Lorenza Mondada, University of Basel

This paper deals with a significant advance the use of video and the study of multimodality within conversation analysis have made possible: the possibility to analyse in detail the sequentiality of actions that are achieved by other resources than talk, and more precisely a diversity of embodied practices. A close attention to how we transcribe these silent embodied actions enables us to better understand their specific temporal unfolding, spatial arrangements, and sequential organization. The paper starts discussing silent second actions (responses to requests); then moves to silent first actions (requests achieved by handing over objects); and finally discusses embodied sequences fully realized in silence. The proposed transcript notation enables reflection upon the complex emergent and sequentially unfolding temporality of multimodally formatted actions. Taking into account the details of embodied conduct, it discusses the consequences for the principled notions of temporality, sequentiality and multimodality.

1. Introduction

The use of video has considerably increased opportunities for studying human action and, more particularly, for re-considering the role of the body in the organization of social interaction. While language has traditionally been considered the main resource for communicating, the use of video recordings and the exploitation of video data has made it possible to fully develop a multimodal conceptualization of action, communication and social interaction, and has shown that reducing these things to language alone is a theoretical and methodological artifact (Goodwin, 2017).

The notion of *multimodality* refers to the resources that participants of social interaction mobilize to make their actions publicly intelligible: these resources include language in its various aspects (prosody, syntax, lexis, etc.), as well as body conduct such as gestures, gazes, facial expressions, body postures, movements, and so on (Deppermann, 2013; Goodwin, 2017; Nevile, 2015; Streeck, Goodwin & LeBaron, 2011). This conception of multimodality treats the diversity of resources in a unified and holistic way: it includes more conventional resources (language, and some gestures) and more improvised and contingent resources (including vocalizations, movements and object manipulations) which largely depend on the specificities of the context (Mondada, 2014a); it also considers how these resources are combined and arranged together in complex multimodal Gestalts, although they are characterized by distinct temporalities, trajectories and projections (Mondada, 2016).

This approach to multimodality in social interaction raises several innovative questions for approaches that have long been based on audio recordings and/or focused on verbal resources. For instance, the focus on talk has allowed conversation analysis to be developed in a rigorous way, showing the systematic and methodic organization of turns, actions, and sequences (Sacks, Schegloff, Jefferson, 1974; Schegloff, 2007) – although the main objective has been, from the beginning, not to study language *per se*, but social action (Sacks, 1984), and some early studies show an interest in the body (Sacks & Schegloff, 2002). Nowadays, the availability of video data makes it possible to enlarge and elaborate on fundamental notions of conversation analysis, such

as temporality, sequentiality and indexicality (Mondada, 2016; Heath 2013; Streeck et al., 2011). Several questions are raised in this respect: how are turns at talk and speakership to be understood when considering embodiment? Where/when does a turn begin? Where/when does an action begin? How is an emergent action made recognizable and accountable for the participants, and when? What counts as a response, and how is responsiveness identified by co-participants? How are embodied and verbal resources arranged together, within distinct but interrelated temporalities, which at the same time are holistically recognizable as contributing to the same action? How are embodied and verbal resources distributed within a situated activity, when moments of talk alternate with moments of silent action? How should the sequential organization of silent actions be accounted?

This paper aims to tackle some of these questions. Two methodological prerequisites are needed in order to do so. One the one hand, video recordings are indispensable (Heath, Hindmarsh, Luff, 2010; Mondada, 2006). Recordings have improved, thanks to technological innovations, although technology alone is not enough to define the quality of video data, which ultimately depends on the interactional relevance of the shootings. On the other hand, detailed transcriptions of video recordings are also needed, since subtle details, fine-tuned temporalities and micro-sequential phenomena can neither be imagined or be directly observed with a naked eye. Transcripts are part of the professional vision (Goodwin, 1994) of the video analyst: they allow her to identify and order details with the necessary temporal precision and to discover the locally relevant features of action, making sequential analysis possible.

In this paper, I discuss some of the challenges for sequential video analysis on the basis of multimodal transcripts, by adopting the conventions I have developed over the years (for a conceptual discussion, see Mondada, 2018a). The analysis focuses on a particular challenge for transcription –embodied actions achieved without saying any words– and shows possible variations in transcribing them, and their analytical payoffs. In so doing, the paper focuses on actions that have the particularity of being silently embodied. While sequence organization has been largely systematized on the basis of turns-at-talk (Schegloff, 2007), silent embodied actions have been less considered in this respect. They constitute a challenge to a linear conception of sequentiality: albeit being formatted by multiple simultaneous embodied movements, they rely on sequential orders and projections. They constitute an exemplary phenomenon for further exploring sequentiality as a fundamental principle organizing human action, not bound to talk, but relevant for body actions without talk too.

Within Conversation Analyis, silence has been categorized in different ways, as the absence of talk at different sequential locations with respect to the organization of turns (Hoey, 2017). An intra-turn silence occurring in the middle of a turn, before its completion and before a transition-relevance point TRP is reached, is a *pause*; an inter-turn silence that occurs within a TRP is a *gap*; and an extended silence after a TRP is a *lapse*. This terminology is not straightforward when considering actions that are silently achieved: for example, when considering embodied compliance to a directive, it is problematic to treat the second pair part as occurring within a gap or a lapse. This shows that silence is not always the absence of talk. Silence invites us to consider the temporality of embodied actions more closely.

If we turn to other approaches within ethnomethodological and conversation analytic, video studies have shown the importance of actions that are bodily formatted without saying a word, within very diverse contexts. Workplace settings have invited scholars to take not only talk into consideration, but embodied actions central to the ongoing work (such as passing instruments in surgery, (Heath et al., 2018)). In ordinary interactions too, the study of silent embodied actions has allowed scholars to recognize the methodic organization of activities without a word (such as skate board sessions, (Ivarsson & Greiffenhagen, 2015) or touching and tasting, Mondada (2018b)). It has also made it possible to show the active participation of persons other than speakers in an interaction – for instance by demonstrating how responsiveness, and more generally agency, characterize silent parties while speakers produce their turns-at-talk, such as in the form of gazing behavior, nods and other gestures, etc., (Goodwin, 1979; Goodwin, 1980)– and even of non-human participants, like animals (see Mondémé (2018) on interspecific interactions with dogs, see Mondada (2018a) for a discussion of a transcript of interactions among baboons). The role of silent embodied actions has also been highlighted in relation to multiactivity (Haddington et al., 2014), that is, the possibility that people engage in more than one activity at a time, which is often organized on the basis of a complementary distribution of multimodal resources in parallel actions. More specifically, silent responses within sequences have been recognized, especially for requests and directives (see Mondada, 2014b, Goodwin & Cekaite, 2014, Lindwall & Ekström, 2012, etc.). This paper draws on this literature to reflect on silent embodied actions in more general terms, constituting either the first or the second pair part within a sequence, or both.

In order to be explored in a systematic way, these configurations need to be analyzed on the basis of a detailed multimodal transcription – considered here the fundamental empirical tool with which to document and conceptualize the temporality, projectability, sequentiality, and accountability of action in social interaction. In particular, the analyses in this paper show the importance of the transcription of time to the analysis of sequentiality. Various attempts to represent embodied actions are reported in the literature, with various degrees of precision and granularity concerning the relationship between one action and another, and thus the temporality of one set of resources implementing that action in relation to another set. Some transcripts choose to rely only, or mainly, on images for annotating embodiment (Heath et al., 2018, Laurier, 2014, Goodwin 2017), others choose to rely on forms of representation of segments of time in the form of dashes representing tens of seconds (Luff & Heath, 2015); or in the form of numeric indications of segments (Mondada, 2018a).Other transcripts are much looser as far as the temporality of embodied actions is concerned (often adapting verbal conventions for ((comments)) or [overlaps] for transcribing embodied conducts). In this paper I reflect on the consequences of transcription choices for the conception of silent embodied actions on the basis of my conventions, which are compatible with other notations of time and with existing annotation software.

2. Data

The data examined in this paper are all involves shop encounters, in which a customer interacts with a salesperson. This institutional setting is characterized by routine sequences of actions that can be variously multimodally implemented. Their apparent simplicity enables a systematic exploration of what makes them possible, accountable, and unproblematically intelligible for the participants.

The corpus comprises a series of video recordings in three kinds of shops –bakeries, cheese shops and kiosks (convenience stores)– which are characterized by the fact that most of the products are requested at the counter rather than being self-collected, as in supermarkets (De Stefani, 2011), thus engendering sequences of requests (Fox, 2015; Fox & Heinemann, 2015; Mondada & Sorjonen, 2016). The recordings considered here were made in French, Italian and German speaking countries (Switzerland and France, Germany and Italy).

In all cases, the shop owners were active partners collaborating with the study; salespersons and customers were informed about the video recordings and the aims of the project and gave their written consent for the use of data in videos and transcripts for research purposes.

The fragments considered in this paper all deal with requests for products that orient to an immediate response (vs. compliance in the future), and that imply the manipulation of material objects. Nonetheless, the analysis is less interested in the format of the requests than in the forms of bodily coordination between the participants, and in situated practices of giving and taking (Heath et al., 2018; Mondada & Sorjonen, forthcoming), which are often achieved in silent ways.

More particularly, the analysis explores the multimodal arrangements characterizing silent embodied actions within sequences in which either the second (Section 3) or the first action (Section 4) is achieved in an embodied and silent way, without any word; it also discusses sequences in which both actions are silently implemented (Section 5).

3. Silently fetching an object in response to a request

The possibility to respond to requests with an embodied action has been identified and discussed in the literature (Mondada, 2014b; 2017; Goodwin & Cekaite, 2014; Rossi, 2014). Typically, when a request targets an action to be done immediately, the response is constituted by the embodied action granting it. Thus, in this case, the second pair part is an embodied silent action. This kind of response shows interesting variations relative to their temporality: they can be performed early or late, depending on multiple contingencies (Mondada, 2017). This section considers how to represent embodied responses to requests in transcripts adopting different granularities: it shows how these variations affect the possible analysis of the temporality of responses, and thus the very conception of what counts as a response.

The first, simplest occurrence of the phenomenon is provided here, which will serve as a baseline for the analysis and for the illustration of the multimodal conventions used. At a cheese stand in a market in Italy, a customer requests some mozzarella:

(1a) (FRO_I_PAD_mat-2-08-54_bufala)

```
1   CUS    dam+mi+   #una    + mo+zzarella %di% bufala:
            give me   one        bufala mozzarella
    cus        +..+thumb ptg+,,,+
    sal                                    %nods%
    fig            #fig.1
2          *    (1.8)  * (0.2) # (0.6)
    sal    *goes to mz*brings back mz-->
    fig                    #fig.2
3   CUS    ce la mangiamo a pranzo* dai
            we will eat it for lunch PRT
    sal                          ->*
```


1

The customer's request (1) is made while pointing with the thumb towards the area in the refrigerated case where this product is stored (Figure 1). In response, when the generic name of the product is uttered, the salesperson first nods; and secondly, as soon as the request is completed, he walks towards the mozzarella and brings one back to the customer (2-3, Figure 2), while the latter produces an account about when it will be eaten.

The multimodal transcription of this simple sequence shows the precise timing of the thumb pointing gesture which co-occurs with "+una+"/"one" (1, bracketed by the two same symbols, here +), is prepared slightly before (indicated by +…+) and retracts on the first syllable of the noun (+,,,+). In a similar way, the seller's nod co-occurs with "di" and is delimited by the symbol %. These conventions show how embodied conducts are temporally located in relation to the ongoing turn, how the turn is multimodally formatted by the speaker, and also how it is silently acknowledged by the co-participant. Moreover, the same convention enables an annotation for how the silent response to the request is produced: the silence that follows (3) is segmented in temporal fragments corresponding to the salesperson's walk towards the mozzarella "*(1.8)*", then to the time it takes to bring it back – which begins in the last part of the silence, "*(0.8)", but continues in the next line (this continuation is indicated by an arrow at the end of the notation "brings back mz-->" which is an instruction to look for the final arrow ending with the same symbol

"-->*", here on line 3). The description of action, shortly indicated in the notation, is complemented by the images, which are also precisely located within the action (by the symbol #): while the textual description is necessary selective and limited, analytically distinctive for each resource, the visual representation of an instant within the flow of action synthetically demonstrates the coordination of all the resources (see below for a more in-depth comment on the figures).

The convention is based on the fundamental principle that all embodied conducts are precisely located in time and are coordinated together: they are generally not isochronous –in the sense that their temporalities most often do not coincide– but they are finely-tuned and adjusted one to another. Their precise location enables an analysis of how responsive they are one to another, and therefore the identification of which participant, with which resource, has initiated the action.

Transcripts generally vary concerning their granularity, depending on how detailed the analysis. In this case, the transcript could also be enhanced in the following way:

(1b) (More detailed transcript of 1a)

```
1  CUS    *dam+mi+ una+ mo+zzarella* %di% bu•fa•±la:*
a  cus        +..+thumb ptg+,,,+
b  sal    *leans twd CUS-----------*,,,,,,,,,,,,,,,,*
c  sal                              %nods%
d  sal    >>looks at CUS--------------------•,,•looks at mz->
e  cus    >>looks at SAL----------------------±gz fllws SEL->
2       *    (0.9)    *          (0.9)      *(0.5)• (0.3)
a  sal    *walks to mz*leans over fridge*brings back mz->
b  sal                                     ->•looks at CUS->
3  CUS    ce la mangiamo a pra•nzo* da±i
a  sal                  -->•
b  sal                ->*
c  cus                          ->±
```

In this version, the gaze of both participants is added. This shows how the customer looks first at the salesperson (1e) –who leans over towards her (1b, Figure 1), displaying attention and availability– while uttering the request, and how her gaze then follows him until he comes back (1e-3c, Figure 2). This enables a demonstration of the fact that the seller sees the customer's gesture and that the customer recognizes the movement of the seller as compliance with her request. The annotation of gaze is thus crucial for the documentation of the participants' understanding and orientation to embodied actions. This raises the question of when a responsive action begins: as we will further observe in the next fragments, responses emerge progressively step by step, for example, first with a shift of gaze, then a preparatory gesture or movement, displaying an understanding of, and alignment to, what is going on, and finally by an embodied action complying with the initial request. Depending on the (analyst's as well as participant's) issues at stake, the preparations and projections of the action are more or less crucial for a global appreciation of what is going on.

The detailed transcript and the choice of granularity is therefore consequential for demonstrating both the ongoing action and the participants' orientation to it. Textual *vs.* visual representations of these details produce different affordances. Whereas images in transcripts (clearly showing the coordination of the seller's walk, hand grasping, and gaze on the target, and the corresponding

following gaze of the customer) are crucial for offering a *synthetic* visual representation of all resources at a precise instant, textual annotations (clearly showing the exact moment of the beginning/end of a movement) instead show the *analytical* relationships between distinct but interrelated timings of multimodal conducts.

In order to demonstrate the recurrent pattern of an action followed by a silent embodied response and some possible variation in the temporality of the sequence, the following fragment shows another sequence, where all the actions are performed earlier and faster. In a kiosk in France, a customer requests a cigarette box located on the shelf behind the salesperson:

(2a) (KIO_F_VIL-1-23-48 malb light)

```
1  CUS      bonjou[r +madame. ±
           good mor[ning m'am
2  SAL          [bonjou:r
                [good morning
   cus      >>gazes at SAL----±gaze at money->
   cus            +extends RH w money--->
3  CUS      un malboro* ligh#t s'il v+ous        + pl±aît
           a   malboro  light   please
                              --->+money on c+
                                           ->±gz cig->
   sal               *turns to shelf---->
   fig                   #fig.1a/1b
```


1a=1b

```
4        (0.2) *(0.1)#(0.2)*+(0.2)+(0.1)+* (0.2)  *(0.1)+± (0.2)± (0.4)*#
   sel       ->*grasps cig*takes cig*..,,*reads bc*puts cig on c-------*
   cus                    +ext LH+,,,,,+              +ext LH->
   cus                                           -->±gz SAL±gz cig->
   fig                #fig.2                                   fig.3#
```

3

The encounter begins with an exchange of greetings (1-2), while the customer is already putting some money on the counter, projecting payment (the gesture emerges on "+madame" and is completed on "+vous+"). She only then moves to the reason for the visit, requesting some cigarettes (3). This constitutes the first pair part of the sequence. The salesperson promptly grants the request, within a second pair part that is bodily implemented: she turns to the shelf to fetch the cigarettes. The granting response is initiated by the pivoting movement of the salesperson turning to the shelf as soon as the brand is uttered (Figure 1a/1b) –without waiting for the specification of the kind of cigarette targeted.

In this way, the second pair part emerges as the first is still ongoing. The movement of the seller's hand is incrementally responsive to the turn progression, and orients more and more precisely towards the relevant target, thanks to continuous micro-adjustments. The progressivity of both actions is organized in the form of two successive, sequentially organized simultaneities (Mondada, 2018a).

Compliance is completed by picking up the cigarettes (Figure 2), moving them towards the bar-code reader and finally placing them on the counter (Figure 3a/3b). This second paired action – fetching the cigarettes- is constituted by a diversely articulated gesture. The segmentation in the transcript (4) does not merely reflect different trajectories (the left hand moves towards the shelf, comes back from the shelf with the cigarettes, which are then passed to the right hand towards the bar-code reader, and finally moves to the counter) constituting the fetching of the cigarettes,

in response to the request. These trajectories are oriented by the customer herself: she gazes back at the seller's hand when it is almost on the cigarettes, and monitors its grasping. This gaze shows the participant's orientation towards the relevance of the moment in which the cigarettes are picked up (which is also a way to orient to the prior movement as still possibly repairable). Next, the customer orients to the gesture of taking back the cigarettes in another way, by beginning to extend her left hand towards them. When the seller begins to move the packet towards the bar code (annotated as "`*takes cig*..,,*reads bc*`") the customer retracts her gesture: this shows an orientation to the former trajectory as possibly directed towards the counter, and an adjustment to the change of trajectory towards the bar-code as projecting another, preliminary gesture, before the packet is moved back to the counter. These responsive movements ground the segmentation of the embodied actions in the transcript.

The seller's granting action is bodily initiated in the same position as where the previous seller (extract 1) was nodding: she displays her understanding and acknowledgment of the request with the early initiation of her response. A closer look at the video, however, shows another, even earlier, form of response – constituting an instance of another micro-adjustment that can be transcribed in the following way:

(2b) (detail)
```
1   CUS     bonjou[r +ma*dame.
            good mor[ning m'am
2   SAL            [bonjou:r
                   [good morning
a   cus              +extends RH--->
b   sal                 *extends RH twd CUS->
3   CUS     un malb*oro* light s'il v+ous+ plaît
            a   malboro  light please
a                                 --->+puts money on c+
b   sal           ->*,,,*turns to shelf---->
```

The extension of the customer's hand (2a) is first responded to by the seller extending her hand in a reciprocal way (2b). This movement shows an orientation towards the client as *giving* something, immediately responding with an incipient *taking* gesture. These paired actions are frequently relevant in this setting when customers hold a coupon to be checked (see below, extract 4), however, in this case, the client's movement is *not* an action requesting to check a lottery ticket, but an early anticipation of the action of paying, which will become visible later on, when the money is put on the counter (that is, when what the customer hands over becomes visible as a bill and some coins, which also make some sound) (3a). The description of the action in the transcript accompanies this emergent character, without anticipating what becomes intelligible only later (in line 2a the object in the extended hand is not specified, contrary to line 3a). The action of requesting (and the gesture of the customer) becomes clear when the client's turn is emerging (3): as soon as the name of the brand is recognizable, the salesperson stops her initial movement, retracts it (*,,,* 3b) and instead turns to the shelf.

The granularity of the transcription enables the documentation of an instance of embodied repair –with an initial response to a giving hand being transformed into a response to a verbal request. This embodied repair makes visible the ongoing seller's understanding –and its transformations– of the kind of first action that is being emergently achieved.

The second pair part is completed when the cigarettes land on the counter, however, in this case too, the trajectory is responded to with some discontinuities by the customer, displaying her progressive (re)interpretations of the final trajectory of the gesture. The closing of this sequence is made more complex by the fact that behind the customer (CUS), another client (CLI, in grey), was still musing over her purchases, and picks up her cigarettes and leaves at the same time:

(2c) (continuation)

```
4           * (0.2)£ *(0.1)+± (0.2)±+    (0.2)    + (0.2)* #
    sel     *reads bc*puts cig on c---------------------*
    cus                     +ext LH--+palm open up+grasping H->
    cus                  -->±gz SAL±gz cig->
    cli            £extends grasping H---->
    fig                                              fig.4#
```

4=detail

```
5   CUS     [mer+c£*i#::
6   CLI     [mer+c£*i#::
    cus       ->+takes cig and exits->>
    cli          £takes cig and exits->>
    sal            *takes the money->
    fig              #fig.5
7   SAL     au[r'voir merci
8   CLI       [>au r'voir.<
```

5=detail

The customer (CUS) orients to the approaching seller's hand: she not only extends her own arm, but opens her hand, palm up, in a 'receiving' gesture (Figure 4). The seller's trajectory does not orient to the customer's hand, however, rather it ends by placing the cigarettes on the counter (vs. in the customer's hand). The customer readjusts to this trajectory, by changing the position of her hand, now taking the form of a 'grasping' gesture (Figure 5). Interestingly, the client nearby (CLI), also extends her hand at the same moment, which from the beginning takes the form of a grasping gesture (on the left side of Figures 4-5).

Both customers thank the salesperson (5-6): it is significant that the customer (5) thanks her after she has changed her hand posture. Both take their cigarettes, and leave at the same time. Moreover, at the same time, the salesperson takes the money as soon as she has put the cigarettes on the counter (Figure 4). In this way she also closes the paying sequence that was initiated by the customer even before her request. It is also remarkable that the price of the cigarettes is never announced: the customer puts the exact amount on the counter and this is seen by the salesperson, who just picks it up after the cigarettes have been given and taken.

Multimodal transcripts of this sort can be realized manually, but can also be computer supported. There are various solutions for the aligned annotation linking text and audio-video recordings, which use either a musical score presentation (such as ELAN, Exmaralda, etc. – see Schmidt & Wörner, 2009) or a list presentation (such as CLAN – see MacWhinney & Wagner, 2010). All transcripts for this paper have been made with the help of ELAN software [1]. The software can be used either as a simple tool, able to precisely calculate the temporal fragments indicated in the textual transcript, or as a transcription tool producing a fully-fledged annotation. The multimodal conventions used here produce a textual transcription that is fully compatible with the ELAN visualization. For instance, the following screen-shot is the ELAN annotation corresponding to the transcription of Extract 2:

(2d) (ELAN screen-shot)
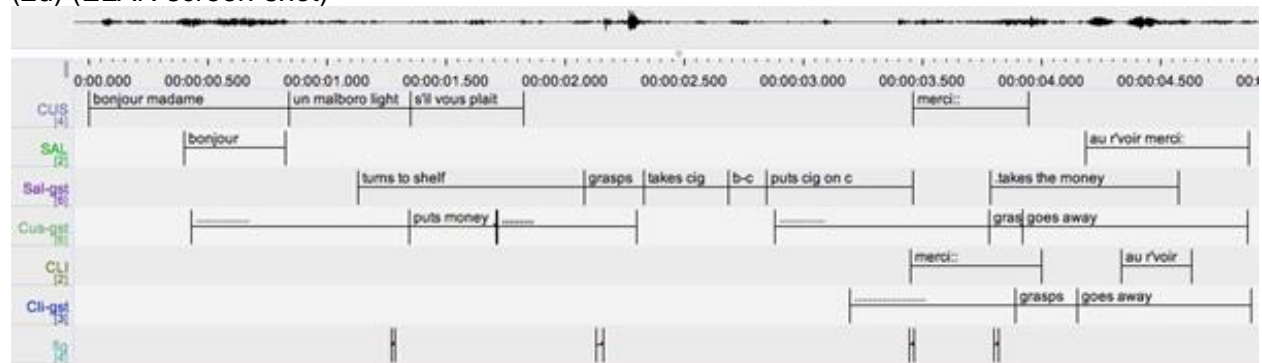
ELAN associates textual annotations with intervals of time, along a continuous chronological line. This is visualized as a "musical score", in which each participant, and each type of phenomenon, occupies a line. By contrast, the textual transcript visualizes the interaction in a "line by line" representation, which highlights the sequential organization of the unfolding actions (for a comparison see Mondada, 2007). The textual transcript, thanks to the convention used, also makes all temporal relationships more explicitly marked (with the symbols working as temporal

anchors precisely located in time). This is why the latter representation is used in the remainder of the article.

As shown in Extracts 1-2, there might be considerable variation relative to the way a request is performed by the customer (pointing at the product vs. already looking at their purse, projecting paying) and the way the seller's response is provided (early vs. later on). Various responses can be provided, implemented in different ways, and formatted by means of complementary multimodal resources, also exploiting different temporal arrangements:

(3a) (BAK_F_STL_avr_2-00)

```
1  SAL      bonsoir m[adame
            good evening m[a'm
2  CUS               [bonjou:r
                     [good morning
3           (0.7)
4  CUS      (alors) deux petits pains au chocolat
            (so) two little chocolate bread rolls
5           s'il vous pl•aît.=
            please
   sal →                •moves R_leg->
6  SAL →    =oui:,•
            =yes
       →        ->•steps laterally, twd R/pastries->>
```

After the exchange of greetings (1-2), the customer requests two chocolate breads (4-5). In response, the salesperson utters a positive response token ("oui:"/'yes' 6) and moves towards the area where the pastries are. The second pair part is therefore constituted of two actions: a short positive verbal response (6) and a movement towards the requested product (5-6). These two actions are formatted in different ways, with different resources, which also have different temporal features. Typically, the response token is short, uttered immediately after the completion of the request (latching). It is also produced with a rising intonation, indicating that there is more to come. By contrast, the actual action granting the request takes more time: it first requires a movement to the area where the product is, which projects the compliance proper, the gesture consisting of grasping it. As in the previous extract, the geography of the products within the shop (where they are distributed by type in different places) is a resource that makes movements in space recognizable as beginning the requested action. So, the distribution of verbal vs. embodied resources orients here to specific temporalities characterizing the responses. The verbal response is characterized by its immediacy and rapidity. The incipient embodied movement is characterized by its projectability. Finally, the completion of the response is implemented in the final, projected, grasping gestures (see the transcript below). These different practices show a temporal expansion and distribution of the response(s). It also shows the importance of considering (in the analysis as well as in the transcription) the emergent accountable projection and trajectory of a response (rather than reducing it to the outcome alone, which comes rather late).

The transcript can be further developed, with more granularity, showing the detail of the salesperson as well as the customer's unfolding actions:

(3b) (In-depth transcript)

```
1  SAL     bonsoir m[adame
           good evening m[am
2  CUS              [bonjou:r
                    [good morning
   sal     •>>walks twd counter->
   cus     ±>>steps twd counter->
3          (0.7)
4  CUS     (alors) *deux petits ±pains au choco*lat±
           (so) two  little chocolate bread roll
   sal            *LH on paper---------------*grasps paper-->
   cus                     ->±leans forward-----±
5          s'il vous •pl±aît.=
           please
                       ±withdraws from c and turns to pastries->
   sal            ->•moves R leg->
6  SAL     =oui:,*•
           yes
               ->•steps laterally to the pastries->
               ->*,,,,->
7          (0.8) * (0.15) ± (2.5)• (0.7) * (1.4) º (1.4) º
   cus                   ->±
   sal                        -->•
   sal           ->*                        *LH takes bag---->>
   sal                                       ºRH grasps tongºchoc br->>
```

This more detailed transcript shows how both participants bodily orient themselves to the request.

The customer was waiting a step away from the counter, and takes a final step towards it when the salesperson approaches the counter, walking in from the storeroom behind. The customer also leans forward a bit during the request, towards the salesperson. In this way, the customer organizes her maximal face-to-face involvement with the salesperson at that point (cf. Harjunpää, Mondada & Svinhufvud, 2018). As soon as the request is completed, however, at pre-completion of the final token "please" (5), the customer withdraws from the counter and turns to the display case where the pastries are. This movement anticipates the response of the salesperson, who is slower in walking around the counter towards the same area.

The salesperson approaches the counter and makes a gesture that anticipates a possible request: she grasps the bakery paper on the counter, used to wrap breads. She abandons that paper when she moves to the pastries, which are products wrapped in a paper bag. This gesture exhibits readiness to serve and orientation to an expected action by the customer, which is the request for a product.

The detailed transcript also shows how the request is complied with: the salesperson moves to the pastries, grasps a bag with one hand and tongs with the other. These artifacts and their mobilization are the conditions for finally completing the requested action, grasping the product and initiating the transfer of property from the display case to the customer's hands. Their visible and timed manipulation produces the accountability of the seller's action and of the customer's waiting for the request to be completely fulfilled.

4. Silently requesting a service by handing over an object

In the previous section, I have shown how a first multimodal action realized within a multimodal turn can be responded to in an embodied silent way. In this section, I show that the first action can also be silent. I also show that it can be responded to in a multimodal way, with, among other resources, the utterance of "thank you", co-occurring with various giving and taking gestures. There is a growing body of literature on thanking expressions and practices (Floyd et al., 2018), which have previously been described in closing environments (Schegloff & Sacks, 1973; Aston, 1995), and as a second pair part responding to different actions, such as offers (Schegloff, 2007: 76, 128). Schegloff (2007: 136) also comments on a "thank you" responding to an embodied action (actually as a closing third, acknowledging a complying action that responds to a request). Here I show several instances of "thank you" uttered in second position, in response to silent first actions in which the customer handles over a product or money.

The next extract shows six occurrences of thanking, within a recurrent pattern and some variations. The extract comes from a kiosk shop in Basel, in German-speaking Switzerland. A customer has been waiting for a while, and has already put a self-collected drink on the counter, while the seller serves another client, at the other counter ("merci" is the thanking expression in Swiss German, and "danke" is German).

(4a) (KIO_CH_BS_2_1-01-51-cli30) (simplified)

```
1            • (0.2) * (0.7)      *(0.2)
  sel             *extends RH*grasps drink---->
2 SEL     mer+ci*
          thanks
  sel         ->*brings to bc->
  cus        +hands over money->
3         (0.7)*(0.1)*   (1.0)   *(0.5)*(0.2)
  sel         ->*bar-c*back on c*,,,..*grasps money->
4 SEL     °mer*ci°
          thanks
  sel       ->*
5         (3.3)* (0.7) + (0.2)
  sel          *hands change over---->
  cus                +extends RH----->
6 SEL     danke schön
          thank you very much
7 CUS     danke*+ au
          thank you too
  sel        ->*
  cus         ->+grasps change-->>
8 SEL     adieu danke
          good bye thank you
```

The only talk in this encounter involves a series of thanking (2, 4, 6, 7, 8), which responds to a first, previous action, consisting of handing over either a product or some money (with the exception of 6, in which the seller thanks while giving back the change). The first *thank you* is uttered in response to the customer having put a drink on the counter. The seller produces it (2) after having grasped the drink (1). The second (4) is produced when taking the money from the customer (the grasping of the coins begins slightly before, 3). The third (6) is uttered by the seller when giving their change to the customer (5) and before the customer has taken it: the customer

responds with another *thank you* when grasping the coins (7). Finally, the encounter is closed by the salesperson's farewell and thanking (8). We can thus observe that *thank you* is produced when some object is being exchanged, given or taken – with the exception of the final thanking, which retrospectively treats the exchange as a whole and closes it. All but one thanking are produced by the seller, even when he is the one giving something (giving the change back, 7). The only thank you produced by the customer responds to the handing over of the change. These are also the only words proffered by the customer in the entire exchange.

In order to better understand both what *thank you* does and how it identifies and treats previous, first actions, a new version of the transcript might be useful.

(4b) (KIO_CH_BS_2_1-01-51-cli30) (more detailed)

```
1           (0.5)    ± (0.5) •± (1.3)  •
a sel    >>turns to till-•gz screen•
b cus    >>gz CLI±gz SEL--±gz away----->>

2         •  (0.1) + (0.1) * (0.7)+ £••(0.2)|(0.1)•£+(0.1)|*
a sel                        *ext-RH----*grasps drink---------*
b sel    •gz prod------------------•gz cus------•gz prod---->>
c sel                             £eyebrow-------£
d SEL                                      |merci-------|
e cus            +puts money on c+,,,,,,,,..........+hands money over->>
```

A more detailed annotation of the embodied conducts shows a specific distribution of gaze, with the customer looking away at the beginning of the encounter (1) and never looking straight at the salesperson (she only briefly looks at the change received, 4e, before moving away). The only moment she looks at him is when he turns away from the previous client (1a): but instead of gazing at her, he orients to the till and its screen, where he performs a last operation. The customer was staring for a long time at the previous client (1b), who actually by-passed her in the queue, looking at her with an angry face: when the seller turns in her direction, projecting serving her, she looks at him, but sees that he is not yet available, and looks away. She will maintain this posture for the entire encounter. This peculiar transition from one encounter to the other might also explain the absence of greetings, in the absence of any moment of visual reciprocity.

The second version of the transcript is not only more detailed, but also treats the action of thanking in a different way, embedding it in the course of the ongoing actions instead of singling it out on a separate line. In this way, thanking is treated as one resource among others, contributing to the responsive action alongside them. This enables a different appreciation of responsive thanking.

As observed in the first version of the transcript, the first thanking ("merci") is produced in response to the client putting the drink on the counter, and while the seller grasps it (2a), before reading its bar-code (3d). The second version of the transcript enables another vision of the environment in which "merci" is uttered. It shows that the second, responsive, action is not only constituted of the reception of the object, but co-occurs with a quick gaze on the customer (2b) and raised eyebrows (2c). This holistic pattern implements an action that does more than just taking or thanking: it addresses the customer for the first time in a personal way – maybe orienting

towards the absence of greetings and a mutual gaze. Interestingly all this clusters around the moment in which the object is grasped.

If we now turn to the subsequent actions, comparable clusters can be observed:

(4c) (cont.)

```
3           *    (0.7)    +*(0.1)*   (1.0) ** (0.5)*(0.2)+|(0.2)*+(0.3)|
a cus                --->+holds money twd SEL---------+,,,,....+takes drink->
b sel                                 ->•gz money-->
c sel       *takes to bc*bar-c*back on c*,,,,...*grasps money*
d SEL                                             |--°merci°---|

4           (0.6)+(0.2)+(2.9)*(0.2)•(0.1)+±(0.2)|(0.4)±(0.2)|(0.2)*+(0.3)|
a sel                     *hands change over-------------------*,,,->>
b sel                  -->•gz CUS----------------------------------->>
c SEL                               |danke schön|
d cus        ->+,,..+holds RH twd SEL--+extends RH-------------+grasps change->>
e cus                                  ±gz money----±turns away---->>
f CUS                                             |-danke au--|

5 SEL    adieu danke
```

The customer projects the paying of her drink very early –hence hinting at the fact that it is the only product she wants– by handing over the money for paying (2e) as soon as the seller has taken the drink to process it at the bar-code reader. She formats this action in two successive, self-repaired ways, first by putting some coins on the counter, then by extending her arm with the money. This second way of formatting relies on the reciprocal gesture of the seller, and also augments the temporal pressure on him (see Mondada and Sorjonen (in press) for a systematic analysis of these two ways of transferring an object over the counter). When the seller is available again, after having put the drink back to the counter (3c), the customer is visibly holding the money (3a). The seller gazes at the coins (3b) and grasps them (3c), then saying "°merci°" (3d). Here too, the responsive action to the initiation of paying is comprised of these three practices, which are successively orchestrated, one projecting the other. In this environment, "°merci°" is the final element closing the complex multimodal Gestalt, organizing the response to the customer.

The final two paired *thank you*s are produced, when the seller brings back the change and the customer grasps it. Here the seller thanks while giving, rather than taking, the money, to the customer who receives it. Although it initiates the sequence of giving back the change, it also responds to the customer beginning to hold her hand towards him very early (4d, almost 3 seconds before he hands her the change – cf. 2e when she was also holding the money to pay very early on). He gazes at her (4b) before thanking her but, although she reorients to him, she looks at the money instead, before rushing away (4e) – in such a way that visual reciprocity is again not established. Reciprocity is instead established by the response to "danke schön" (4c) with "danke au" (4f), by the use of the same form ("danke" instead of "merci", more colloquially and in the local dialect) and the addition of "au"/"too". We notice that the seller was using the more informal "merci" in the two previous cases, and that now he uses the more formal "danke schön" – perhaps orienting to the persistent absence of response from the customer. The final greeting is produced in a mixed form, combining "adieu", in the local Swiss German dialect, and "danke" in standard German.

The notation of thank you in the same way as other multimodal resources, along the time line, makes it clear that thanking is embedded within the embodied actions constituting the complying action proper, temporally fitted to the grasping gesture but also to the (re)establishment of a possible mutual gaze. The precise position of thank you in this complex multimodal Gestalt orients to the moment in which the object changes hand.

Another instance of a silent first action being responded to by thanking is the following, from a kiosk encounter in Fribourg (French/German-speaking Switzerland) in French, in which it is performed in a rather delayed way:

(5) (KIO_CH_FRI_080515_left_31-08)

```
1  CUS     bonjour
           hello
   cus     >>approaching the counter from the shop entrance-->
   sal     >>walking to the counter from the backroom-->
2          (1.1) + (0.9) ±  (0.15) * (0.15)+ (0.2)
   cus            +hands over the drink on c+
   cus                  ->±stops and stands---->
   sal                           *...->
3  SAL     °bon°jour*
           °hell°o
   sal            ->*graps drink->
           (0.2) * •
   sal          ->*moves drink twd barcode->
   sal            ->•stops in front of CUS-->>
4  SAL     mer:ci,
           thanks
```

The customer approaches from the shop entrance with a drink in his hand, which he has taken from the self-service area, and which he ends up putting on the small counter. Meanwhile the seller comes out from the backroom. The customer is the first to greet (1) and the seller responds with some delay (3), as she is still walking towards the customer, and has also already grasped the product. Here "mer:ci"/"thanks" is proffered after having taken the product, but before the product is fully processed: the thanking token is placed at the moment in which the seller stops in front of the customer and her hand moves the product towards the bar-code, projecting its processing. So, despite being produced slightly later than in the previous extract, the token is again placed in a transition phase, just after the product has transited from the customer's to the seller's hands, and as the processing is projected.

In these two cases, the temporal position of *thank you* shows to which first action it responds and contributes to. It shows how participants retrospectively treat first actions, identifying them in the flow of ongoing movements, and the way sequences of actions are locally and temporally configured by them. The embeddedness of the token within dynamic multiple multimodal trajectories constituting a complex multimodal Gestalt accounts for the token doing more than just thanking.

5. Embodied silent sequences

It not uncommon that both first and second actions constituting a requesting sequence are achieved in silence, in embodied ways. This is the case of the following except, recorded in

Lugano, in Italian-speaking Switzerland, in which the only verbal actions are the greetings and the final price announcement:

(6a) (LUG_35-10)

```
1           (2.7) +*(0.1)
   cus            +hands over newspaper->
   sal             *extends LH->
2  SAL      buongiorno signo#ra
            good morning m'am
   fig                        #fig.1
3           (0.2)*
   sal          ->*grasps->
4  CUS      'ng+*iorno+
            'd morning
   cus          ->+,,,,,,+
   sal           ->*brings coupon to b-c->>
5           (1.4) *(0.5)*              (1.3)         *
   sal               ->*scans*puts nwp on the counter*
6  SAL      due e ven:ti, signora
            two and twenty m'am
```


1

The customer hands over a newspaper she picked up in the kiosk, displaying her purchase of the product. Almost at the same time, the seller extends her left hand projecting the grasping of the object – aligning with the projected action of the customer (1). Only at this point do they greet (2, 4), with the greeting sequence being initiated by the salesperson (2). In the meanwhile, the seller grasps the newspaper (Figure 1), and takes it back, bringing it to the bar-code reader (5). After the scan, she puts it back on the counter and announces the price (6).

This is an example of the unproblematic handling and taking of an object, in silence. Greetings are not initiating the encounter, they are not displaying availability – they come later on. This could be interpreted as a situation of multiactivity (in which various activities are happening in parallel, see Haddington et al., 2014) rather than a classical instance of opening (in which the greeting sequence would be followed by the reason for the visit, the purchase). The way the transcript is presented can demonstrate this secondary role of the greetings. As in Extracts 4b and 4c, the transcript is organized on the basis of the unfolding timeline, rather than on unfolding talk:

(6b) (alternative transcript of lines 1-5)

```
1             (2.7) +*(0.1) |     (0.7)     #(0.2)| (0.2) * |(0.2)*+   |+
a   cus             +hands over newspaper-----------------------+,,,,+
b   sal             *extends LH----------------------*grasps-*takes it back->
c   SAL                   |buongiorno signo#ra|
d   CUS                                                  |'ng*+iorno|
e   fig                               #fig.1
2             (1.4) *(0.5)*            (1.3)           *
a   sal             ->*scans*puts nwp on the counter*
```

In this alternative transcript, the handing over of the product (1a) and the salesperson's grasping it (1b) are displayed first, beginning with the initiating action (1a) and then the response (1b). These two actions are symmetrically organized.

The greetings are exchanged during these actions, and they are temporally positioned in such a way that they highlight precise moments within the sequence. The first greeting is given as both participants begin to reciprocally extend their arms and hands (1c), the second one as the seller grasps the object handled over by the customer (1d). This representation highlights the location of the verbal turns in relation to the ongoing embodied trajectories –instead of the reverse (in Transcript 6a)– that is, precisely as giving/taking gestures are just emerging and exactly as they are completed.

This version makes it clear that giving/taking the product happens not only before the greetings, but also with both participants extending their arms at the same time (1a-1b). How are these perfectly synchronized gestures possible? In which circumstances can these two actions happen simultaneously (vs. one participant initiating one and the other one responding to it, as in the first extracts analyzed here and in particular in extract 2b)?

The circumstances that make this simultaneity possible are not documented in the previous transcripts: they can be discovered in a further version that expands the initial 2.7 seconds of line 1, which features not only the customer (CUS) and the salesperson, but also the previous client CLI:

(6c) (in-depth transcript of line 1, initial 2.7 seconds)

```
1        (0.4) ±          (0.2)     £(0.2)£(0.5)±(0.2)*(0.6)#(0.2)£(0.4)*
   cli    >>puts coupon in purse£        £steps lat--------------£
   cus          ±gz at CLI-------------------±gz at SAL--->
   sal                                       ->*closes till-----*
   fig                                           #fig.2
2        +*(0.1) |          (0.9)      | (0.2) * |(0.2)*+    |+
   cus    +hands over newspaper------------------------+,,,,+
   sal     *extends LH---------------------*grasps*takes back->
   SAL            |buongiorno signora|
   CUS                                     |'ng**+iorno|
```


2

Just before the participants give/take the newspaper the previous encounter closes. The transition to the next encounter –with our target– emerges as the seller is still engaged with the previous client (CLI), who is putting away her purchase in her purse, and steps aside. Her actions are monitored by the imminent customer (CUS) who waits to one side and looks at her. It is remarkable that as soon as CLI makes a step away, CUS shifts her gaze to the seller (Figure 2), orienting to the completion of CLI's purchase. At this point, the seller, who was ordering bills in her till, closes the till. This action is seen by the customer; moreover, it is audible since it makes a clear sound. The closing of the till is treated by both imminent co-participants as making them available for the first step in their mutual engagement: the customer hands over the product and the seller extends her hand towards the customer at the same time (2). Immediately, the seller also greets the customer. This timed reciprocity is made possible by the publicly visible and audible steps towards the closing of the previous encounter, clearly projecting the precise moment of a possible next opening and visibly monitored as such by both co-participants. This shows not only the relevance of preliminary actions and events, and their projective potential for the achievement of reciprocity in the form of simultaneous symmetric embodied actions, but also of prior sequences of actions, as they are glanced at, monitored, and witnessed by the relevant parties. Careful analytical attention to the embodied details shows how subtle the multimodal organization of sequence and sequentiality can be, and how finely and deeply embedded it is within ongoing flows of embodied actions.

6. Conclusion

This paper has proposed an analysis of the organization of silent embodied actions within sequences of paired actions and highlighted the importance of the representation of their timeliness and granularity in adequate transcripts. In so doing, the aims are both to address issues in multimodal transcription and issues in the analysis of silent actions in social interaction –the latter constituting an interesting challenge for the former.

While silent embodied actions responding to requests have been increasingly noted in the literature, silent first actions have been less studied, as have silent pairs of actions, with regard to sequence organization – although they have been explored in particular settings, such as workplace interactions (see for example the importance of silent actions for the understanding of coordinated actions in surgery: Mondada, 2014b, Heath et al., 2018). This paper has explored some variations within the same transcript conventions, in the visual-textual representation of: a) their respective fine-grained temporality, projectability, and trajectories; b) their public accountability, locally achieved by the co-participants' gazes, glances, and reciprocal monitoring; and c) their situated arrangement with other actions, achieved through verbal resources (such as greeting and thanking).

The proposed transcript notation enables reflection upon the complex emergent and sequentially unfolding temporality of multimodally formatted actions. Taking the preparatory movements into account enables consideration of how early an action emerges/is initiated. This depends in a crucial way upon what counts as a preparatory move, projecting that action, but it also depends on how that incipient action is interpreted by the co-participant and responded to (as in Excerpt 2), as well as on how trajectories of action are not only designed and interpreted, but also eventually revised (as in Excerpts 2, 3 and 4). Taking into account the trajectories of these movements also enables consideration of how emergent responses are accountably designed and diverse resources are mobilized to make them visible (such as nodding before stepping, Excerpt 1, "ouais" before walking/before the movement of the leg is visible/seen, Excerpt 3) and audible (such as by thanking and greeting while grasping an object in Excerpts 4-6).

The distributed arrangement of multimodal details, as well as their temporal precision and location, build complex multimodal Gestalts which are both indexical –depending on the local ecology and on the contingencies of the ongoing action– and systematic (as shown by the co-occurrence of grasping/taking gestures, gaze, and thanking or greeting in Excerpts 4-6). In order to highlight their temporal relations, different forms of spatialized annotations in the transcripts were explored, and their consequences made explicit.

In more general terms, the paper shows how the temporality of embodied details –when adequately documented and represented– enables the organization of sequence and

sequentiality to be revisited. Emergent embodied movements, sometimes minimal and microscopic, have preparatory and projective features that make them identifiable and witnessable very early on as incipient first or second actions – so early on that their recognizability, intelligibility, and accountability are still very fragile, and uncertain. Early preparatory movements can be the locus of embodied repairs, showing how revisable and renegotiable a trajectory of embodied action is. These movements can also be further stabilized and fully expanded into evident and normatively expected actions. This shows how the nature of an action can be initially undetermined (and even invisible) during its trajectory, and only progressively become relevantly recognizable and categorizable. This affects the interpretations of both analyst and the participant regarding when an action begins, when a response emerges as such, when it is recognized and when it becomes intersubjectively accountable. The evidence provided by the embodied details of social interaction show that the nature of actions, their form, their meaning, their relevance, and their sequential import are a dynamic phenomenon, acquiring a meaningful potential quite early but assuming a definitive meaningfulness only later. This is a general characteristic of multimodally formatted actions; silent embodied actions constitute a particular kind of action that is exemplary in this respect. They are thus a stimulating challenge for multimodal transcriptions and analysis. This article attempts to show that it is possible to deal with them in ways that highlight their finely tuned sequential organization. This in turn enables an analysis to treat verbal and embodied resources in equal and symmetric ways, and to reflect upon their situated use, interpretation and distribution by the participants.

Conventions

Talk has been transcribed following the conventions of Gail Jefferson (2004). Embodied conducts were transcribed following the conventions of Lorenza Mondada (see 2018a for a conceptual discussion, and for a tutorial).

References

Aston, G. (1995). Say 'thank you': Some pragmatic constraints in conversational closings. *Applied Linguistics* 16(1): 57-86.

De Stefani, E. (2011). *'Ah petta ecco, io prendo questi che mi piacciono'. Agire come coppia al supermercato. Un approccio conversazionale e multimodale allo studio dei processi decisionali.* Roma: Aracne.

Deppermann, A. (ed.) (2013). Conversation analytic studies of multimodal interaction. Special Issue, *Journal of Pragmatics*, 46:1, 1-172.

Floyd S., Rossi G., Baranova J., Blythe J., Dingemanse M., Kendrick K., Zinken J., Enfield N.J. (2018). Universals and cultural diversity in the expression of gratitude. *R. Soc. open sci.* 5: 180391.

Fox, B. (2015). On the notion of pre-request, *Discourse Studies*, 17(1): 41-63.

Fox, B. & Heinemann, T. (2015) The alignment of manual and verbal displays in requests for the repair of an object, *Research on Language and Social Interaction*, 48:3, 342-362,

Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 97-121). New York: Irvington Publishers.

Goodwin, C. (1994). Professional vision. *American Anthropologist* 96(3): 606-633.

Goodwin, C. (2017). *Co-operative action.* Cambridge: Cambridge University Press.

Goodwin, M. & Cekaite, A. (2014). Orchestrating directive trajectories in communicative projects in family interaction. In P. Drew & E. Couper-Kuhlen (Eds.) *Requesting in social interaction* (p. 181-210). Amsterdam: Benjamins.

Goodwin, M. H. (1980). Processes of mutual monitoring implicated for the production of description sequences. *Sociological Inquiry, 50*(3-4): 303-317.

Haddington, P., Keisanen, T., Mondada, L. & Nevile, M. (eds.). (2014). *Multiactivity in social interaction: Beyond multitasking.* Amsterdam: Benjamins.

Harjunpää, K., Mondada, L. & Svinhufvud, K. (2018). Coordinated entry into service encounters in food shops: Managing interactional space, availability, and service during openings. *Research on Language and Social Interaction*, 51(3), 271-291.

Heath, C. (1986). *Body movement and speech in medical interaction.* Cambridge: Cambridge University Press.

Heath, C. (2013). *The dynamics of auction: Social interaction and the sale of fine art and antiques.* Cambridge: Cambridge University Press.

Heath, C., Luff, P., Sanchez-Svensson, M., Nicholls, M. (2018). Exchanging implements. *Sociology of Health & Illness.* 40(2): 297-313.

Heath, C., Hindmarsh, J. & Luff, P. (2010). *Video in qualitative research.* Sage, London.

Hoey, E. (2017). *Lapse Organization in Interaction.* PhD dissertation. Nijmegen: MPI.

Ivarsson, J., Greiffenhagen, C. (2015). The organization of turn-taking in pool skate sessions. *Research on Language and Social Interaction*, 48:4, 406-429.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the first generation* (pp. 13-31). Amsterdam: Benjamins.

Laurier, E. (2014). The graphic transcript: Poaching comic book grammar for inscribing the visual, spatial and temporal aspects of action. *Geography Compass*, 8, 4, 235-248.

Lindwall, O., Ekström, A. (2012). Instruction-in-interaction: The teaching and learning of a manual skill. *Human Studies*, 35: 1, 27-49.

Luff, P., Heath, C. (2015). Transcribing embodied action. In D. Tannen, H. E. Hamilton, and D. Schiffrin (eds.). *The handbook of discourse analysis.* New York: Wiley.

MacWhinney, B. & Wagner, J. (2010). Transcribing, searching and data sharing: The CLAN software and the TalkBank data repository. Gesprächsforschung, 11, 154-173.

Mondada, L. (2006). Video recording as the reflexive preservation and configuration of phenomenal features for analysis, in Knoblauch, H., Raab, J., Soeffner, H.-G. & Schnettler, B. (eds). *Video Analysis* (pp. 51-68). Bern: Lang.

Mondada, L. (2007). Transcript variations and the indexicality of transcribing practices. *Discourse Studies*, 9/6, 809-821.

Mondada, L. (2014a). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics, 65,* 137-156.

Mondada, L. (2014b). Instructions in the operating room: How surgeons direct their assistant's hands. *Discourse Studies, 16/2,* 131-161

Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics.* 20:2, 2-32.

Mondada, L. (2017). Imperatives in action. Precision timing and timed embeddedness in embodied courses of action. In Sorjonen, M.L., Raevaara, L., Couper-Kuhlen, E. (eds.). *Imperatives in Interaction* (pp. 65-101) Amsterdam: Benjamins.

Mondada, L. (2018a). Multiple temporalities of language and body in interaction: challenges for transcribing multimodality, *Research on Language and Social Interaction,* 51:1, 85-106.

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

Mondada, L. & Sorjonen, M.-L. (2016). First and subsequent requests in French and Finnish kiosks. *Language in Society*, 45, 733-765.

Mondada, L. & Sorjonen, M.-L. (in press). Trajectories of actions with products in sales encounters: Putting/giving products on/over the counter at convenience store. In Fox, Mondada, Sorjonen (eds.). *Encounters at the counter*. Cambridge: Cambridge University Press.

Mondémé, C. (2018). Comment parle-t-on aux animaux ? Formes et effets pragmatiques de l'adresse aux animaux de compagnie, *Langage et société*, 163, 77-99.

Nevile, M. (2015). The embodied turn in research on language and social interaction, *Research on Language and Social Interaction*, *48:2*, 121-151,

Rossi, G. (2014). When do people not use language to make requests? In P. Drew & E. Couper-Kuhlen (eds.). Requesting in Social Interaction (pp. 303-334). Amsterdam: Benjamins.

Sacks, H., & Schegloff, E. A. (2002). Home position. *Gesture, 2*:2, 133-146.

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. (2007). *Sequence organization in interaction: A primer in conversation analysis* (Vol. 1). Cambridge: Cambridge University Press.

Schmidt, T. & Wörner, K. (2009). Exmaralda – Creating, analyzing and sharing spoken language corpora for pragmatic research. *Pragmatics*, 19:4, 565-582.

Streeck, J., Goodwin, C., & LeBaron, C. (eds.). (2011). *Embodied interaction, language and body in the material world*. Cambridge: Cambridge University Press.

---

[1] https://tla.mpi.nl/tools/tla-tools/elan/