette.
                  ->*stops-->
                        oRH twd R--------------o,,,,,,,oLH twd L-->
   sop                                  +pivots & stops+
   mar                                   %pivots & stops---%
   lis                  &slows down, looking at YVE---&stops----------&
   ann                 ->$slows down, looking at YVE---$stops--------$
   fig                       #fig3.3                          fig3.4#
```


3.2


3.3


3.4

line 2) used by Sophie to complete her question, thereby tying back the second pair part to the first and providing a confirmation with it. However, he does more than confirming. The first part of the turn is uttered with a prosody that projects more to come: after some hitches, he produces a syntactically complete utterance, synthetizing the train project (lines 3–4). During this utterance, he modifies his way of walking: he slows down and begins to move sideways, turning his whole body, not only his head/gaze, towards the group (Figure 3.2). He gradually comes to a halt (line 4), and in a stationary position, turned to the group, he uses his body as a signpost, indicating two opposite directions with his left and right arms while uttering the two toponyms of *Bâle, (0.3) Ferrette* (Figures 3.3–3.4).

The co-participants respond to Yves' new positioning: first, they look at him (Figure 3.2, line 3), then they slow down their walk (Figure 3.3, line 4), and finally, still looking at him, they stop (Figure 3.4, line 4). They align with him in a double sense: as recipients of his answer, they align to his action; as mobile bodies, they align with his new position, stopping on a line in front of him.

Again, the camerawoman following the group anticipates what Yves will do, moving very early on to the side (line 3), in such a way that later on she is able to film from behind the group, facing Yves (line 10). In this way, she prefigures and adjusts to the new interactional space Yves is achieving by transforming a mobile walk side-by-side into a sedentary position of his body facing the group. Again, the cameraperson acts as a participant, anticipating and interpreting what is going on – and as such is integrated into the transcript.

Finally, the fragment continues in Excerpt 3c, with Yves engaging in a longer explanation.

Excerpt 3c

```
5          (0.4)
6  ?       ah:, c'est (  )o
           oh:, it's (  )
   yve                 -->oholds LH, RH shows from R to L->
7  SOP     avant la premoière?
           before the first?
   yve                 ->o
8  YVE     oui.
           yes.
9          (0.6)
10 YVE     .h et ©après, avec la guerre, on a abandonné.
           .h and after, because of the war, they abandoned.
   cam         ->©stops in front of YVE, behind the audience--->>
11         (0.3) % (0.6)
   mar           %comes closer to the group->
12 YVE     .h et, et voilà, on n'a# (.) jamais% repris cette idée ((cont.))
           .h and, and that's it, they never went back to this idea ((cont.))
   mar                                        -->%
   fig                              #fig3.5
```

3.5

Yves's short summary of the train project (line 4) is responded to verbally in various ways (lines 6, 7). At this point the sequence could be closed, but Yves adds a new detail ('.h and after, because of the war, they abandoned', line 10), and continues his explanation (line 12).

As he engages in a more substantial multi-unit turn, Marie, who stopped at the same time as the others (line 4), but a bit ahead (Figure 3.4), walks back and joins the group (lines 11–12, Figure 3.5), aligning with the other recipients. Now, the new interactional space is achieved as a clear symmetric figure in which two *parties* are facing each other, the 'performer' and the 'audience' (Goffman 1981; Sacks 1992: Vol II 104ff.). This reflexively supports Yves' public speech, suspending the turn-taking machinery for a moment, giving him the right to talk extensively, and shaping his talk as a type of lecture. The normativity of this new configuration is displayed in Marie's step to join the others as a hearer.

In sum, in this excerpt, the finely tuned movements of different parts of the body, coordinated with talk, achieve together several transformations concerning the ongoing action and relevant participation framework – from a question-on-the-move involving two participants and overheard by the others, to a stationary explanation addressed by one party, the performer, to another party, the audience. These spatial and embodied constellations are

complex – involving multiple layers of multimodal resources – but they are methodically produced by the participants within an articulation that mediates between their situatedness anchored in the specificities of the local ecology and their systematicity. This is observable in the next excerpt, Excerpt 4, showing how the same praxeological and sequential dynamic environment – the transition from a question on the move into a stationary explanation – can be achieved within another configuration, involving a set of complementary resources and practices.

In Excerpt 4, we join a group of citizens inspecting a construction site within an urban project they have been involved in. The architect (Ligour) is walking ahead; a citizen (Fonseca) follows at some distance (Figure 4.1). In the excerpt, Fonseca accelerates to pose a question to Ligour. After completing the answer provided on the move, the latter transforms the mobile interactional space into a stationary disposition of the audience in front of him as a performer.

About to self-select, Fonseca accelerates towards Ligour. He begins to ask a question about a wall situated behind them, as he is still behind Ligour (Figure 4.2). The format of his turn – a noun phrase (NP), *le: le: début du mur là* 'the: the: beginning of the wall there' (line 2) – is adjusted to this particular spatial configuration: he uses a turn-initial article stretched and self-repaired, introducing the name of the referent, followed by a locative deictic uttered with a pointing gesture. All these elements delay the progress of the turn and favor the attention of the recipient both on the speaker and on the referent/place, *before* he moves on to the core content of the question (*euh euh vous allez la raser complètement? ou:::* 'ehm ehm you're going to erase it totally? or:::', line 6), which retrospectively configures the initial NP as a left-dislocated element.

Ligour turns back even before the first part of the question is initiated (line 2), probably having heard the steps quickly approaching from behind. This preparatory move is perceptible as such by its addressee – which shows the value of analyzing (still largely neglected) features of audible sounds and their spatialization (particularly challenging for video analysis since recording these would rely on the positions of the microphones and their possible artefactual effects).

First, Ligour turns his head (Figure 4.2); then, he pivots with his entire body (Figure 4.3) and starts to walk backwards (Figure 4.4). Only after this reorientation, he responds verbally with *ouais?* 'yeah?' (line 4), as Fonseca reaches him (Figure 4.5). When they are side by side, one walking backwards and the other walking forwards, Fonseca produces the core of his question (lines 6–7). So, the distribution of his turn in two parts (line 2 vs lines 6–7) adjusts the syntax to the mobile configuration of the participants. Ligour initiates a repair (line 8) on spatial reference, and points at a wall as possible referent (lines 8–9, Figure 4.6). Fonseca achieves the repair, pointing to an alternative wall (lines 9–11) and pivoting back towards the pointed at direction (Figure 4.6); he begins to walk backwards too.

Excerpt 4

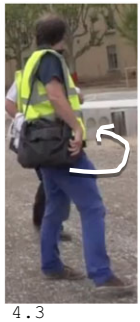
```
1          (64.0)            +# (4.9)
   fon     >>walks behind LIG+accelerates->
   fig                        #fig4.1
```

FONSECA LIGOUR

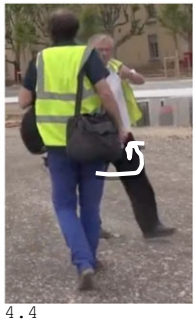
4.1

```
2  FON     l*#e: le: début# du mur ++là.*#
           the: the: beginning of the wall there.
                                    ++points--->
   lig      *turns back and pivots------*walks backwards-->
   fig       #fig4.2      #fig4.3        #fig4.4
```

4.2



4.3



4.4

4.5

```
3          (0.6)
4  LIG     ouais?+ #
           yeah?
   fon        -->+reaches LIG->
   fig             #fig4.5
5          (0.3)
6  FON     euh euh vous++ allez +la raser complètement? ou:::
           ehm ehm you're going to destroy it totally? or:::
                   --->++
                              -->+
7          (0.3) vous le descende:z? e:t=
           (0.3) you take it down? and:=
8  LIG     o=c'ui qui eu:::#[h le]
           =the one that eh:::[h PRO]
9  FON                       [non ] le +le °premier qui est⊥ là.°#
                             [no] the the °first that is there.°
   lig     o...points in front of him->
   fon                               +points and pivots----->
   fon                                                  ⊥walks bckwd-->
   fig                        #fig4.6                   fig4.7#
```


4.6

4.7

```
10          (0.5)
11 FON      °c'ui qui part de là-bas.°+⊥
            °the one that starts from over there.°
                                  --->+⊥
12 LIG      eu:::::::::h là:?o y a- le projet n'est pas fini.
            eh:::::::::m there:? there i- the project is not finished.
   lig                    -->o
13          on va rempli:r, et on sera au niveau.
            we'll fill it in, then it will be level.
14          (0.2)
15 FON      °vous êtes à niveau.°
            °you will be level.°
16 LIG      on sera à niveau.
            we'll be level.
17          *(0.4)
   lig      *looks in front of him--->
18 LIG      on peut faire# un tout* petit arrêt ici, peut-être
            we can make a quick stop here, maybe
                                --->*looks on his left->
   fig                    #fig4.8
```


4.8

```
19          au cen*tre,o (0.7)
            in the middle, (0.7)
                ->*looks in front, to the group-->>
                   --->ostops->>
```

Now both participants walk backwards side by side (Figure 4.7), while Ligour finally delivers an answer (lines 12–13). Fonseca walks forwards again as soon as the repaired reference has been established and the answer has started, while Ligour holds this position much longer. Both participants adopt a 'mobile body-torque' (Mondada in press) exhibiting a double orientation, back and forward. But while walking backwards for Ligour constitutes a way of maintaining a dual orientation towards the progress of the visit, of which he is in charge, and the possibly delaying sequence initiated by Fonseca, for Fonseca, walking backwards merely secures the repair of the referent and constitutes a prolongation of his pointing gesture. This difference is visible in the temporality of their movements, which last only over the completion of the repair sequence for Fonseca, as opposed to over the entire sequence for Ligour.

Moreover, as soon as Ligour closes the sequence (line 17), he moves his gaze from Fonseca to the group. He suggests stopping (line 18), chooses a location (line 19), actually stops and projects more to come (line 19). The group stabilizes and Ligour gives a general explanation (not shown here), explicitly targeting the plurality of citizens now assembling around him (Figure 4.8). In this way, Ligour transforms his walking backwards into a stationary position facing the group, transitioning from a private question/answer-on-the-move to a public explanation delivered within a stabilized and immobilized interactional space.

As shown by these analyses, mobility does not just constitute an expansion of sedentary actions; it impacts on the way actions and sequences are initiated and completed, turns are formatted, and relevant resources for organizing them are chosen. Mobility involving the entire body shows how turn formats, action formation, sequence organization and the organization of participation are emergently and progressively built, how they mobilize a variety of multimodal resources (widely exploiting the contingencies of the dynamic ecology of the activity), and how they design specific interactional spaces adjusting to changing participation formats. If the analysis of turns-at-talk reveals a neat turn-by-turn progressivity of the interaction, the analysis of sequences of actions in the move reveals how the latter might rely on various forms of embodied progressivity – most notably the step-by-step progressivity of the walk – and exploit them for the multimodal packaging of multiple orientations, multiactivity, and multiple participation frameworks.

The adjustment of walking and talking – both activities being organized in a stepwise projectable way – casts light on how actions are multimodally recipient-designed and adapted moment by moment (literally, step by step) to changing participation frameworks and interactional spaces. The manner of stepping and walking and their flexible transformations impinge upon the formatting of turns: hitches, self-repairs, perturbations, delaying the progressivity, as well as grammatical constructions used for their projective and expansive potential, appear to be closely coordinated with walking practices. Ways of walking and talking are also crucial for establishing the

conjunction of participants, for instance in openings (Mondada 2009), or their disjunction, for instance in closings (Broth and Mondada 2013), as well as for designing actions as private vs. public (Mondada in press). These flexible arrangements of bodies also reveal how the participants' categories – such as 'guide' and 'guided' – as well as their rights, obligations, responsibility, and authority, are visibly displayed and reflexively achieved through ways of initiating and responding to actions, positioning and repositioning, stepping forward and following, moving or resisting, walking ahead or backwards.

4. SENSING BODIES: SENSORIALITY

As previously demonstrated, the visual turn made possible by the use of video materials offers important potential for investigating embodied conduct and expanding the notions of temporality and sequentiality. However, it also presents some limitations, which are worth exploring. If video invites us to rethink a logocentric view of language and interaction, it also calls into question another form of reductionism, that of embodiment to audible-visible features. Audible features have been extensively explored in the form of linguistic forms (though non-linguistic sound features, including their perspectivization and spatialization, remain under-studied – see the comments for Excerpt 4, line 2, above); visible features have been exploited for analyses of embodied conduct conceived as mutually visibly accessible by the co-participants (and by the overseeing analyst). However, not all embodied conduct is reducible to visual cues. This notably concerns other, largely understudied, senses, such as touching, tasting and smelling. Sensoriality represents a new challenge for the analysis of social interaction for several reasons: on the one hand, because it tends to escape traditional video analyses; on the other hand, because it might appear to involve an individual sensing body (proprioceptivity) rather than socially interacting bodies.

Touch and tactility constitute one largely understudied form of sensoriality. Within interactional studies, it has been tackled mainly in the form of persons touching other persons, as in everyday hugs (Goodwin in press) and parental tactile control of children (Cekaite 2015), or in institutional settings, as in health-related activities (Nishizaka 2007; Nishizaka and Sunaga 2015). The case of participants touching objects has been even less studied (but see Kreplak and Mondémé 2014). The recent renewed focus on the study of the manipulation of objects (see Nevile et al. 2014) does not *per se* entail attention to tactility. Moreover, whereas touching other persons constitutes a form of embodied sensorial social interaction, in which others may respond in a tactile way, touching material objects is not interactionally organized *per se* and might appear to be mostly private.

In order to consider these issues, we briefly turn to a series of instances in which sellers in cheese shops touch their products (Mondada in prep.). The

following excerpts show that this material tactility goes beyond private sensations; it involves a form of embodied knowledge and professional expertise, as well as an intersubjective engagement with the customer.

We join the first example, Excerpt 5, in a cheese shop in Paris, where the seller is offering a series of possible options to a customer who has requested some dry goat cheese.

Excerpt 5

```
1         (0.6)
2  SEL    on a: euh*::
          we have: eh::
                    *...opens window--->
3         (0.3) * (0.4)        *(2.0)#
   sel        ->*LH reaches mzc*palpates--->
   fig                                 #fig5.1
4         ou*ais, *les mau#zacés vont être très *bien,
          yeah,    the Mauzacés will be   very  good,
          ->*points*,,,,,,,,,,,,,,,,,,,,,,,,,*gesticulates--->
   fig                  #fig5.2
5         i vont être un peu fermes aussi, un peu denses,
          they'll be a bit firm too, a bit dense,
6         sans être trop trop secs,
          without being too dry,
7         (0.5)
```

5.1

5.2

Prior to this excerpt, the seller has already mentioned a first possible cheese and described its characteristics. Now he moves to the second one, starting his utterance with a verb (*on a: euh::* 'we have: ehm::', line 2) projecting its argument, possibly a cheese name. However, what follows is not the projected noun, but a pause (line 3), during which the seller opens the window of the glass case where goat cheeses are refrigerated. While the ongoing turn is suspended,

his hand reaches a piece, and touches and palpates it (Figure 5.1). This palpation is visibly displayed for the customer, who remains behind the seller; it is also exhibited as the tactile evidence on the basis of which the progression of the seller's turn is produced. When the talk is resumed, the turn continuation does not provide the form of the cheese name, projected by the previous verb: instead the seller re-starts with *ouais* 'yeah' (line 4) – which is hearable as a response to the cheese palpation – and then continues, while pointing (Figure 5.2), with the description of that particular cheese (line 5). Touching is presented as the evidential basis for assessing its quality (softness vs. dryness); the way of touching manifests the embodied expertise of the cheese master.

Excerpt 6

```
1  SEL   pues pasame uno?
         so   pass me one?
2        (0.2) + (0.3)
   cli         +grasps cheese and leans it-->
3  CLI   huh huh
4        (1.2)
5  SEL   eh#::: +*piensa que eso *que #es muy lí*quido
         eh::: be aware that this that is very liquid
   cli       ->+
   sel          *takes it-------*palpates above/below*palpates laterally->
   fig    #fig6.1                  #fig6.2
6        (0.2) # (0.2)
   fig         #fig6.3
```


6.1


6.2


6.3

```
7  CLI   ah* okei
         oh okay
   sel   ->*palm up horizontal on top of cheese-->
8  SEL   y que:: (0.4) lo abres por aquí
         and that:: (0.4) you open it from here
9        (0.4)
10 CLI   pa+ra:,* (0.4)
         for:,    (0.4)
           +'opening' gesture--->>
   sel        ->*'opening' gesture--->
11 SEL   efectiva*mente no lo ab*ras entero* sabes,
         actually you don't open it entirely you know,
             --->*              *cutting gesture*
```

Whereas the customer does not always respond to the seller's touch and description, as in the previous fragment, in some cases s/he does, as in the next one, Excerpt 6 from a cheese shop in Madrid. Here, the 'liquidity' characterizing the selected piece of cheese is consequential for its future manipulation, and is mentioned by the seller in explaining how the client will have to cut it.

The client has earlier expressed his decision to buy that piece of cheese, a kind of *torta del casar*. After some explanations, the seller requests that he pass her one item (line 1), and he responds by grasping it and handing it to her (Figure 6.1). When receiving it (line 5), the seller relates information which is also a warning, mentioning a specific property of the cheese – its 'liquidity'. Her turn is slightly delayed by an initial 'eh:::' so that it starts only when she grasps the cheese: the imperative verb (*piensa* 'be aware', line 5) projects an assessment of the cheese which co-occurs with her palpating it. Her professional touching gesture consists first of palpating it with her fingers on the upper and the lower sides (Figure 6.2) and second with her fingers around the cheese, continuing the palpation during the following pause (Figure 6.3). The client does not respond immediately: he does not respond to the overt description of the 'liquidity', as indicated by the pause (line 6); he responds only a bit later, after *seeing* the seller visibly *palpating* the cheese during the pause. So in this case, the seller's activity of touching is oriented to and responded to – touching is visibly displayed and seen in an intersubjective way, producing an understanding of the quality of the product, consequential for the way it has to be handled (notice the fact that immediately afterwards the client responds to the seller by doing the gesture of opening the cheese, gesturally echoed and confirmed by her).

This *intersubjectivity of touch* is also noticeable in Excerpt 7, from an Italian cheese shop. The client requests a *caciotta*, indicating that this a regular purchase (line 1), occasioning several confirmations (lines 6–8). Finally, the seller grasps a piece of cheese from the fridge below him (lines 6–8), while the customer checks again if this is the right one (line 10). He does that by mentioning a sensorial quality (*morbidina*, using the diminutive of *morbida* 'soft') and by making, at the same time, a 'palpating gesture' in the air (Figure 7.1), with his right hand opening and closing. The seller confirms immediately, with a turn that both asserts and relativizes the softness of the cheese, while he visibly displays the palpation of the piece with both hands (Figure 7.2). After an extra confirmation the client decides to take it (line 13).

In this case, the tactile quality of the cheese is oriented to in a verbal description and in an embodied way by the client and by the seller. The former does not have access to the piece of cheese, but performs the 'palpating gesture' while uttering the adjective *morbidina*; the latter holds the *caciotta* in his hands and palpates it in a visible way for the customer, while describing and confirming its softness – even despite the coldness of the piece in the fridge. Both co-participants use a similar gesture, and recognize it as it is produced by the other.

Excerpt 7

```
1  CLI   m: io volevo mezza di quella *caciotta che c'ha:[::
         m: I wanted half of that caciotta that you have:[::
((5 lines omitted))
6        +(1.0)
   sel   +leans twd the product-->
7  CLI   *è [+questa?]+ ¡sì?¡*
         is [this one?]   yes?
8  SEL      [+eccola qua]
             [here it is]
   cli   *points-------------*
   sel                ¡nods¡
   sel     ->+grasps it+suspends his movement--->
9        (0.6)
10 CLI   *(che) fa# morbi+dina [(sì?)
         (which) doPRES3S ADJ-DIM [(yes?)
         (which) is very soft [(yes?)
11 SEL                         [sì. *adesso+ è fredda, però è #mor+bida.
                               [yes. now it is cold, but it is soft.
   cli   *palpating gesture---------*
                 ->+takes it to counter+palpates w two H-------+holds RH->
   fig            #fig7.1                              fig7.2#
12       quea s*olita che* le davo alla si[gnora.
         the usual one that I gave to your [wife
13 CLI                                     [ah *beni*ssimo +allora °(  )°
                                           [oh excellent so °(  )°
               *nods----*                      *nods*
   sel                                                 --->+puts on counter->>
```

7.1


7.2

In this activity, touching an object is not a private perception or sensation; rather, it is produced for self and for the other, visually displayed, recognized as such, achieving the intersubjectivity of touch and the alignment relative to the quality of the touched object, revealed by touching it. This intersubjectivity is achieved by exploiting the tactile properties and the visual display of the action of touching, as well as its verbal description, in a form of synesthetic communication. Moreover, the instances of touching observed here show how touch may constitute both a lay experience and an expert practice, a 'professional

touch' (see the 'professional vision' described by Goodwin 1994). Both meet in the shop encounter, resulting in subtle negotiations and confrontations of embodied access, perception and knowledge within an institutional context.

As shown in these last fragments, sensoriality is a crucial dimension of embodied action. Communicating bodies are also sensing bodies. Sensorial experiences and practices do not constitute a mere proprioceptive dimension, through which the individual agent establishes a private relation to the world, but a social, intersubjective dimension, through which a participant publicly and accountably mobilizes material objects in specific, skillful and knowledgeable ways, recognizable for the co-participants. Touching is not reducible to mere visual cues, and this raises methodological issues of how to document the haptic practices of the hand; but touching is also experienced in a way that is visibly displayed for the other and is actually noticed and seen by the other. In this perspective, sensing in interaction is always synesthetic, mobilizing an array of senses – within sensorial Gestalts constituted through words and other audible sonorities, and visible touching-tasting-smelling gesture. Sensing in interaction is organized by the co-participants as a collective engagement, conferring on sensorial experience an intersubjective dimension, which remains to be further studied.

5. CONCLUSION: CHALLENGES

In this article, I have sketched some challenges for the study of social interaction that originate from the use of video data, considering both their contribution and their limitations. The 'visual turn' has made possible a huge expansion of the field, opening up new avenues for the study of interacting bodies rather than only talking heads or communicative torsos. In the article, I have discussed some of these advances, which concern the study of the entire body instead of only language, mobility in interaction rather than only sedentary activities, and intersubjective sensorialities and their role in social practices rather than only verbal-visual cues. This widening of the field of social interaction does not merely represent the addition of new data, settings and activities to be studied. It also invites an adjustment of the existing conceptual frameworks for studying them. In particular, it presents some challenges for multimodal analysis, calling for a perspective that includes language as well as the body, and verbal-visual cues as well as other sensorial cues. This integrative and holistic conception of multimodality, rejecting verbal or visual reductionism, responds to the original invitation of CA and ethnomethodology to study the *accountability* of action (Garfinkel 1967; Sacks 1992): how action is accountably produced and how it is intersubjectively received and understood. It concerns the resources publicly mobilized by doers (body movements and postures, gesture, gaze, facial expressions, language) in order to implement their action, as well as the resources publicly oriented to by recipients (verbal, visual, and other sensorial cues).

This analysis depends heavily not only on the detailed scrutiny of the video materials at hand, but also of their detailed transcription, allowing a precise interpretation of the trajectories, temporalities and qualities of these multiple resources. In this respect, it is interesting to note that transcripts tend to be biased in favor of doers rather than recipients (for instance, the way an embodied action is looked at by the recipient is often transcribed and then ignored in the analyses). Multimodal transcripts raise again the question of their relevance, both for the co-participants and for the overhearing/seeing/sensing observer (analyst – but also cameraperson, see above), and thus of their relevant description (Sacks 1963).

This way of approaching accountability conceives of 'resources' for building social action as not only *used* by interactants, but more radically (re)*constituted* and (re)*negotiated* by them, being largely occasioned and transformed through their use. Consequently, they are situated, being indexical, depending on their sequential position, and on the praxeological context (i.e. the context of action) in which they emerge and are mobilized; in particular, their communicative relevance and intelligibility depend on the ecology of the ongoing action, which might foster the choice of some types of resources over others. Their indexicality does not exclude their methodical establishment and selection, as well as conventionalization and grammaticalization through repetitive use. Indexicality and systematicity constitute their Janus faces, requiring both an analysis of the specificity of their occasioned instances and an analysis of the way they are methodically mobilized and interpreted within specific sequential environments.

In order to make sense of and identify the relevance of multimodal cues, co-participants as well as analysts must rely on the way they are assembled in an orderly way, and their composition as well as their position. Sequentiality remains the fundamental principle on which the organization of human action is based. In this respect, too, multimodality presents us with conceptual and analytical challenges. For instance, the 'next-turn proof procedure' (Sacks, Schegloff and Jefferson 1974), which allows CA analysts to demonstrate how a turn and its compositional details have been understood and locally treated by the participants, is more generally a 'next-action proof procedure', allowing the analyst to show how participants respond to any multimodal Gestalt. As we observed in the empirical analyses in this paper, this means not only considering responsive verbal actions in the next turn, but also embodied responsive actions that might happen within and along the previous one. Multimodally formatted actions are characterized by multiple simultaneous temporalities, rather than strict successivity. Nonetheless, the principle of sequentiality still holds: participants fundamentally orient to emergent actions and their sequential positioning, to prior and next, to initiating and responding actions. This shows how the challenges of multimodality require a careful analytical focus (and consequently a detailed and precise temporal annotation in transcripts) on complex sequential arrangements of multiple resources, as

well as, more generally, a sophisticated conception of time – conceived both as *chronos* (referring to the precise timing of the multimodal assemblage of resources) and *kairos* (referring to their occasioned, flexible, reflexive production and interpretation; Heidegger 1996 [1927]).

This has consequences not only for the study of language and interaction but also more globally for the social and the cognitive sciences. There is a general interest in language and the body across disciplines; the careful study of the finely tuned coordination of talk and body movements in interaction casts light on how utterances are produced bodily and received in real time, in a situated way; how bodies mutually adjust together in a variety of practices, like walking, looking around, sensing objects, eating together; how they explore the world and occupy space in a collective intersubjective way. In turn, these situated embodied practices are the *locus* in which broader issues relative to language, body, cognition, action, culture, knowledge, social relations and identities, spatiality and temporality are locally shaped, implemented and transformed, and, as such, made observable for the analyst.

NOTE

1. This paper has been supported by my project 'Multimodality: Reconsidering Language and Action through Embodiment' within my Finland Distinguished Professorship funded by the Academy of Finland (SA grant no 282931), as well as by Swiss National Science Foundation Project 'Speaking in Public' (SNF grant no. 100014_144376 and 100012_162689/1).

REFERENCES

Birdwhistell, Ray. 1970. *Kinesics and Context*. Philadelphia, Pennsylvania: University of Pennsylvania Press.

Broth, Mathias, Eric Laurier and Lorenza Mondada (eds.). 2014. *Studies of Video Practices: Video at Work*. London: Routledge.

Broth, Mathias and Lorenza Mondada. 2013. Walking away: The embodied achievement of activity closings in mobile interaction. *Journal of Pragmatics* 47: 41–58.

Cekaite, Asta. 2015. The coordination of talk and touch in adults' directives to children: Touch and social control. *Research on Language and Social Interaction* 48: 152–175.

Clark, Andy. 1997. *Being There: Putting Brain, Body and World Together Again*. Boston, Massachusetts: MIT Press.

Couper-Kuhlen, Elizabeth and Margret Selting (eds.). 1996. *Prosody in Conversation: Interactional Studies*. Cambridge, U.K.: Cambridge University Press.

Cresswell, Tim and Peter Merriman. 2011. *Geographies of Mobilities, Practices, Spaces, Subjects*. London: Ashgate.

Crossley, Nick. 2001. *The Social Body: Habit, Identity and Desire*. London: Sage.

Deppermann, Arnulf (ed.). 2013. Conversation Analytic Studies of Multimodal Interaction. Special issue of *Journal of Pragmatics* 46(1).

Ford, Cecilia E., Barbara A. Fox and Sandra A. Thompson (eds.). 2002. *The Language of Turn and Sequence*. Oxford, U.K.: Oxford University Press.

Garfinkel, Harold. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, New Jersey: Prentice-Hall.

Goffman, Erving. 1971. *Relations in Public: Microstudies of the Public Order*. Harmondsworth, U.K.: Penguin.

Goffman, Erving. 1981. *Forms of Talk*. Philadelphia, Pennsylvania: University of Pennsylvania Press.

Goodwin, Charles. 1979. The interactive construction of a sentence in natural conversation. In George Psathas (ed.) *Everyday Language: Studies in Ethnomethodology*. New York: Irvington. 97–121.

Goodwin, Charles. 1981. *Conversational Organization: Interaction between Speakers and Hearers*. New York: Academic Press.

Goodwin, Charles. 1994. Professional vision. *American Anthropologist* 96: 606–633.

Goodwin, Charles. 2000. Action and embodiment within situated human interaction. *Journal of Pragmatics* 32: 1489–1522.

Goodwin, Charles. 2003. Pointing as situated practice. In Sotaro Kita (ed.) *Pointing: Where Language, Culture and Cognition Meet*. Hillsdale, New Jersey: Erlbaum. 217–241.

Goodwin, Charles and Marjorie Harness Goodwin. 2004. Participation. In Alessandro Duranti (ed.) *A Companion to Linguistic Anthropology*. Oxford, U.K.: Blackwell. 222–244.

Goodwin, Marjorie Harness. In press. Haptic sociality: The embodied interactive constitution of intimacy through touch. In Christian Meyer, Jürgen Streeck and J. Scott Jordan (eds.) *Intercorporeality: Emerging Socialities in Interaction*. Oxford, U.K.: Oxford University Press.

Haddington, Pentti, Tiina Keisanen, Lorenza Mondada and Maurice Nevile (eds.). 2014. *Multiactivity in Social Interaction: Beyond Multitasking*. Amsterdam, The Netherlands: Benjamins.

Haddington, Pentti, Lorenza Mondada and Maurice Nevile (eds.). 2013. *Interaction and Mobility: Language and the Body in Motion*. Berlin, Germany: De Gruyter.

Hakulinen, Auli and Margret Selting (eds.). 2005. *Syntax and Lexis in Conversation: Studies on the Use of Linguistic Resources in Talk-in-Interaction*. Amsterdam, The Netherlands: Benjamins.

Heath, Christian. 1986. *Body Movement and Speech in Medical Interaction*. Cambridge, U.K.: Cambridge University Press.

Heath, Christian and Paul Luff. 2000. *Technology in Action*. Cambridge, U.K.: Cambridge University Press.

Heath, Christian, Jon Hindmarsh and Paul Luff. 2010. *Video in Qualitative Research*. London: Sage.

Heidegger, Martin. 1996 [1927]. *Being and Time*. New York: SUNY.

Hindmarsh, Jon and Christian Heath. 2000. Embodied reference: A study of deixis in workplace interaction. *Journal of Pragmatics* 32: 1855–1878.

Ivarsson, Jonas and Christian Greiffenhagen. 2015. The organization of turn-taking in pool skate sessions. *Research on Language and Social Interaction* 48: 406–429.

Kendon, Adam. 1990. *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge, U.K.: Cambridge University Press.

Komter, Martha. 2006. From talk to text: The interactional construction of a police record. *Research on Language and Social Interaction* 39: 201–228.

Kreplak, Yaël and Chloé Mondémé. 2014. Artworks as touchable objects: Guiding perception in a museum tour for blind people. In Maurice Nevile, Pentti Haddington, Trine Heinemann

and Mirka Rauniomaa (eds.) *Interacting with Objects: Language, Materiality, and Social Activity*. Amsterdam, The Netherlands: Benjamins. 289–311.
Kress, Gunther and Theo van Leeuwen. 2001. *Multimodal Discourse*. London: Oxford University Press.
Leeds-Hurwitz, Wendy. 1987. The social history of *The Natural History of an Interview*: A multidisciplinary investigation of social communication. *Research on Language and Social Interaction* 20: 1–51.
McIlvenny, Paul, Mathias Broth and Pentti Haddington (eds.). 2009. Communicating Place, Space and Mobility. Special issue of *Journal of Pragmatics* 41(10).
McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago, Illinois: University of Chicago Press.
McQuown, Norman A. (ed.). 1971. *The Natural History of an Interview*. Chicago, Illinois: University of Chicago.
Mondada, Lorenza. No date. Multimodal transcription conventions. Last accessed at https://franz.unibas.ch/fileadmin/franz/user_upload/redaktion/Mondada_conv_multimodality.pdf
Mondada, Lorenza. 2006. Video recording as the reflexive preservation and configuration of phenomenal features for analysis. In Hubert Knoblauch, Jürgen Raab, Hans-Georg Soeffner and Bernt Schnettler (eds.) *Video Analysis: Methodology and Methods*. Bern, Switzerland: Lang. 51–68.
Mondada, Lorenza. 2007. Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies* 9: 195–226.
Mondada, Lorenza. 2009. Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. *Journal of Pragmatics* 41: 1977–1997.
Mondada, Lorenza. 2013. Video as a tool in sociology and anthropology. In Cornelia Müller, Alan Cienki, Ellen Fricke and David McNeill (eds.) *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*. Berlin, Germany: de Gruyter. 978–988.
Mondada, Lorenza. 2014a. The local constitution of multimodal resources for social interaction. *Journal of Pragmatics* 65: 137–156.
Mondada, Lorenza. 2014b. Pointing, talk and the bodies: Reference and joint attention as embodied interactional achievements. In Mandana Seyfeddinipur and Marianne Gullberg (eds.) *From Gesture in Conversation to Visible Utterance in Action*. Amsterdam, The Netherlands: Benjamins. 95–124.
Mondada, Lorenza. 2014c. The temporal orders of multiactivity. In Pentti Haddington, Tiina Keisanen, Lorenza Mondada and Maurice Nevile (eds.) *Multiactivity in Social Interaction: Beyond Multitasking*. Amsterdam, The Netherlands: Benjamins. 33–75.
Mondada, Lorenza. 2014d. Bodies in action: Multimodal analysis of walking and talking. *Language and Dialogue* 4: 357–403.
Mondada, Lorenza. In press. Questions in the move. In Arnulf Deppermann and Jürgen Streeck (eds.) *Temporalities and Modalities*. Amsterdam, The Netherlands: Benjamins.
Mondada, Lorenza. In prep. The intersubjectivity of touch.
Mondada, Lorenza and Kimmo Svinhufvud. 2016. Writing-in-interaction: Studying writing as a multimodal phenomenon in social interaction. *Language and Dialogue* 6: 1–53.
Nevile, Maurice. 2015. The embodied turn in research on language and social interaction. *Research on Language and Social Interaction* 48: 121–151.
Nevile, Maurice, Pentti Haddington, Trine Heinemann and Mirka Rauniomaa (eds.). 2014. *Interacting with Objects*. Amsterdam, The Netherlands: Benjamins.

Nishizaka, Aug. 2007. Hand touching hand: Referential practice at a Japanese midwife house. *Human Studies* 30: 199–217.
Nishizaka, Aug and Masafumi Sunaga. 2015. Conversing while massaging: Multidimensional asymmetries of multiple activities in interaction. *Research on Language and Social Interaction* 48: 200–229.
Ochs, Elinor. 1979. Transcription as theory. In Elinor Ochs and Bambi B. Schiefflin (eds.) *Developmental Pragmatics*. New York: Academic Press. 43–72.
Ochs, Elinor, Emanuel A. Schegloff and Sandra A. Thompson (eds.). 1996. *Grammar and Interaction*. Cambridge, U.K.: Cambridge University Press.
Ryave, A. Lincoln and James N. Schenkein. 1974. Notes on the art of walking. In Roy Turner (ed.) *Ethnomethodology*. Harmondsworth, U.K.: Penguin. 265–274.
Sacks, Harvey. 1963. Sociological description. *Berkeley Journal of Sociology* 8: 1–16.
Sacks, Harvey. 1992. *Lectures on Conversation [1964–72]* (Vols. I and II). Oxford, U.K.: Blackwell.
Sacks, Harvey and Emanuel A. Schegloff. 2002. Home position. *Gesture* 2: 133–146.
Sacks, Harvey, Emanuel A. Schegloff and Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50: 696–735.
Scheflen, Albert E. 1972. *Body Language and Social Order: Communications as Behavioral Control*. Englewood Cliffs, New Jersey: Prentice Hall.
Schegloff, Emanuel A. 1984. On some gestures' relation to talk. In Maxwell Atkinson and John Heritage (eds.) *Structures of Social Action*. Cambridge, U.K.: Cambridge University Press. 266–296.
Schegloff, Emanuel A. 1998. Body torque. *Social Research* 65: 535–586.
Schegloff, Emanuel A. 2007. *Sequence Organization in Interaction: A Primer in Conversation Analysis* (Vol. 1). Cambridge, U.K.: Cambridge University Press.
Shilling, Chris. 2003. *The Body and Social Theory*. London: Sage.
Sidnell, Jack and Tanya Stivers (eds.). 2012. *The Handbook of Conversation Analysis*. Oxford, U.K.: Wiley-Blackwell.
Streeck, Jürgen, Charles Goodwin and Curtis LeBaron (eds.). 2011. *Embodied Interaction, Language and Body in the Material World*. Cambridge, U.K.: Cambridge University Press.
Thrift, Nigel. 2007. *Non-Representational Theory: Space, Politics, Affect*. London: Routledge.
Urry, John. 2007. *Mobilities*. London: Polity.

Address correspondence to:

Lorenza Mondada
Department of Linguistics and Literature
University of Basel
Maiengasse 51
Basel 4056
Switzerland

lorenza.mondada@unibas.ch

APPENDIX: Transcription conventions

Talk has been transcribed according to conventions developed by Gail Jefferson (see 2004 for an extensive presentation). An indicative translation is provided line by line, in italics. The aim is to provide help with reading the original, rather than an idiomatic rendering of the language used. Multimodal details have been transcribed according to the conventions developed by Mondada (n.d.).

**	Gestures and descriptions of embodied actions are delimited between
++	two identical symbols (one symbol per participant)
ΔΔ	and are synchronized with correspondent stretches of talk.
*--->	The action described continues across subsequent lines
---->*	until the same symbol is reached.
>>	The action described begins before the excerpt's beginning.
--->>	The action described continues after the excerpt's end.
.....	Action's preparation.
----	Action's apex is reached and maintained.
,,,,,	Action's retraction.
luc	Participant doing the embodied action is identified when (s)he is not the speaker.
fig	The exact moment at which a screen shot has been taken is indicated
#	with a # sign showing its position within turn at talk.