Methodological explorations of interpreter-mediated interaction: novel insights from multimodal analysis

Abstract

Research in Dialogue Interpreting (DI) has traditionally drawn on qualitative analysis of verbal behaviour to explore the complex dynamics of these 'triadic' exchanges. Less attention has been paid to interpreter-mediated interaction as a situated, embodied activity where resources other than talk (such as gaze, gestures, head and body movement, proxemics) play a central role in the co-construction of the communicative event. This paper argues that understanding the complexity of DI requires careful investigation of the interplay between multiple interactional resources, i.e. verbal in conjunction with visual, aural, embodied and spatial meaning-making resources. This call for methodological innovation is strengthened by the emergence of video-mediated interpreting, where interacting via screens without sharing the same physical space adds a further layer of complexity to interactional dynamics. Drawing on authentic extracts from interpreter-mediated interaction, both face-to-face and video-mediated, this paper problematizes how the integration of a multimodal perspective into qualitative investigation of interpreter-mediated interaction can contribute to the advancement of our understanding of key interactional dynamics in DI and, in turn, broaden the scope of multimodality to include new, uncharted territory.

Keywords: multimodal analysis, dialogue interpreting, methodology, face-to-face interaction, video-mediated interpreting

Introduction

In the present era of globalisation, migration, mobility and web-connected communication, opportunities for people of different cultural and linguistic backgrounds to interact and communicate with one another have grown exponentially. Dialogue Interpreting (DI) has become a key communicative practice in real-life scenarios to enable two or more parties from different linguistic and cultural backgrounds to interact with each other. In contrast to simultaneous interpreting, where interpreters operate from a booth and reach the audience through the audio channel, DI is carried out consecutively, mostly face-to-face, with participants sharing the same interactional space, which may be physical or virtual (as is the case in video-mediated

interpreting, VMI). During DI, all participants, interpreter included, interact with one another directly through multiple verbal and embodied semiotic resources (such as gaze and head movement, posture and body orientation, gestures, facial expressions). Their use is strongly linked to the specific ecologies of action in which such resources are produced. In VMI, interacting through a screen adds further multimodal complexity to the interactional dynamics. DI and VMI can therefore be considered very specific types of multilingual, multiparty spoken interaction.

Multimodality offers a new perspective of "representation, communication and interaction as something more than language" (Jewitt 2013:1), and has been adopted by researchers from diverse theoretical and methodological perspectives to investigate how verbal and embodied resources interact to create meaning (Jewitt 2014). However, most research on DI has focused on the verbal dimension, with little account of how resources other than language contribute to meaning making in such complex scenarios. A multimodal turn has not yet been fully undertaken in this field and would certainly "throw up some provocative issues for qualitative research methodology" (Dicks et al. 2011: 227). In turn, through the study of DI, multimodality can broaden its scope to a new, uncharted territory, and apply its methodological tools to 'a-typical' sites presenting their own potentialities and constraints. This paper aims to contribute to the growing literature on the interactional use of multimodal resources through analysis of a specific case study. Undertaking a *methodological exploration* in multimodality, the

paper proposes a rigorous and holistic approach to account for integrated verbal and embodied resources in DI.

Dialogue Interpreting as practice and research

Finding a suitable research approach to account for integrated semiotic resources in DI requires a clear conceptualization of the object of study. DI scenarios are also known as triadic exchanges (Mason 2001) or communicative *pas de trois* (Wadensjö 1998), two expressions that recognise the interpreters' visibility and fundamental responsibility with regard to the negotiation of meaning in interaction.

Literature refers to DI in many different ways, including as *community-based*[1], *liaison,* and *bilateral* interpreting, among others, indicating the nuanced differences in specific aspects of the interpreting process, such as setting, mode[2], professional status or the interpreter's area of specialization (Hale 2007: 27-30). *Dialogue* interpreting (Mason 1999; Wadensjö 1998) is used in this paper, as it "seeks to encompass a group of activities seen as sharing an overall mode of interaction rather than a particular term", thus transcending boundaries and focusing on the "characteristics of a particular mode of interaction, shared in many, quite diverse socio-professional contexts" (Mason

[1] Hale (2015) refers to community interpreting as the mode of interpreting that involves people who are part of the same 'community', society or country but who do not share a common language, including, for instance, public service, legal and religious settings.

2 The term 'mode' in Interpreting Studies (IS) refers to the specific method in which interpreting is delivered, which can range from monologic modes, such as consecutive and simultaneous, to dialogic ones, such as DI. This overlapping of IS and multimodal research terminology poses a particular challenge.

2009:81). It also highlights the focus on dialogic communication, as opposed to the mostly unidirectional nature of conference interpreting.

DI interaction has been investigated from diverse disciplinary perspectives as a linguistic and socio-cultural practice (for a review, see Hertog and Van der Veer 2006), and these works have enriched our knowledge of DI interactional dynamics. The variety of approaches applied to the investigation of DI suggests it is a field at the intersection of different disciplines and research traditions, with divergent, but complementary, insights from studies that have adopted very different theoretical and methodological lenses. Particularly relevant to the present paper are micro-analytical studies of DI dynamics based on transcripts of recordings from naturally-occurring interaction, an approach which gained momentum in the late 1990s and led to the development of the "dialogic discourse-based interactionist paradigm" (Pöchhacker 2004: 73). This paradigm encompasses descriptive, qualitative, empirical and discourse-oriented studies that rely on a variety of (mainly ethnomethodological) methods for microanalytical investigation of DI, such as Conversation Analysis (CA), Discourse Analysis (DA) and Critical Discourse Analysis (CDA), depending on the specific focus of the research. This analytical approach to DI was pioneered by two scholars: Roy (1989, 2000) analysed instances of student-teacher meetings mediated by a sign-language interpreter, with a particular focus on turn-taking dynamics; Wadensjö (1992, 1998) applied a micro-analytical approach to authentic interpreter-mediated immigration and medical

interviews to unveil how dialogue interpreters are active participants in the communicative event, rather than invisible language conduits.

The micro-analytical focus on what interpreters *actually do* in interaction was revolutionary in that it deconstructed traditional views prescribed by (inter)national organisations and codes of conduct, where interpreters were regarded as neutral, discrete, passive and transparent (Mason and Ren 2012). Conversely, this research highlighted the role played by interpreters in creating relationships between primary parties - acting not only as translators, but also as coordinators and intercultural mediators (Wadensjö 1998). Another merit of this body of research is to have shown how *all* parties contribute to the construction of meaning and understanding in the social activity, where "the outcome of the interpreter's work is dependent on the primary participants, on their mutual relations, on how they relate to the interpreter and on their communicative style" (Wadensjö 1999: 248).

The insights into the multifaceted nature of the interpreter's tasks and activity were mostly gained through close scrutiny of how verbal resources (including linguistic and paralinguistic ones) are used to coordinate social interaction (e.g. code-switching, Anderson 2012; minimal responses, Gavioli 2012; repairs and repetitions, Straniero Sergio 2012), how these shape variation in interpreter rendition (e.g. Wadensjö 1998; Baraldi 2012; Braun 2016; Jacobsen 2003) as well as turn-taking management (e.g. Davidson 2002). More recently, it has been argued that the multilingual and multiparty nature of DI makes it suitable for multimodal fLang (Pasquandrea 2011), yet this entails

a reconceptualization of DI as embodied activity where meaning is co-constructed through the integration of multiple semiotic resources:

> A cornerstone in a dialogical theoretical framework is the embodiment of spoken language. Applied to studies of dialogue interpreting, this means that there is a need to highlight participants' bodily orientation, gestures and gaze.
>
> (Wadensjö 2004: 108)

The call for multimodality to be applied more systematically to DI research reflects a growing trend to explore the mutually constitutive roles of embodied resources and talk in interaction, yet work in this field has encountered methodological challenges, as selectively discussed below.

Multimodal approaches to DI: a critical overview

A select but growing body of research in DI has investigated how embodied resources can complement, replace or integrate talk, and trigger, disambiguate or modulate specific interactional moves. The main 'pioneer' researchers in multimodal studies of DI can be identified in Lang (1976, 1978), Poyatos (1997, 2002) and Apfelbaum (1998). Lang (1976, 1978) investigated the gaze, posture and gesture patterns of an interpreter-mediated court case, noticing the systematic use participants made of (averted or direct) gaze and gesture (such as hands outstretched) as devices for signalling involvement and

exclusion, and for distributing turns-at-talk. Lang also found that the interpreter's preference for averted gaze, intended to signal neutrality or detachment, may interfere with smooth turn-taking in that important cues may be missed. Poyatos' (1997, 2002) tripartite model of multichannel communication investigated the interplay between the verbal, paralinguistic and kinesics, and was successfully applied to simultaneous and consecutive interpreting modes. Apfelbaum (1998) focused on the impact of physical position on participants' rhythmic synchronization of talk in interpreter mediated interaction based on the interpreter's projection of a next turn.

Research investigating resources beyond the verbal dimension grew throughout the 2000s, with most grounded in the analysis of naturally occurring data in face-to-face scenarios, although more recent studies focus on VMI. This body of literature can be selectively reviewed by placing studies on a timeline in chronological order (Figure 1) and reviewing them according to four main dimensions: settings explored, embodied resources integrated, phenomena investigated and multimodal frameworks adopted.

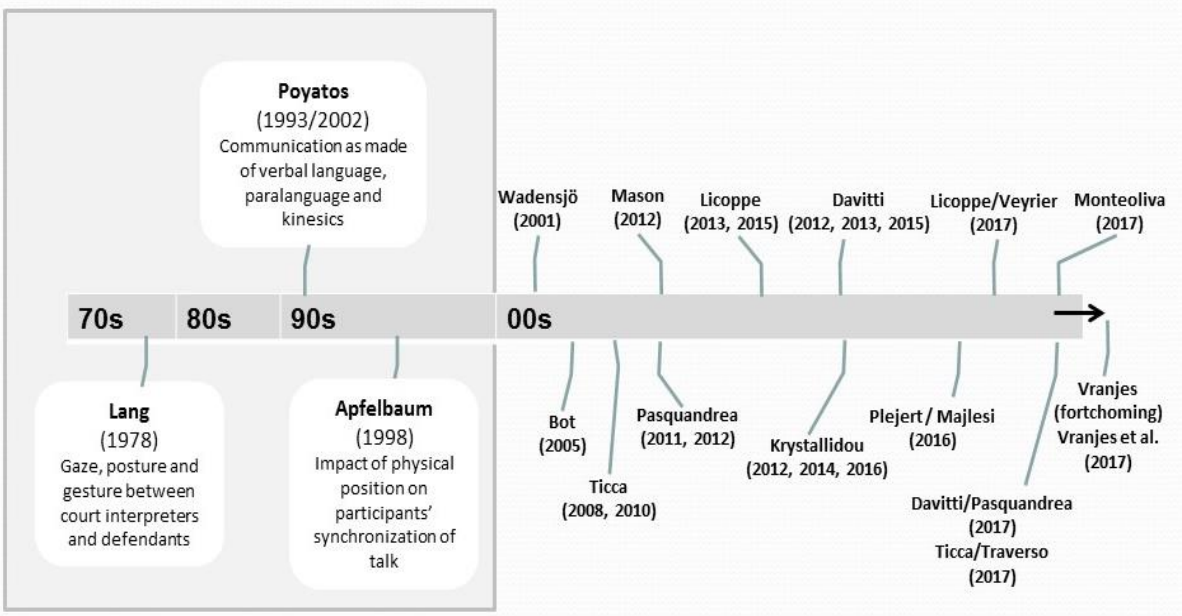


Figure 1: Evolution of multimodal approaches of DI on a timeline

Regarding the first dimension, i.e. settings explored, there is a predominance of research in the medical field, ranging from doctor-patient encounters (e.g. Pasquandrea 2011, 2012, Ticca 2008, 2010, Krystallidou 2012, 2014, 2016) to psychotherapeutic (e.g. Wadensjö 2001, Vranjes forthcoming, Vranjes et al. 2017), mental health (e.g. Bot 2005) and dementia evaluation settings (Plejert and Majlesi 2016). These are followed by studies in legal settings, including police interviews (e.g. Monteoliva García 2017) and courtrooms (e.g. Licoppe 2013, 2015; Licoppe and Veyrier 2016), alongside pedagogical settings (e.g. Author 2012, 2013; Author et al. 2017). In addition, two studies have looked at interaction in asylum seeking (Mason 2012) and social service (Ticca and Traverso 2017) settings.

Focusing on the embodied resources dimension, gaze and head orientation have been integrated with analysis of verbal interaction, with Bot (2005), Mason and Davitti (2016) highlighting striking differences in how participants engage in mutual gaze and gaze aversion, and Wadensjö (2001) noting how participants' seating arrangement and relative position to one another impact on the rhythmic regularity of talk and construction of joint narratives. Recent studies have integrated gesture and proxemics (e.g. Pasquandrea 2011, 2012; Krystallidou 2012, 2014), and handling of artefacts (e.g. Davitti and Pasquandrea 2017; Plejert and Majlesi 2016, Ticca and Traverso 2017) in the analysis of interactional dynamics.

Studies that have investigated the diversity of semiotic resources used in interaction have mostly focused on the dynamics of inclusion and exclusion of primary parties (a key analytical focus in encounters that are characterized by a certain level of power differential) and on how the coordinating role of the interpreter manifests itself (one of the main themes in DI research). This research has highlighted the embodied nature of how shifts between triadic and dyadic interactional formats are negotiated (e.g. Davitti and Pasquandrea 2017; Monteoliva García 2017) and how specific actions are achieved collaboratively, for example when giving instructions (Plejert and Majlesi 2016), providing assessments (Davitti and Pasquandrea 2017), explaining administrative procedures (Ticca and Traverso 2017) or revisiting phenomena already explored through verbal resources only, such as interpreter-produced expansions (Davitti 2013) .

These studies have relied upon very different multimodal research traditions to conduct their analyses. A substantial group has adopted a multimodal approach to CA (e.g. Davitti and Pasquandrea 2017; Licoppe and Veyrier 2017) to study the organisation of social interaction with a focus on talk and bodily action, whilst others have integrated embodied cues within a DA framework (e.g. Wadensjö 2001; Mason 2012), or adopted a mixed approach (e.g. Bot 2005 integrates fine-grained DA with interviews and concept maps; Krystallidou 2012, 2014 combines CA with Norris' 2004 framework for multimodal discourse analysis (MDA); Vranjes (forthcoming) uses multifocal eye-tracking to investigate gaze direction as an active communicative signal).

This selective review of multimodal approaches to DI suggests the potentialities of a multimodal approach to re-envisage known phenomena in DI, and to make visible the complexity of DI interaction, which is often underestimated by practitioners and users of interpreting services. However, it also reveals an eclectic methodological and theoretical landscape, with the lack of a systematic multimodal method for DI, making comparability across findings and settings problematic.

Towards a systematic approach to multimodal analysis in DI

As discussed, the diversity of frameworks used by multimodal studies of DI has led to a lack of transparency about how specific models or concepts are used, which is perpetuating a state of "doorstep interdisciplinarity" (Gile 1999:41), where existing frameworks (or parts of them) are opportunistically applied to DI. It is therefore timely

to reflect on the "divergences in conceptual and methodological orientations" (Pöchhacker 2004:75) and to work towards developing a more comprehensive framework for multimodal investigations of DI. Advocating a 'multimodal turn' in DI, this paper embraces a view of "language and talk as fundamentally embodied" (Mondada 2016: 340), viewing verbal and semiotic resources as intrinsically connected and talk-in-interaction resulting from their constant interplay. Building on this premise, a number of issues need to be addressed to advance the multimodal turn in DI.

A first principle of multimodal research is to include *all* semiotic resources contributing to a specific phenomenon, rather than prioritising one resource (e.g. speech). In other words, the verbal dimension of interaction is no longer the unique springboard for analysis of specific interactional phenomena, but other "interactionally loaded" (Sidnell 2009: 397) phenomena, such as gaps in conversation or moments where specific embodied behaviours are displayed (in combination or not with talk) need to be accounted for. For instance, Davitti and Pasquandrea (2017) identified artefact manipulation as a starting point for their investigation of shifting participatory frameworks in mediated interaction in pedagogical settings. Exploring how a specific object (school report) was handled in two selected sequences from the same corpus of parent-teacher meetings (PTMs), they identified that the same action (reading and signing the school report), which is prototypical in PTM encounters, was carried out differently within different ecologies of action, with different repercussions in terms of participants' inclusion and exclusion.

Broadening the scope of the analytic gaze to include multiple modalities in DI research leads to challenges about how to describe the diverse semiotic resources contributing to the phenomena identified. A common analytic vocabulary is needed to characterize different resources consistently, in an attempt to break away from the tendency, which is still widespread in DI studies, to encompass all resources *other than talk* under the umbrella term '*non-verbal*', thus reinforcing a misleading dichotomy between language and other modes (Kendon 1972; Mondada 2014). An initial systematization of terminology could draw on Enfield's (2005) categorization of vocal-aural and visuo-spatial resources, with vocal-aural encompassing verbal and prosodic resources, and visuo-spatial including embodied resources and spatial arrangements. This latter category would enable the inclusion of semiotic resources that have not been accounted for systematically in DI studies, such as the material constraints of the physical environment in which the interaction takes place (e.g. seating arrangement, presence of artefacts).

Secondly, there needs to be recognition in DI research that interactionally-produced actions are laminated entities (Goodwin 2013) with "no principled priority of one type of resource over the others" (Mondada 2016: 341). However, in DI, embodied resources are often *de facto* treated as ancillary to talk rather than interrelated and interdependent resources having equal weight and forming complex multimodal *Gestalt* (Mondada 2014: 139-140) or *ensemble* (Kress, 2010).

Thirdly, broadening the analytical scope to resources other than talk also entails investigating *how* these multimodal resources work together. A productive concept in this respect is *temporality*, a term borrowed from CA (Depperman and Günthner 2015), and relates to when and how different resources appear in the unfolding of the specific sequence under investigation. Temporality can be taken as an umbrella concept encompassing *simultaneity* and *sequentiality*: (Mondada 2016; Streeck et al. 2011). This requires a fine-grained, micro-analytic approach to reveal the complex and subtle dynamics that would be discarded *a priori* if the same principle of "strict successivity" (Mondada 2016: 361) were adopted as is normally applied to analysis of talk.

Fourthly, the analytical 'spotlight' needs to shift from the interpreter only to how *all* parties-at-talk orient and adapt to 'previous and next' in interaction, not only verbally, but also in terms of embodied display of interaction. As argued by Lee (2016: 672), "*both* speakers and hearers accomplish socially organised participation not only through joint construction of talk but also through nonvocal conduct that is recognisably tied to the temporally unfolding interaction".

This shift in stance presents profound methodological and epistemological challenges, for example, identifying and isolating moments of interaction that are multimodally salient from an interpreting perspective, *what* semiotic resources contribute to a specific (course of) action and *how* they work together. Insights into a variety of interactional phenomena have already been provided by multimodal analysis of monolingual interaction: for instance, the understanding of turn-management system

developed by CA for monolingual interaction (Sacks et al. 1974), and subsequently applied to triadic encounters (e.g. see Davidson 2002, Metzger 1999), has been fundamental to capture the essence of the coordinating role played by the interpreter in interaction. Multimodal studies of monolingual interaction have gone further in shedding light on turn-taking phenomena such as turn allocation and self-selection [3] through a range of semiotic resources. Such insights can serve as a springboard to extend the exploration and understanding of coordinating actions to DI through multimodal lenses.

This paper advocates the need for a staged approach to multimodal investigation of DI data that integrates distinct levels of analysis. Such a model should present rich descriptive layers in the initial analytical phase, and this is the focus of the current paper. It is key that DI research reach a deeper level of description in order to ground the assessment of such practices in what *actually* happens in interaction and *how* it happens, rather than on an *a priori* assumption of what should (not) happen..

Video-recorded data enable DI researchers to broaden their analytic gaze to take account of embodied resources, which "run parallel to talk, showing how they are coordinated with it but not coinciding with its units" (Mondada 2016: 342), with clear methodological implications in terms of transcription. A multimodal turn in DI entails enriching the traditional Jeffersonian system with insights from research on monolingual multimodal data (e.g. Mondada 2007; Licoppe 2016). Furthermore, given the complexity of video-recorded interaction, multimodal investigation requires "a

[3] For a discussion of these phenomena in monolingual interaction see, among others, Lee (2016), Ford and Stikle (2012), Mondada (2007), Mortensen (2009), Streeck (2009), Streeck, Goodwin and Lebaron 2011

careful analytical focus" (Mondada 2016: 361), and the selection of salient moments of interaction from the broad scope of interactional patterns already identified in DI research. This paper extends multimodal exploration to a specific form of turn-taking in DI, and proposes relevant concepts to explore the interplay between different multimodal resources. Extracts from interpreter-mediated interaction are used to exemplify the different points raised.

Revisiting interactional concepts in DI: chunking as a case in point

By way of example, I bring multimodal exploration to a turn-regulating practice that is inherent in the interpreter task, also known as *chunking*. This terms refers to interpreters' ability to understand how and when to intervene during a multiunit turn to provide their rendition (self-selection) and to give the turn back to the speaker (turn allocation) while ensuring a smooth communicative flow. The concept of chunking was originally developed for monologic forms of interpreting, such as simultaneous interpreting (Ilg 1978; Setton 1999), where it refers to the "process whereby interpreters segment the input into smaller fragments that can be encoded without having to wait for the entire sentence to unfold" (Seeber 2011: 194). Given the dialogic nature of DI, chunking is extended here to identify a practice aiming to split either party's input into manageable processing units, mostly with a view to keeping the flow of the dialogue.

Chunking is a subtle activity often tacitly negotiated on a moment-by-moment basis through verbal, prosodic and embodied resources, such as gaze-shift and head

movement away from the speaker and/or gesture projecting self-selection and signalling when it is time for the speaker to stop. The interactional significance of gaze/head shifts has been investigated in monolingual research, where it has been found to be an explicit method of next speaker selection (e.g. Rossano, Brown and Levinson 2009; Schegloff 1996). Finding appropriate translation places (TRPs, Sacks, Schegloff and Jefferson 1974), i.e. moments projecting possible completion of the ongoing turn constructional unit (TCU) to insert a rendition, represents a challenge for interpreters wanting to initiate chunking. Decision-making as to when and how to intervene and give the floor back depends on multiple factors, including the nature of the turn-at-talk, prosodic and other multimodal features, and visual access to participants.

Initiating chunking in face-to-face interpreter-mediated interaction

To exemplify this point, I introduce two extracts where the initial phase of chunking, i.e. self-selection to take the floor, is negotiated entirely through embodied resources. Extracts 1 and 2 are taken from a corpus of approximately four hours of naturally-occurring, video-recorded interpreter-mediated PTMs (Author 2012), carried out face-to-face in Italian and English pedagogical settings. PTM1 was gathered in the UK and features a mother speaking Italian, two teachers speaking English and an interpreter; PTM2 was gathered in Italy and presents a mother speaking English, two teachers speaking Italian and an interpreter (Figure 2). The analytical focus is on the resources

displayed by the interpreter to regulate speakers' turn length and find a relevant TRP to take the floor and start their rendition.

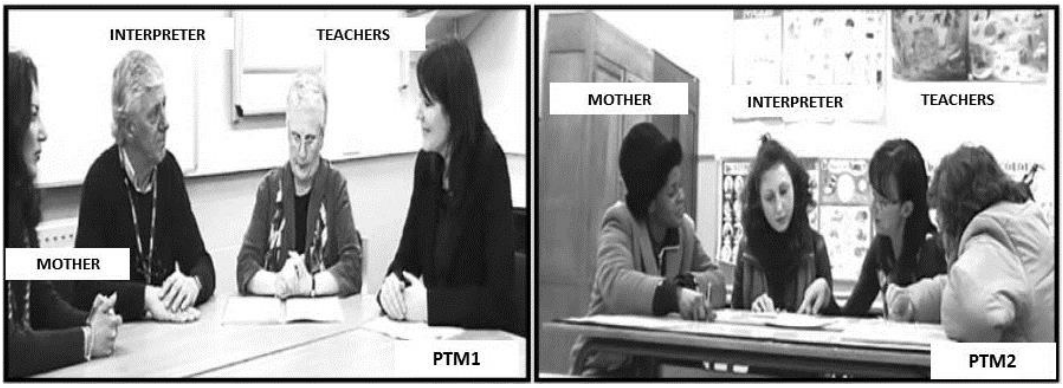


Figure 2: Interpreter-mediated PTMs

Figure 3 (PTM1) shows one teacher (T) assessing the child's performance in class – he has difficulties understanding and expressing himself in English, which is not his mother tongue. The transcript (see Appendix I) illustrates where and how chunking is initiated by the interpreter (INT).

```
1    T     so I believe     from speaking to Ax that (.) the understanding/
     comm  ((M and INT looking at T))
      t    +gaze at INT--> +T gaze at M ------------------> +T gaze at INT ----->

2    T     is there (.) *.h
      t    +gaze at M
→    int                 *head-turn towards M*
                         #fig3.1

3    INT   parlando con Ax la sua impressione è che lui capisce (.)
           talking to Ax     her impression       is that he understands
     comm  ((mutual gaze between INT and M is established))
→    int   *head-turn towards T – mutual gaze with T established*
            #fig3.2

4    T     but it's (.) obviously the language that's holding him back
      t    +gaze at M -------------> +T gaze at INT ------> +T gaze at M ----

5          he's not able to: (.) express himself (.) *as he would like\
      t    ------------------------------------------------->+
     comm  ((M and INT looking at T))
→    int                                                  *head-turn towards M*

6    INT   (XXX la lingua) non è in grado di esprimersi come vorrebbe
                 the language he is not able to express himself as he would like
```

Fig 3.1

Fig 3.2

Figure3: Extract 1 (PTM1)

In line 1, INT is gazing at T while the latter is talking; T alternates her gaze between M and INT. At the end of line 2, despite T's gaze being directed at M and no clear indication on her part to relinquish the turn, the micro-pause and in-breath seem to be treated by INT as a TRP marking the completion of a TCU and, therefore, offering a natural opportunity for INT to initiate his head movement and gaze-away shift towards M (Figure 3.1). Methodologically, it is important to point out that INT's movement

starts *before* his verbal rendition starts, which shows how embodied resources can be used to project INT's embodied self-selection before he has begun to speak.[4]

Chunking implies a second phase, i.e. that the turn is given back to the main speaker once rendition is completed. In Extract 1 line 3, INT shifts his head back to T after completing his rendition, mutual gaze is established between them after which T shifts her gaze to M and continues her turn. Turn-management is handled by INT entirely through embodied resources.

A similar pattern recurs when it comes to rendering the second part of T's turn (lines 4-5), where INT turns his head towards M to provide his rendition, this time in partial overlap with T's talk (…*as he would like*). In this case, from the point of view of content, a potential point for insertion of T's rendition could have been at the end of line 4. Nevertheless, a closer look at multimodal resources reveals that at that point T is looking at M and no micro-pauses or in-breaths are providing a natural break to the communicative flow. INT's head turn starts when a micropause is produced after another idea is expressed in full (i.e. that the child is not able to express himself); this is in partial overlap with an additional specification made by the teacher (*as he would like*, line 5), which is heard by INT despite the partial overlap and conveyed to M in the rendition (line 6).

---

[4] *Non-speech* turn-taking systems (i.e. how bodily actions can achieve turn-taking) have already been the object of scrutiny of multimodal research. Relevant to the practice highlighted here is Mondada (2007), which focused on pointing gestures predicting possible turn completions and projecting the emergence of possible next speakers and Ivarsson and Greiffenhagen (2015), which focused on how bodily actions can accomplish pre-beginnings.

A similar way of handling chunking is found in Extract 2 (Figure 4) from PTM2: the teacher is producing multi-unit turns to provide information on a music project. The same resources are displayed by INT to start her rendition, i.e. head-turn towards M in partial overlap with T (…*le informazioni*, line 6), which projects the beginning of the actual rendition starting at line 7, after mutual gaze with M is established.

1 **T** **poi finito questo progetto abbiamo un progetto di musica/ (.)**
then after this project we have a music project
int *gaze at T ---->
m $gaze at T ---->
2 **di:: che inizierà con un'insegnante esterna di musica (.) percui**
of which will start with an external music teacher so
3 **impareranno a:: usare alcuni piccoli strumenti\(.) $un progetto**
they will learn to use some small instruments a project
m --->$ gaze down--
4 **di musica al quale partecipano tutti i bambini\(.) attualmente-**
of music all children will take part in currently
5 **e: alla fine dell'anno ci sarà un piccolo saggio +comunque**
and at the end of the year they will perform all together anyway we
t +gaze at INT-->>
#fig4.1
6 **anche lì daremo *le informazioni**
will provide the information
→ int --> *head-turn towards M*
m -->>$
#fig4.2
7 **INT** **$.h $ (and) later there will be also a music project that will**
m $gaze up at INT$
start (.) where there will be a professional musician/ (.) a p- a professional music teacher/ (.) who will actually: teach the children how to play some instruments (.) (some small instruments)
8 **M** **$mh (.) *ok$**
m $nodding-->$
→ int *head-turn towards T*
#fig4.3

#fig4.1 #fig4.2 #fig4.3

Figure 4: Extract 2 (PTM2)

The nature of teachers' turns in Extracts 1 and 2 is different, with Extract 1 providing an evaluative assessment of the child, and Extract 2 providing a more factual account of school activities. Compared to Extract 1, INT in Extract 2 seems to be 'missing' a

number of potential TRPs, mostly marked in the transcript by a descending tone followed by a micropause (lines 3 and 4). Although not marked throughout the transcript, T keeps shifting gaze between M and INT while uttering her turn. INT gazes at T throughout the turn and starts turning her head to M at a point in line 6 where it is possible to predict how the teacher is going to complete her utterance after the verb "daremo" (*daremo le informazioni*, i.e. we will provide information, is a predictable collocation in Italian). At this point, INT has enough information to be able to provide a rendition, and gaze/head shift is chosen as a way to self-select for the floor.

As highlighted by Mondada (2007: 208), some embodied pre-beginning may be treated as having "an 'interruptive' potential or effect". The interpreter's chunking activity requires that an interruption to the main speakers' turn is made in cases where the latter does not chunk their own talk independently. Unless participants have experience working with interpreters, it is indeed very common that they lack awareness of the need to produce manageable chunks for the interpreter to render. It is therefore up to the latter to identify appropriate times to intervene, deliver the rendition and give the floor back, in the least disruptive possible manner. This requires sensitivity not only to what is being said, but also to what multimodal features are being used (in the extracts, for instance, we saw how micropauses or in-breaths can create natural breaks in the communicative flow), as well as awareness of the coordinating power of embodied resources (such as gaze shifts, head turns and gestures) in specific sequential points, and of the implications that this may have on the unfolding of the interaction.

To sum up, extracts 1 and 2 exemplify how interpreter-initiated turn regulation via embodied resources only can be highly effective for chunking in face-to-face mediated scenarios. Close analysis of the four-hour dataset revealed that chunking is mostly initiated by interpreters through a combination of the embodied resources of gaze shift and head movement initiated at possible TRPs, when there is enough information to a unit of meaning to enable them to provide a rendition. Although occasionally produced in partial overlap with teachers' talk, these moves enable the flow of communication to progress smoothly, with participants adjusting to them as the interaction unfolds.

Multimodal analysis of monolingual interaction has shown that the resources used in an ongoing action are heavily dependent on and shaped by their local material conditions, including the ecology of the activity (Mondada 2014). It can be argued that the efficacy of the embodied management of chunking exemplified in Extracts 1 and 2 may be partially ascribed to participants' seating arrangement and body positioning: in both extracts, the interactional space is triangular, with INT seated between the parties. Even when T is gazing at M, INT is included within the communicative radius and peripheral view, making it easier for participants to perceive slight body shifts and movements. A different seating arrangement or ecology of action could make the same move less or more conducive to turn regulation, as will be exemplified in the following section in relation to remote interpreting.

Initiating chunking in remote interpreter-mediated interaction

When interaction is mediated via technology, such as VMI, elements of system design also need to be taken into account in the analysis of interactional phenomena, including type of equipment (e.g. static or dynamic cameras; touch-screen), number and position of cameras and screens (e.g. camera-face distance and angle, implications of seating position and angle of all participants towards the cameras and screen), screen display (e.g. presence of multiple images, picture-in-picture), and screen size (which is gradually reducing as devices become increasingly mobile). These factors can support or hinder mutual visibility and access to embodied cues, thus influencing the way people orient to one another, how they come across through the technological medium and, ultimately, the construction of a rapport among interlocutors.

Building on the assumption that "some ecologies and types of activities might favour verbal resources along with gestures and body movements, whereas other ecologies and activities might favour distinctive and specific embodied resources over talk" (Mondada 2016: 341), a different participant constellation and ecology of action might pose restrictions on the type of resources that can be used effectively by interpreters for the purpose of chunking. This section focuses on a specific form of VMI, video remote interpreting (VRI), where the primary parties share the same interactional space while the interpreter is connected from a remote site via video-link and provides short, two-way consecutive interpreting (Braun 2015). Extract 3 (Figure 6)

has been selected purposefully to exemplify how environmental constraints can impact on chunking in DI during VRI.

The data presented in Extract 3 form part of a mediated lawyer-client consultation scenario collected during the SHIFT project[5]. The use of (partially) simulated interaction, where role-players act as professionals (e.g. police officers or doctors), and interpreters are invited to take part in semi-scripted simulations that are based on real-life audio recordings[6], is fairly common practice in DI studies mostly due to the sensitivity and confidentiality of visual data, which often require careful ethical negotiation with participants to gain permission to video-record (e.g. Wiles et al. 2008). This approach contradicts the ethnographic and ethnomethodological basis of CA, which is adamant on the use of recordings of naturally occurring data as the empirical basis of analysis (Heritage and Clayman 2013). Although controversial, semi-authentic roleplay scenarios have arguably been used successfully in DI firstly to overcome the difficulty of obtaining permission to video-record naturally-occurring data in sensitive settings; and secondly to help "to create a controlled environment which would support the isolation of relevant problems" (Braun 2013: 205). Extract 3, for instance, comes from a simulation where interpreters were not aware of the simulated nature of the

[5] SHIFT in orality (Shaping the Interpreters of the Future and of Today) 2015-2018 is an Eramus+ funded project on remote interpreting which aims to develop training solutions for remote interpreting in Higher Education and Lifelong Learning (2015-1-IT02-KA203-014786).

[6] This approach has been adopted, for instance, to compile the SimDik corpus of simulated interpreted doctor-patient interaction (Bührig, Kliche, Meyer and Pawlack 2012),the corpora of simulated remote interpreting developed within the framework of two EU-funded projects: AVIDICUS 1 and 2 (Assessment of Video-Mediated Interpreting in the Criminal Justice System) in relation to police interviews (Braun and Taylor 2012, 2014; Braun 2015; see http://wp.videoconference-interpreting.net/) and the simulations in business, administrative and healthcare contexts carried out within the framework of the ongoing SHIFT project (SHaping the Interpreters of the Future and of Today, http://www.shiftinorality.eu/), where Extract 3 is taken from.

encounter, and role-players were professionals in their respective fields who had been not been asked to follow a script, but to act naturally as they would in a similar scenario. It may therefore be argued that this data can still shed light on the interpreter's chunking behaviour in remote interpreting, particularly on how this practice seems to be sensitive to the spatial and visual ecology of actions.

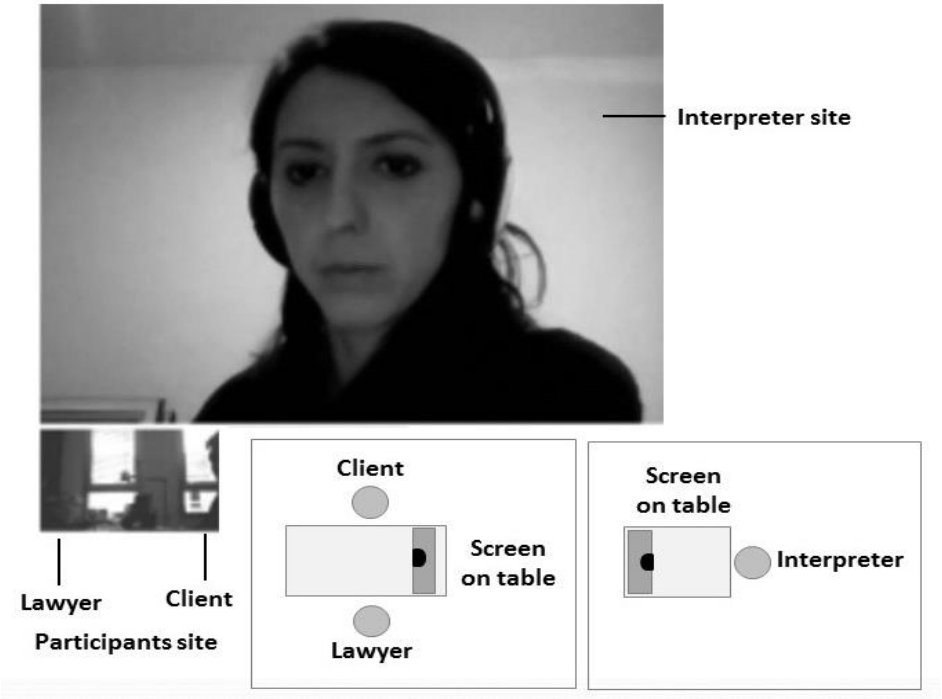


Figure 5: RI lawyer-client consultation (set-up for Extract 3)

The RI configuration of Extract 3 (Figure 5) poses constraints on the use of embodied resources to perform chunking. The primary parties (i.e. lawyer and client) are in the same physical location, and from the interpreter's perspective, they appear on the same screen, thus nullifying the need for full head turns to gaze at one party or the other, as minimal gaze shift is sufficient to fulfil this purpose. Furthermore, they are

only partially visible on screen and the lighting conditions are poor, which may affect the extent to which embodied resources can be used effectively to negotiate turn-taking. For the participants, the small size and positioning of the screen (tablet on the table) may impede them to fully capture the multimodal behaviour displayed by the interpreter. Hence it can be argued that the two-dimensional nature of the hardware restricts the visual field, reduces peripheral view and this limitation may desensitise the interlocutors to each other's physical conduct (Heath and Luff 1993).

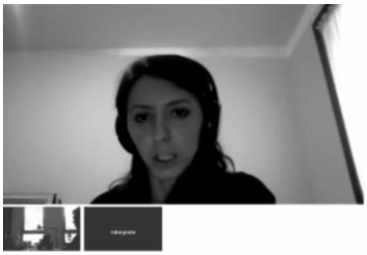
Extract 3 (Figure 6) captures the first part of a sequence where the lawyer is asking for clarification about the nationality and residence place of the client's wife. In terms of chunking, the extract highlights the problem faced by the interpreter to determine appropriate TRPs to start her rendition. Similarly to Extract 2, this task is made more complex by the length of the turns produced by the lawyer, and I argue that it is further exacerbated by the absence of a shared physical space.

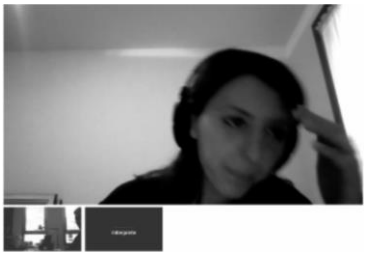
1 **LAW** **e: bhe intanto chieda se [eh al fine di capire quale procedimento:**
in the meantime ask if in order to understand what action
comm ((int gazing down at notepad))
2 **eventualmente: intraprendere .h eh: vorrei sapere se: diciamo eh**
we can possibly undertake I would like to know if let's say
3 **hanno già trovato un accordo (.) sulle eventuali condizioni .h eh**
they have already reached an agreement on the possible terms of
4 **di separazione/ oppure/ se sono in conflitto tra loro (.) .h**
the separation or if they are in conflict between themselves
5 **e poi la seconda [domanda se hanno figli perchè**
and then the second question if they have any children because
6 **INT** **[°ok°]**
→ comm ((nodding))
7 **LAW** **devo chiedere un po' varie: (.) varie cose insomma**
I need to ask a few various various things
8 **[se hanno] figli se han[no] del patri-**
if they have children if they have any assets
9 **INT** **[certo]**
→ *sure*
*gaze up at screen gazing at LAW-->
#fig6.1
10 **INT** **[ok]**
→ *leaning towards screen, gazing at client-->
#fig6.2
11 **LAW** **mh (.) comunque risiede in Italia risiede in I[ta]lia la signora/**
anyway does she have her residence in Italy in Italy the lady
int *leaning backwards gazing at LAW-->
#fig6.3
12 **INT** **[ok]**
→ *gazing at client-->
#fig6.4
13 **LAW** **ho capito bene/ (.) risie[de] in Italia la: la mo[glie]/**
I understood it correctly does she have her residence in Italy the wife
int *gazing down at notepad --->
#fig6.5
14 **INT** **[ok]**
15 **INT** **[eh so] ha detto**
she said
16 **che che è in Italia[na]**
that that she is in Italy/an
17 **LAW** **[che è in] Italia (.) [è italiana]/**
that she is in Italy she is Italian
18 **INT** **[adesso glielo chie]do(.)**
→ *now I'll ask*

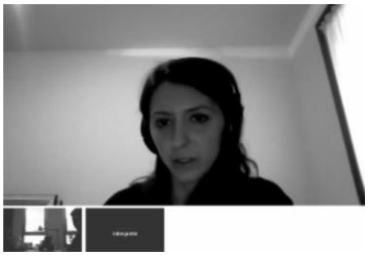
19 **faccio questa domanda**
→ *I'll ask this question*
20 **LAW** **però è anche in Ita[lia risie]de/**
but she is also in Italy she has her residence
21 **INT** **[ok the lawyer]**
*gazing at client-->
#fig6.6
22 **(1.7)**
int *gazing at LAW-->
#fig6.7

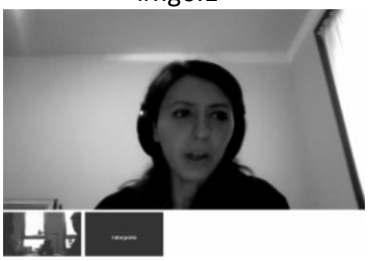
23 **INT** **scusi/**
sorry
*gazing at LAW-->>
#fig6.8
24 **LAW** **no no chiedevo se è anche qui in Italia ris- ha la residenza anche**
no no I was asking if also here in Italy res- has a residence also
25 **in Italia cioè .h oltre a essere italiana (.) in questo momento**
in Italy I mean beyond being Italian in this moment
comm ((INT nodding but continuing to gaze up at LAW screen))
26 **(.) è in Italia/**
she is in Italy
27 **INT** **ho capito**
→ *I understand*
comm ((INT nodding))
#fig9
28 **LAW** **mh**
29 **INT** **certo (.) certo adesso glielo chiedo subito**
→ *sure sure now I will ask him immediately*

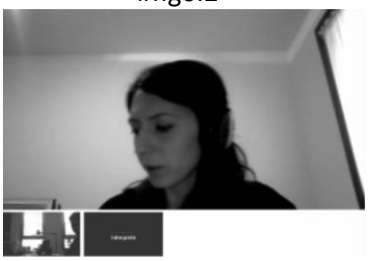
Figure 6: Extract 3

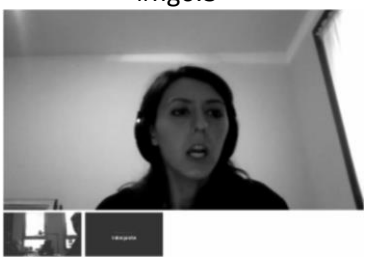

#fig6.1

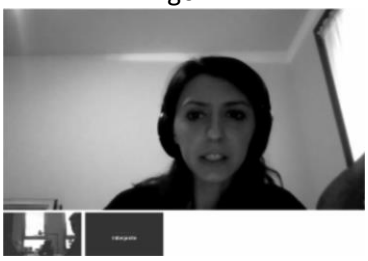

#fig6.2

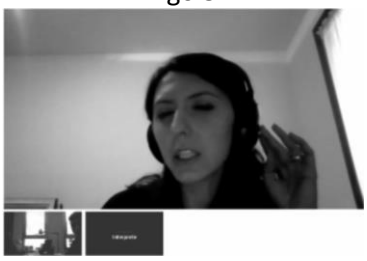

#fig6.3

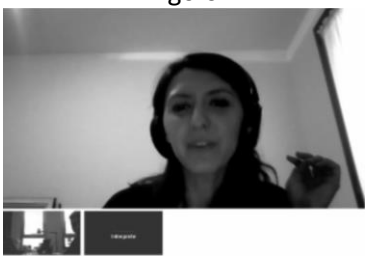

#fig6.4


#fig6.5


#fig6.6


#fig6.7


#fig6.8


#fig6.9

Extract 3 shows several acknowledgment tokens produced by INT (lines 6/9/10/12): a multimodal perspective on these micro-interventions on the part of INT shows that while the tokens produced at line 6 in overlap with LAW (uttered in a low tone of voice while gazing down at the notepad) signal understanding only, from line 10 these devices seem to be used also with a view to chunking the turn and self-selecting for the floor. At line 9, *certo* produced in partial overlap is paired up with gaze away from the notepad and up to the screen, towards the lawyer direction; at line 10, *ok* is paired up with a change in gaze direction and leaning of the upper body towards the left

part of the screen, where the client is displayed. Despite these embodied activities, LAW continues talking, and the acknowledgment token at line 12 (*ok*) is yet another attempt on the part of INT to take the floor as it is produced in overlap and in conjunction with gaze up at the client on screen. At this point we notice a change in the range of resources used by INT to try and intervene: at lines 18-19, INT makes her intention to interrupt the turn explicit through verbal resources (*adesso glielo chiedo – faccio questa domanda*). Turn transition, however, only happens at line 28 where INT adopts the same strategy of verbalising her attempt to take the floor (*certo certo adesso glielo chiedo subito*). From a multimodal perspective, however, different 'preparatory' groundwork has been undertaken: from line 24, INT keeps her gaze up at LAW on screen (instead of mostly looking down at her notepad), at line 26 she clearly expresses her understanding while nodding (*ho capito*) and only then signals her willingness to take the floor. This sort of preparatory phase seems necessary for INT to achieve her goal and provides some evidence of adaptive skills developed 'on the go' and within contextual constraints.

In contrast to the face-to-face interaction in Extracts 1 and 2, INT's attempt to initiate chunking through embodied resources (e.g. via nodding or gazing up from the notes or changing gaze direction) paired up with minimal verbal resources (e.g. acknowledgment tokens) do not lead to smooth turn transition in this case. This resonates with Braun and Taylor (2012), who showed that in remote interpreting data the use of embodied resources for regulating turn-taking was less impactful than in face-

to-face encounters. This may be ascribed to a variety of reasons, such as network problems causing delays and lack of acoustic clarity (as evidenced at lines 22-23 by the silence followed by *scusi*) and the lawyer's tendency to over-elaborate, which might be partially explained as an attempt to compensate for the communication difficulties caused by the remoteness of the situation. However, such a difficult unfolding may also be partially attributed to a 'latent uncertainty' about what the other party can see (Braun 2015): the two-dimensionality of the screen combined with access to the speakers' profile and poor lighting conditions hinder mutual access to embodied features.

Analysis of this data set suggests that the regulatory potential of embodied conduct may be at least partially undermined by the material constraints of the ecology of action in which the participants were operating, requiring the interpreter to adapt her strategies to perform chunking and other recurring interpreting practices. This seems therefore a fertile area for further exploration through complementary methods (such as eye tracking technique combined with introspective methods) and triangulation with information about participants' visual access and impact of the affordances of diverse technologies on how DI is managed.

Conclusions

This paper has argued that a turn to multimodal analysis is essential for a more comprehensive understanding of DI interaction. Following the embodied (Nevile 2015) or visual turn (Mondada 2016) in monolingual communication studies, it seems timely

to call for DI studies to develop a more holistic approach to the study of interpreter-mediated interaction. The goal is not to discover universally valid patterns of multimodal behaviour in DI, but to gain more in-depth knowledge of the dynamics and variability of specific practices through which mediated interaction is produced. As multimodal resources have been found to play a key role in "monitoring the ongoing interaction, displaying engagement in the activities performed and reorienting the participant's constellation" (Pasquandrea 2012: 150), this paper proposes that DI research needs to embrace a multimodal approach more systematically to delve into the complexity and nuances of interpreting processes in face-to-face and video-mediated environments.

The brief literature review provided in this paper indicates a need to develop a more consistent, yet flexible, approach for multimodal analysis of DI and the analysis of data extracts presented point to the need for progressive stages of analysis. Furthermore, there is also an urgent need to integrate the observation of videos with analysis of the verbal dimension of interaction with a view to capturing the precise "trajectories, temporalities and qualities of these multiple resources" (Mondada 2016: 361; see also Mondada 2008). Describing (ir)regularities and patterns in the multimodal formatting of actions more systematically would contribute to making research across different settings more comparable as well as gaining more nuanced insights into how specific DI phenomena unfold. This could facilitate the creation of *collections* (Mondada 2005) of

similar phenomena across different settings, showing the variety of ways in which DI is negotiated through multiple modes and through different technologies.

In this paper I have argued that a multimodal approach to DI has the potential to provide innovative, in-depth and comprehensive insights into interpreting practice, to recognise and scrutinise its complexity while raising awareness of the multiple semiotic choices and restraints that interpreters experience when negotiating specific interactional phenomena and sites. Multimodal analysis can therefore provide empirical evidence of how complex DI interaction is, which is often underestimated. This is particularly relevant in VMI, where some of the hyperbolic and business-(rather than research-)led claims about the affordances of such environments promoted by the current 'on demand' culture can be questioned by rigorous research of this kind. The difficulty of implementing chunking in VMI discussed in this paper is just one small but indicative example of the complexity of DI, and signals the need for systematic multimodal analysis to identify challenges and potential solutions.

To this end, a sound methodological framework is required, and some 'explorations' have been presented in this paper. Given the cross-linguistic and cross-cultural nature of DI encounters, and the complex interactional and cognitive task of interpreting, a multimodal methodology needs to be inherently interdisciplinary and bring together different tools and concepts to be woven into a coherent, multidimensional model that works for DI. In bringing a multimodal lens to CA, this paper has focused on the first layer of a potential model, arguing the need for a robust

descriptive bedrock to ultimately build a more complete understanding of DI practices. This will, in turn, need to be further complemented with additional analytical layers. To conclude, multimodality presents DI with the timely and unique opportunity for theoretical reconceptualization and for the development of new frameworks that can raise awareness of the multiple factors that influence decision-making for interpreters in action. Through DI, multimodal analysis can extend its remit to interlingual and intercultural communication, and benefit from application to increasingly complex forms of interaction that unfold across diverse technologies.

Appendix 1: Multimodal transcription conventions

Conventions draw on Jefferson (2004) and Mondada (2007)

(1.5)	silence expressed in seconds
(.)	micropause of less than 1 second
°XXX°	segment produced very softly
:	sound elongation
-	sound cut-off
XXX/	ascending tone
XXX\	descending tone
[XXX]	square brackets mark the beginning and end of a turn that overlaps with a preceding turn
()	uncertain transcription of poorly audible talk
**	delimitate INT's actions description
++	delimitate T's actions description
$$	delimitate M's actions description
*-->	action described continues across subsequent lines
*-->>	action described continues until and after extract's end
-->*	action described continues until the same symbol is reached

>>-	action described begins before extract's beginning
int/t/m	lower case for participant whose gesture is identified when s(he) is not the speaker
fig	the exact point where a screenshot (figure) has been taken is indicated
#	with a specific sign showing its position within a turn at talk
comm	commentary on other participants' behaviours in double brackets

References

Anderson L (2012) Code-switching and coordination in interpreter-mediated interaction. In: Baraldi C. and Gavioli L (eds) *Coordinating Participation in Dialogue Interpreting*. Amsterdam: Benjamins, pp. 115-148.

Apfelbaum B (1998) Instruktionsdiskurse mit Dolmetscher-beteiligung. Aspekte der Turnkonstruktion und Turnzuweisung. In: Brock A and Hartung M. (eds), Neuere Entwicklungen in der Gesprächsforschung. Tübingen: Narr, pp. 11-36.

Davitti E (2012) Dialogue Interpreting as Intercultural Mediation: Integrating Talk and Gaze in the Analysis of Mediated Parent-Teacher Meetings, PhD thesis, University of Manchester

Davitti E (2013) Dialogue interpreting as intercultural mediation: Interpreter's use of upgrading moves in parent-teacher meetings. Interpreting 15(2): 168-199.

Davitti E (2015) Interpreter-mediated parent-teacher talk. In: Alatriste L et al. (eds). *Linking Discourse Studies to Professional Practice*, Clevedon: Multilingual Matters, 176-200

Davitti, E. (2016) Dialogue Interpreting as a multimodal activity in community setting. In: Crawford-Camiciottoli B and Bonsignori V (eds) Multimodality across communicative settings, discourse domains and genres Newcastle: Cambridge Scholars, pp. 116-143

Davitti E and Pasquandrea S. (2017) Embodied participation: What multimodal analysis can tell us about interpreter-mediated encounters in pedagogical settings. *Journal of Pragmatics* 107: 105-128.

Baraldi C (2012) Interpreting as dialogic mediation: The relevance of expansions. In: Baraldi C, Gavioli L (eds) *Coordinating Participation in Dialogue Interpreting*. Amsterdam: Benjamins, pp. 297-326.

Bezemer J and Mavers D (2011) Multimodal transcription as academic practice: a social semiotic perspective. *International Journal of Social Research Methodology* 14(3): 191-206.

Bot H (2005) Dialogue interpreting in mental health. Amsterdam: Rodopi.

Braun S (2016) What a micro-analytical investigation of additions and expansions in remote interpreting can tell us about interpreter's participation in a shared virtual space, *Journal of Pragmatics* 107: 165-177.

Braun S (2013) Keep your distance? Remote interpreting in legal proceedings: A critical assessment of a growing practice. *Interpreting* 15(2): 200-228.

Braun S (2015) Remote Interpreting. In: Mikkelson H and Jourdenais R (eds) *Handbook of Interpreting*. London: Routledge, pp. 352-367.

Braun S and Taylor J (eds) (2012) *Videoconference and remote interpreting in legal proceedings*. Cambridge: Intersentia.

Braun S and Taylor J (eds) (2014) *Advances in videoconferencing and interpreting in legal proceedings*. Cambridge: Intersentia

Bührig K, Kliche O, Meyer B and Pawlack B (2012) The corpus "Interpreting in Hospitals": Possible applications for research and communication training. In: Schmidt T and Wörner K (eds) *Multilingual Corpora and Multilingual Corpus Analysis*. Amsterdam: Benjamins, pp. 305-318.

Cowan K (2014) Multimodal transcription of video: examining interaction in Early Years classrooms. *Classroom Discourse* 5(1): 6-21.

Davidson B (2002) A model for the construction of conversational common ground in interpreterd discourse. *Journal of Pragmatics* 34(9): 1273-1300.

Deppermann, A and Günthner S (eds.) (2015) Temporality in Interaction. Amsterdam: Benjamins.

Dicks B, Flewitt R, Lancaster L and Pahl K (2011) Multimodality and ethnography: working at the intersection. Introduction: *Qualitative Research Special Issue: Multimodality and ethnography: working at the intersection* 11(3):227-237.

Enfield NJ (2005) The body as a cognitive artefact in kinship representations: Hand gesture diagrams by speakers of Lao. *Current Anthropology* 46(1): 1–26.

Gavioli L (2012) Minimal responses in interpreter-mediated medical talk. In: Baraldi C, Gavioli L (eds) *Coordinating Participation in Dialogue Interpreting*. Amsterdam: Benjamins, pp. 201-228.

Gile D (1999) Doorstep Interdisciplinarity in Conference Interpreting Research. InL Álvarez L, Fernández Ocampo A (eds) *Anovar/anosar. Estudios de traducción e interpretación*. Servicio de publicacións, Universidade de Vigo. I: 41-52.

Goodwin C (2013) The co-operative, transformative organization of human action and knowledge. *Journal of Pragmatics* 46 (1): 8-23.

Heath CC and Luff P (1993) Disembodied conduct: Interactional asymmetries in video-mediated communication". In: Button G (ed.) *Working Order: Studies of Work, Interaction, and Technology*, London: Routledge, pp. 35–54.

Heritage J and Clayman S (2010) *Talk in action: Interactions, identities and institutions*. Malden: Wiley-Blackwell.

Ilg G. (1978) L'apprentissage de l'interprétation simultanée. *Parallèles* 1: 69-99.

Jacobsen B (2003) Pragmatics in court interpreting: Additions. In Brunette LG, Bastin G, Hemlin I and Clarke H (eds) *The Critical Link 3. Interpreters in the Community*. Amsterdam: Benjamins, pp. 223-238.

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

Jewitt C (2013) Multimodal methods for researching digital technologies. In: Price S and Jewitt C, (eds) *The SAGE handbook of digital technology research*. Los Angeles: SAGE, pp. 250-265.

Jewitt C (2014) (Eds) *The Routledge handbook of multimodal analysis*. London: Routledge.

Kendon A (1972) Review of the book Kinesics and context by Ray Birdwhistell. *American Journal of Psychology* 85: 441-455.

Kress, G (2010) *Multimodality: A social semiotic approach to contemporary communication.* London: Routledge.

Krystallidou D (2012) On mediating agents' moves and how they might affect patient-centredness in mediated medical consultations. *Linguistica Antverpiensia, New series – themes in Translation Studies* 11: 75-93.

Krystallidou D (2014) Gaze and body orientation as an apparatus for patient inclusion into/exclusion from a patient-centred framework of communication. *The Interpreter and Translator Trainer* 8(3): 399-417.

Krystallidou D (2016) Investigating the interpreter's role(s): the A.R.T. framework. *Interpreting. An International Journal of Research and Practice in Interpreting* 18(2): 172-197.

Lang R (1976) Interpreters in local courts in Papua New Guines. In: Barr WMO and Barr JFO (eds) *Language and Politics*. The Hague: Mouton, pp. 327-365.

Lang R (1978) Behavioural aspects of liaison interpreters in Papua New Guinea: Some preliminary observations. In: Gerver D, Sinaiko WH (eds) *Language, Interpretation and Communication.* New York: Plenum Press, pp. 231 -244.

Licoppe C (2015) Video communication and "camera actions". *Journal of Pragmatics* 76: 117–134.

Licoppe C and Veyrier C-A (2017) How to show an interpreter on screen? The normative organization of visual ecologies in multilingual courtrooms with video links, *Journal of Pragmatics* 107: 47-164.

Mason I (1999) Introduction. In: Mason I (ed.) *Dialogue Interpreting*, Special Issue of *The Translator* 5(2), Manchester: St. Jerome, pp. 147-160.

Mason, I (2000) Models and methods in dialogue interpreting research. In: Olohan, M. (ed) *Intercultural faultlines*. Manchesterç St. Jerome, pp. 215-232.

Mason I (2001) *Triadic exchanges. Studies in dialogue interpreting.* Manchester: St Jerome.

Mason I (2009) Role, positioning and discourse in face-to-face interpreting. In: de Pedro Ricoy R, Perez IA and Wilson CWA (eds) *Interpreting and translation in public service settings: Policy, practice, pedagogy*. Manchester: St Jerome, pp. 52-73.

Mason I (2012) Gaze, positioning and identity in interpreter-mediated dialogues. In: Baraldi C and Gavioli L (eds) *Coordinating Participation in Dialogue Interpreting*. Amsterdam: Benjamins, pp. 177-200.

Mason I and Ren W (2012) Power in face-to-face interpreting events. *Translation and Interpreting Studies* 7(2): 233-252.

Mondada L (2007) Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers. *Discourse Studies* 9 (2), 195–226.

Mondada, L (2008) Using video for a sequential and multimodal analysis of social interaction: Videotaping institutional telephone calls. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 9(3): Art. 39, http://nbn-resolving.de/urn:nbn:de:0114-fqs0803390.

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

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

Monteoliva Garcia (2017) *Police Interviews with Suspects. The collaborative construction of the stand-by mode of interpreting in police interviews with suspects.* PhD thesis, Heriot-Watt University, Scotland

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

Norris S (2004) *Analyzing Multimodal Interaction: A Methodological Framework.* London: Routledge.

Norris S (2009) Modal density, modal configurations: multimodal actions, In Jewitt C (ed.) *Routledge Handbook of Multimodal Analysis*. London: Routledge, pp. 78-90.

Norris S (2016) Concepts in multimodal discourse analysis with examples from videoconferencing. *Yearbook of the Poznań linguistic Meeting* 2(1): 141-165.

Pasquandrea S (2011) Managing multiple actions through multimodality: Doctors' involvement in interpreter-mediated interactions. *Language in Society* 40, 455–481.

Pasquandrea s (2012) Co-constructing dyadic sequences in healthcare interpreting: a multimodal account. *New Voices in Translation Studies* 8: 132 157.

Plejert C and Majlesi AR (2016) Embodiment in tests of cognitive functioning: a study of an interpreter-mediated dementia evaluation. *Dementia* 0(0): 1-26

Pöchhacker F (2004) *Introducing Interpreting Studies*. London: Routledge.

Poyatos F (1997) *Nonverbal communication and translation: New perspectives and challenges in literature, interpretation and the media*. Amsterdam: Benjamins.

Sacks H, Schegloff EA, Jefferson G (1974) A simplest systematics for the organization of turn-taking for conversation. *Language (Baltim)* 50: 696–735.

Seeber KG (2011) Cognitive load in simultaneous interpreting. Existing theories – new models. *Interpreting* 13(2): 176-204.

Setton, R (1999) *Simultaneous interpretation: a cognitive-pragmatic approach.* Amsterdam: Benjamins.

Sidnell, J (2009) Conversation Analysis: comparative perspectives. Cambridge: Cambridge University Press.

Straniero Sergio F (2012) "You are not too funny". Challenging the role of the interpreter on Italian talkshows. In: Baraldi C. and Gavioli L (eds) *Coordinating Participation in Dialogue Interpreting.* Amsterdam: Benjamins, pp.71-98.

Ticca AC (2008). L'interprete ad hoc nel dialogo mediato medico-paziente: Processi interazionali in una clinica dello Yucata`n indigeno. University of Pisa dissertation, Pisa.

Ticca AC (2010) Configurazioni diadiche nel dialogo mediato medico-paziente. In: Carreras i Goicochea, M., Pérez Vázquez, ME (eds), La mediación lingüística y cultural y su didáctica: Un nuevo reto para la Universidad. Bononia, Bologna, pp. 102-119.

Ticca AC and Traverso V (2016) Participation in bilingual interactions: Translating, interpreting and mediating documents in a French social centre*Journal of Pragmatics* 107: 129-146

Vranjes J (forthcoming) *Eye-tracking in interpreter-mediated dialogue*. PhD thesis, KU Leuven, Belgium

Vranjes J., Bot H., Feyaerts K., Brône G. (2017). Displaying recipiency in an interpreter-mediated dialogue: an eye-tracking study (Accepted). In: Oben B., Brône

G. (Eds.), *Eye-tracking in interaction. Studies on the role of eye gaze in dialogue*. Amsterdam: Benjamins
Wadensjö C (1998) *Interpreting as interaction*. London: Routledge.

Wadensjö C (1999) Telephone interpreting and the synchronization of talk in social interaction. *The Translator* 5(2): 247-264.

Wadensjö C (2001) Interpreting in crisis: The interpreter's position in therapeutic encounters. In: Mason I (ed) *Triadic Exchanges: Studies in Dialogue Interpreting*. Manchester: St. Jerome, pp. 71-85.

Wadensjö' C (2004) Dialogue Interpreting: A monologising practice in a dialogically organized world. *Target* 16(1): 105-124.

Wiles R, Prosser J, Bagnoli A, Clark A, Davies K, Holland S, Renold E (2008) Visual Ethics: Ethical issues in visual research, *ESRC National Centre for Research Methods Review Paper*, http://eprints.ncrm.ac.uk/421/1/MethodsReviewPaperNCRM-011.pdf