Embodied participation: What multimodal analysis can tell us about interpreter-mediated encounters in pedagogical settings

Elena Davitti, Centre for Translation Studies, University of Surrey, UK
Sergio Pasquandrea, Independent researcher, Italy

ABSTRACT: In the last two decades, Dialogue Interpreting (DI) has been studied extensively through the lenses of discourse analysis and conversation analysis. As a result, DI has been recognised as an interactional communicative event, in which all the participants jointly and actively collaborate. Nevertheless, most of these studies focused merely on the verbal level of interaction, whereas its multimodal dimension has not received much attention so far, and the literature on this subject is still scarce and dispersed. By analysing and comparing two sequences, taken from a corpus of face-to-face interpreter-mediated encounters in pedagogical settings, this study aims at showing how multimodal analysis can contribute to a deeper understanding of the interactional dynamics of DI. In particular, the paper shows how participants employ multimodal resources (gaze, gesture, body position, proxemics, object manipulation) to co-construct different participation frameworks throughout the encounters, and how the “ecology of action” (i.e., the relationships between the participants and the surrounding environment) influences the development of interaction.

KEYWORDS: conversation analysis, multimodality, participation framework, ecology of action, interpreting in pedagogical settings

Introduction and focus of the paper

Over the last few decades, the role of embodiment in human communication has been increasingly scrutinised, within a wide range of settings (see, among many others, Deppermann 2013; Goodwin 2000; Hazel et al. 2014a; Kendon 2004; Mondada 2014; Scollon & Scollon 2004; Stivers & Sidnell 2005; Streeck et al. 2011; Nevile et al. 2014a). Despite the different conceptual and methodological tools adopted (see section 3), it has become clear that the production of socially shared meaning needs to be situated within a multi-layered context, including not only human interactants and their verbal exchanges, but also the physical environment in which they operate and the wide range of bodily resources they use in order to communicate. In other words, “human interaction is fundamentally embodied, and as such any research into human social interaction is research into embodied interaction” (Hazel et al. 2014b: 3, italics in the original). As a consequence, the verbal side of interaction has been increasingly integrated with “concurrently relevant semiotic fields” (Goodwin, 2000a: 1499) such as gaze, gesture, posture, body and space orientation, and object manipulation, in an attempt to achieve a holistic model, capable of accounting for the complexity of naturally-occurring communicative events.

Building on these premises, the present paper explores the impact of embodied semiotic resources (especially gaze, object manipulation and body posture/orientation) on participation during
interpreter-mediated interaction (IMI). In particular, the paper contributes to this line of enquiry through a detailed single-case analysis (Schegloff 1987; Mondada 2012a) consisting of two sequences from a corpus of video-recorded face-to-face encounters in Italian and English pedagogical settings (parent-teacher meeting, henceforth PTMs). These have been selected on the basis of the fact that the same specific activity (i.e. reading and signing teachers’ reports) is performed in both sequences. This activity is prototypical of the type of communicative event analysed, and it involves the manipulation of an artefact (i.e. the school report). The paper describes how the different ‘multimodal formatting’ of the activity, i.e. the multimodal resources mobilised throughout its development in each sequence, generate different participation frameworks (Goffman, 1981), and entail changes of the interactional space (Mondada 2009) as well as reconfigurations of participants’ engagement that would escape a purely verbal analysis.

Our aim is to investigate participation as a situated, temporally unfolding process, embedded in its physical environment, actively negotiated and reconfigured by participants through embodied actions. Building on a micro-analytical approach largely based on Multimodal Conversation Analysis (Deppermann 2013; Hazel et al. 2014a), we will show how participation is the result of the concerted efforts of all parties-at-talk to articulate objects, gestures, gaze, body posture and talk within collaborative activities in context. The main research questions are: (a) How is participation negotiated and displayed multimodally in IMI?; (b) What interactional practices do participants (both interpreters and primary parties) implement in order to manage and coordinate participation?; (c) How do participants display, make recognisable and communicate to others their engagement/disengagement through their talk and embodied conduct?; (d) In particular, what is the role of objects (namely, written artefacts) in the negotiation of participation during the interaction?

2. Embodied actions and participation
Participation and embodiment have been treated extensively in the existing literature on monolingual interaction, both separately and in connection with each other. In order to set the background for the analysis, the following sections will provide an overview of how participation has been described as a situated achievement (2.1.), how it has been conceptualised and analysed within IMI, particularly in relation to multimodal semiotic resources (2.2.), and finally, how the ecology of objects (i.e. the peculiar positioning of objects and other physical resources within the environment, in relationship to the interactants; see Gibson 1979, Hindmarsh & Heath 2003; Nevile et al. 2014b) exerts an influence on the organisation of social interaction, which is the specific focus of this article (2.3).

2.1. Participation as a situated achievement
Participating in a conversation is one of the most common experiences in everyday social life. Yet, research has shown that it requires subtle, but complex, coordination among the participants, who need to adjust to and actively negotiate constantly shifting participation frameworks (Goffman 1981).
Such a dynamic conceptualisation of participation opens up the possibility for different encounters to entail different participation frameworks, depending on their goals, participants and settings. Similarly, a single encounter can switch through different participation formats, depending on its local contingencies. Participation is therefore a locally negotiated and co-constructed achievement, which can be shaped in a variety of ways (cf. Goodwin & Goodwin 2004), employing a wide range of multimodal resources. These resources play a crucial role in defining the different participation frameworks which can take place in a conversation, and in negotiating the shifts from one framework to another. For example, displays of engagement and recipiency can be conveyed through gaze (see Goodwin 1980; Heath 1982; Rossano 2012) and body position (Schegloff 1998); mutual orientation can be achieved through gesture (Goodwin 1986; Goodwin & Goodwin 1986) or through an interplay of gaze, gesture and posture (Mondada 2009, 2012a); turn-taking mechanisms are influenced by the multimodal behaviour of the interactants (Mondada 2007). More broadly, participating in a social interaction requires the interplay of different semiotic modalities (see Goodwin 2007; Kendon 1967, 1990; Heath 1986; Kidwell & Zimmerman 2007).

These interactional phenomena have received little consideration in the literature on IMI, where priority has been given to the vocal-aural modality over the visuospatial one (Enfield 2005) for the creation of interactional meaning. Yet, the inherently multiparty, multilingual and socio-culturally situated nature of IMI requires the mobilisation of multimodal resources on the part of all participants, interpreter included, to manage and negotiate complex participation dynamics.

### 2.2. Participation and coordination in IMI

Undoubtedly the presence of an interpreter in a communicative event modifies the very structure and flow of interaction. Since the development of the dialogic discourse-based interactionist paradigm in the late 1990s thanks to Wadensjö’s (1998) groundbreaking work, empirical studies on authentic IMI, particularly in community settings, have characterised the multitasking role of interpreters as one integrating translation, coordination and intercultural mediation (e.g. Angelelli 2004; Baraldi and Gavioli 2012; Davidson 2000; Gavioli 2009; Mason 2001; Roy 2000; Wadensjö 1998).

Such an active role of the interpreter may lead to shifts between two main participation frameworks, namely the triadic format (which is the one expected to be maintained throughout an encounter, according to most codes of conduct) and the dyadic format (which normally involves the interpreter and one of the primary parties, thus leading to the exclusion of the other). Given this fluid and dynamic nature of participation, coordination becomes an unavoidable facet of the interpreter’s activity, who is called upon to manage and negotiate constant shifts between these formats. Coordinating with a view to ensuring inclusion and participation requires not only linguistic, but multimodal monitoring and literacy by all parties-at-talk, interpreters in particular.

Much work still needs to be done in order to fully understand the nexus between participation,
coordination and multimodality in IMI. Literature on this subject is still scarce and fairly dispersed. Besides the pioneer study by Lang (1978) on gaze in courtroom interpreting and research on sign language interpreting, a practice which has always been considered through a multimodal lens due to the visual nature of sign language itself (e.g. import of body shifts, gaze and pointing as turn-management devices – see Metzger 1999; Roy 2000: Metzger et al. 2004), few studies on spoken-language interpreting have attempted to cross the verbal communication boundaries. Apfelbaum (1998) focuses on the rhythmic synchronisation of interpreter-mediated interaction, while Wadensjö (2001) examines the interpreters’ proxemics during joint narratives in psychotherapeutic sessions. Both studies conclude that “rhythmic regularity [...] is intimately connected with how the participants are positioned in relation to one another; that is, whether or not they share the same ‘communicative radius’” (Wadensjö 2001: 82-83). Davitti (2012, 2013, 2015) investigates the role of gaze and body orientation in interpreter-mediated parent-teacher meetings¹ as resources for triggering, modulating and disambiguating specific conversational moves. The gaze patterns identified in her datasets differed considerably from Mason’s (2012) study on interpreter-mediated asylum seekers’ interviews, and Bot’s (2005) study, which analyses the interplay between turn organisation and embodied resources (gaze and gestures) in therapeutic encounters. In the healthcare setting, Krystallidou (2014) and Pasquandrea (2011, 2012) shed light on the role of gaze and body posture in negotiating inclusion and exclusion, while Ticca (2010) takes into account the influence of non-verbal signals in triggering dyadic sequences during medical examinations mediated by ad-hoc interpreters. Although the peculiar patterns of behaviour identified may differ from one study to another, there is widespread consensus on the key role of bodily resources in enhancing or hindering participation.

Such literature, although promisingly growing, does not seem to rely on shared, systematic grounds for the analysis of multimodal resources. Most studies build on the conceptual and methodological underpinnings of Multimodal Conversation Analysis (e.g. Davitti 2012; Pasquandrea 2011, 2012; Davitti & Pasquandrea 2013, 2014; Ticca 2010), or Discourse Analysis (e.g. Wadensjö 2001), or otherwise adopt a mixed approach (e.g. Bot 2005, which integrates a fine-grained discourse analysis with interviews and concept maps, or Krystallidou 2014, which combines Conversation Analysis with Norris’ 2004 framework for multimodal analysis). The diversity of the approaches adopted, although providing a richness and variety of perspectives on data, makes a comparison between the findings a quite problematic issue.

Moreover, one common drawback in the current literature looking at IMI as a multimodal activity is that the embodied dimension often seems to be regarded as ancillary to talk, rather than integrated with it. The manipulation of artefacts in interaction has been particularly neglected in IMI, despite it being a fairly recurrent activity in these encounters.

¹ See also Davitti & Pasquandrea (2013) for a similar study based on a comparison between two different settings, i.e. pedagogical and medical.
2.3. Action and embodiment: the interactional ecology of objects

Objects can be crucial in human interaction, in that they are often not merely used as tools for performing social actions\(^2\), but rather may become constitutive of actions. Objects' affordances (Gibson, 1979)\(^3\) influence the way actions are performed, and are actively employed in turns by the interactants as means for performing actions. As such, they must be regarded as situated resources and as "emergent entities in fieri" (Mondada 2012b: 329; see also Hazel et al. 2014b).

Research shows that objects are used to perform a wide range of actions, within a variety of different activities, including instructing, co-constructing knowledge, ordering interaction, building temporally emergent action, embodying meaningful semiotic resources, claiming relevance, invoking collaboration, constituting shared purposes, etc. (see Hazel et al. 2014a for a thorough collection of studies on objects in interaction). It is therefore necessary to analyse their interactional ecology (as defined in section 2) in order to fully understand their role in human interaction.

In particular, written objects, which are the specific focus of this study, can exert an influence on the organisation of social interaction in a such diverse contexts as workplace meetings (Asmuß & Svennevig 2009; Deppermann et al. 2010), journalism (Hindmarsh & Heath, 2000; Weilenmann & Lymer 2014), archaeological fieldwork (Goodwin 2000a, 2003), medical visits (Frers 2009; Heath 1982; Jones 2009), academic supervision (Hazel & Mortensen 2014; Svinhufvud & Vehviläinen 2013), bureaucracy (Klein et al. 2014) and planning of professional activities, such as plumbing (Sakai et al. 2014) and architectural design work (Mondada 2012a). In all these contexts, a wide range of inscribed artefacts, such as notes, files, reports, IDs, forms, maps, sketches, drawings, plans, are actively employed and negotiated by the participants in order to perform meaningful social actions.

This paper will show how, within the context of PTMs, the way a particular object (i.e. the teacher's report) is handled, in combination with speech, gaze and posture, can be actively constructed as a sense-making device and exert a different influence on the unfolding interaction, thus leading to different shifts and reconfigurations of the participation framework. Section 3 will briefly outline the main methodological principles which will guide the analysis, while section 4 will provide more details about the actual selected sequences.

\(^2\) Conversation Analysis regards human interaction as constituted upon a series of actions (e.g. greetings, proposals, requests, invitations). Such actions are emergent, cooperative and socially shared, in that they require participants to constantly monitor each other's observable behaviour and display their own understanding of it. In this regard, any semiotic resource put in use by the interactants (talk, gesture, gaze, posture, object manipulation, etc.) can be constitutive of a social action.

\(^3\) Gibson first introduced the concept of "affordances", defining it as all "action possibilities" latent in the environment, independent of the individual's ability to recognize them, but always in relation to agents and dependent on their capabilities. Therefore a social actor may or may not perform certain social actions, according to the available resources in their interactional environment.
3. Methodological underpinning

As already pointed out in the introduction, various disciplines have developed different approaches to look at interaction through multimodal lenses: for instance, Conversation Analysis (Deppermann, 2013; Hazel et al. 2014), Social Semiotics (Hodge & Kress 1988; Kress & Van Leeuwen 1996, 2001), Mediated Discourse Analysis (Norris 2004; Norris & Jones 2005; Scollon 2001; Scollon & Scollon 2004), Critical Discourse Analysis (Blommaert 2013), Actor-Network Theory (Latour 1987, 2005). Such approaches differ not only in terms of the methodology employed, but also with regards to the very definition of what is meant by multimodality and what aspects of the context should be taken into account by the analyst.

The approach adopted in this paper largely builds upon some of the basic tenets of Multimodal Conversation Analysis, which broadens up the scope of ‘traditional’ Conversation Analysis to encompass a wider range of resources mobilised in interaction. The starting point for the analysis remains the same, i.e. bottom-up, empirical approach to the observation of data, with a view to collecting and comparing instances of naturally-occurring interaction and identifying recurrent patterns of social practices. The guiding principle is what Schegloff (1992a) calls demonstrable relevance, i.e. social actions (whatever semiotic resources interactants employ to perform them) must be demonstrably linked to their outcome, in terms of observable interactional dynamics.

Nevertheless, a multimodal framework provides the conceptual and methodological tools to take a more holistic approach to human interaction, where “linguistic and embodied resources [are treated] in principle in the same way, without prioritizing a priori one type of resource over other ones”, rather reconstructing in its entirety the interactants’ “complex multimodal Gestalt” (Mondada, 2014: 139-140). This is achieved through another fundamental principle, i.e. that human interaction must be regarded as lamination, i.e. “a set of different semiotic fields organized as layers of diverse resources”, each of which can be selectively decomposed, reused and transformed to build a next action, a proposal for how the future will be organized. Thus human beings build action by combining diverse resources (e.g., language structure, categories, prosody, postural configurations, the embodied displays of a hearer, tools, etc.) to perform both simultaneous and sequential transformative operations” (Goodwin 2013: 21).

As underlined by Mondada (2014: 138), such layers need to be regarded as integrated with each other, thus overcoming the dichotomy between “verbal” and “non-verbal” communication, which organizes human conduct and communication in a bipartite way, opposing language, on the one hand, to other types of conduct, on the other. […] Moreover, the negative particle in ‘non’-verbal seems to imply reference to a body resources which are not related to the verbal. As stated by Kendon, ‘it makes no sense to speak of ‘verbal communication’ and ‘nonverbal communication.’ ‘There is only communication’ (1972:443).
A further step can be made by recognising that social behaviour is very much dictated not only by the actions performed by human interactants (e.g. speech, gaze, posture), but also by the configuration of the settings and environments in which they operate, i.e. what was previously referred to as *(interactional) ecology of objects*. The data analysis will show how such aspects can be integrated in a multimodal investigation of interpreter-mediated interaction.

4. Dataset
The materials analysed in this paper are two sequences taken from a four-hour corpus of naturally-occurring mediated PTMs. To start with, this section will review some general features of this type of encounter, which has been mostly explored by studies on monolingual interaction. The second part of the section will focus on the two sequences specifically selected for the purpose of this study.

4.1 On interpreter-mediated PTMs
PTMs (both monolingual and mediated) are normally characterised as collaborative events during which participants are oriented towards fulfilling common goals and tasks in the interest of a third party who is often not present, i.e. the child. Despite their institutional nature, the "shape, form, trajectory, content or character" (Schegloff 1992b:111) of PTMs seem to be open to local negotiation among the participants, compared to other institutional contexts (e.g. Cuttance and Stokes 2000; Christenson and Sheridan 2001; Walker 1998). Similarly to doctor-patient encounters, parent-teacher interaction “may have an institutional mooring, but it also has an interactional bedrock” (Maynard 1991:486), i.e. all the parties-at-talk engage in actions that are systematically shaped and reshaped over the course of the talk. In other words, the default assumption regarding the collaborative and inclusive nature of these encounters is actually the result of the finely-tuned negotiations of embodied resources among all the parties on a moment-by-moment basis.

In terms of overall structural organisations, PTMs do not follow a rigid agenda, but they have been found to have a fairly ritualised and, to some extent, ceremonial nature (Badger 2007), and to proceed through some prototypical interactional phases, namely (Pillet-Shore 2001: 24):

Co-participants progress through a series of phases, starting with an introductory phase featuring social talk (Phase 1), which leads to the reasons for the conference (Phase 2), which in turn gradually yields to the phase in which the parent questions the teacher (Phase 3), followed by a future-oriented discussion of the student (Phase 4), and finally an orientation toward closing down the conference (Phase 5), which, once ratified by both co-participants, gives way to an exchange of thanks and farewells (Phase 6).

The professional activity analysed in this paper is a recurring practice across our data and a highly predictable one in PTMs in general, i.e. ‘reading and signing’ the record, which is commonly used “as a method of reviewing the student’s current class standing for and with parents, providing an
opportunity for both co-participants to be on the same page, armed with the information they will subsequently use to evaluate and discuss the student’s past, as well as strategize about the student’s future” (ibid.:19). This practice can be integrated at different moments in the overall structural organisation: in our data, it is a prelude to the closing down of the meeting. After discussing the overall achievements, performance, and attitude of each child in the classroom, as well as any potential problems or difficulties they may be confronted with, the school report is presented to the parents, who are requested to ‘read and sign’ it. By doing so, parents acknowledge the content of the report and formalise its acceptance whilst providing evidence that the meeting has taken place.

The actions accomplished in order to perform this practice revolve around an artefact, i.e. a school report, which is discussed and manipulated in interaction. Fulfilling the request to ‘read and sign’ the report in a mediated PTM is more complex than in a monolingual scenario. Such complexity is partly due to the materiality and affordances of the artefact itself (for instance the fact that it is written in a language not spoken by the parent) and partly to the way and sequential moment in which it is introduced in interaction. In particular, three intermediate steps are required, namely (a) the familiarisation with the content of the school report on the part of the interpreter, who has never seen the document before; (b) the rendition of its content into the parents’ language; (c) the explanation of the function and features of such document within the educational system of the host country. The way such actions are sequentially handled and embodied will have an impact on the unfolding of the interaction and on participation dynamics. The artefact therefore plays a key role in the event, in that it is the means whereby a necessary bureaucratic procedure is actively negotiated and eventually accomplished.

4.2 Selected sequences

The two selected sequences are taken from two distinct encounters taking place in two different Italian schools (geographically located in the same region). They both feature four participants, i.e. two teachers (T1/T2), one migrant mother (M) and one interpreter (INT). The two mothers are from India (sequence 1, Figure 1) and Nigeria (sequence 2, Figure 2): they have just moved to the host country (Italy) with their family and they have been invited to the meeting to discuss the performance of their children at school after one semester. The mothers are not proficient in Italian yet, so an interpreter has been called in to facilitate communication with them in English.
The two sequences are perfectly comparable in terms of:

- specific actions implemented, wider activity performed, general purpose and development of the encounter: the request to ‘read and sign’ the school report, which is lying at the centre of the table (as shown in Figures 1 and 2) is in both cases launched by one of the teachers. The ongoing activity revolves around the artefact, which becomes the “centre of collective attention” (Mondada 2007: 198). Similarly, neither the parents nor the interpreter had access to the report prior to the actual meeting;
- number of participants: four in both cases and featuring the same interpreter, who is qualified and regularly works in dialogic settings, but who has never worked with any of the parties before;
- duration of the sequence: over five minutes each, possibly due to the fact that the activity performed, however typical of PTMs, requires more interactional work in mediated encounters (as explained in 4.1);
- ecology of action: the encounter takes place in a classroom, with all parties seated around a table but with a different seating arrangement.

These features were captured via video-recordings obtained with two fixed cameras both enabling a perspective view of the scene from two slightly different angles (Figure 3) to make sure that the facial expression and body posture of all participants sitting around the table, as well as the way the object was manipulated, could be analysed on a moment-by-moment basis.

![Figure 3: Camera frontal and lateral views](image)

The presence of the same interpreter in both sequences may appear idiosyncratic, particularly given the very active role that she takes up in interaction that goes beyond merely interpreting verbal content uttered by the teachers. Despite this, fine-grained analysis of all participants' behaviour in relation to the “read and sign” activity in the two extracts shows contrasting ways of displaying and mobilising multimodal resources to manipulate the artefact and negotiate its function and meaning.
This seems to lead to several changes and reconfigurations of the participation framework and distribution of tasks and responsibilities and, ultimately, it contributes to framing the interaction as more or less collaborative.

5. Sequence 1: a case of constantly shifting participation framework
This section focuses on the analysis of sequence 1, in which the triadic format of the participation framework undergoes a number of reconfigurations before the activity is brought to completion (i.e. the content of the report is conveyed and the document is signed). The analysis illustrates how such shifts in the participation format are often constructed through the manipulation of the school report in combination with several other multimodal resources.

5.1 Introducing the artefact
Figure 4 shows the few seconds immediately preceding the start of sequence 1, i.e. the launch of the request on the part of T1. This is a moment of silent interaction where INT is still involved in a previous action (i.e. jotting down some notes for the mother) while both teachers are monitoring what INT is doing. A close multimodal look, however, shows that T1 is already projected towards the activity which is about to start (i.e. reading and signing the report), even before uttering the first turn. Her right arm is partially stretched towards the report, which is laying midway between her and the interpreter on the table. T1’s gesture is, however, momentarily suspended while waiting for the interpreter to complete the preceding course of action.

The second screenshot in Figure 4 shows that, as soon as INT signals completion of the preceding activity (which is visibly displayed by putting the pen down, setting her elbow on the table and
reorienting her gaze), T1 resumes her handover gesture by stretching the arm towards the report and sliding it towards INT, thus crossing the imaginary line separating her “personal space” (Hall 1966) from the interpreter’s one. While doing so, T1 self-selects\(^4\) also verbally (ecco, line 1, Figure 4) and utters her turn while addressing INT (line 2, Figure 5), thus officially opening sequence 1. As an immediate response to T1’s move, all participants (including M, who does not speak Italian) visibly reorient to the artefact, by repositioning their bodies and gaze towards the artefact itself, which is now being held and handled by INT.

![Figure 5: Sequence 1, part II](image)

In other words, the sequence starts when the artefact, which has been lying on the table since the beginning of the encounter, is made progressively relevant through a process of “multimodal focalization” (Klein et al. 2014: 231). The sequential placement of T1’s arm gesture, particularly its preparatory phase, shows a very close monitoring and understanding of INT’s ongoing action and is used to anticipate and project that a change of speakership is imminent. This shows that it “is not a mere individual isolated gesture, but a finely coordinated one, synchronised with the format of the turn in progress and adjusted to other’s actions” (Mondada, 2007: 202), thus confirming findings from monolingual interaction (Mondada 2007; Streeck 1995; Streeck and Hartge 1992). This is further strengthened in Figure 5 by verbal deixis/identification (questa è la parte finale della scheda, line 2), followed by an indirect request (se vuoi leggere, se vuole firmare, line 2) coupled with an embodied action (stretching arm to slide the document towards INT while gazing up). Through such moves, T1

\(^4\) Turn allocation, and the subtle mechanics which govern it, are traditionally one of the main areas of interest within Conversation Analysis studies (see, among many others, Sacks et al. 1974; Sacks 1992). Basically, two main turn allocation modalities exist: self-selection (i.e., when participants take the turn on their own initiative) and other-selection (i.e., when turn is allocated by another speaker). Both modalities can be conveyed via different semiotic resources, such as speech, gaze, gesture, body movement, etcetera (Lerner 2003; Mondada 2007; Rossano 2012).
steers the overall organisation by marking the beginning of the new sequence.

The formulation of T1’s turn is noteworthy, in that it ratifies INT as the addressee of the request, i.e. as the participant in charge of ensuring that the action is fulfilled. Evidence is provided by the use of the second person to address INT (*se vuoi leggere*) and of the third person singular to refer to M (*se vuole firmare*) coupled with embodied resources (lack of visual engagement with M and handing over gesture). This is multimodally enhanced by T1’s retracting movement into her own personal space after uttering the turn, which seems to emphasise the act of passing the baton of the interaction to INT, who is now in charge. In other words, the physical movements of the artefact, coupled with talk and gaze orientation, not only mark, but actively construct and materialise the transfer of interactional control to INT and, consequently, the shift in the participation format. The artefact therefore is not just a passive tool, but rather a structural component of the ongoing activity, which is influenced by the artefact’s affordances.

5.2 First split in the participation format

At this point in the sequential unfolding of the interaction (Figure 6), INT quickly glances at the artefact, then redirects her gaze to M (whose gaze is fixed on the artefact) and self-selects to produce an autonomous turn (lines 3-7) which deviates in terms of content from T1’s source utterance, thus representing an instance of non-rendition (Wadensjö 1998). In particular, the turn explains the function of the document for M, which is in line with the task of interpreters as intercultural mediators in IMI (see 2.2). After a short pause and hesitation (line 5) during which a minimal embodied uptake is performed by M (*ok* in a low tone of voice combined with slight nodding, line 6), INT utters *let’s see what they wrote* (line 7) while gradually shifting her gaze away from M and down to the artefact. This move accounts for INT’s temporary disengagement and prefaces the next action, i.e. familiarising with the content of the school report, which is fundamental to enable INT to then sight-translate the document for M. Both teachers monitor the sequence without intervening throughout this chunk.
INT’s move leads to a momentary suspension of the progressivity of the turn that is probably perceived by T1, who has been monitoring the ongoing interaction until this moment: as soon as INT shifts her gaze down to the artefact, T1 starts a parallel conversation with T2. To this end, as shown in Figure 7, she reorients her head and gaze towards her colleague, performing a “body torque” 5 (Schegloff 1998).

---

5 A “body torque” is a particular posture, in which “the divergent orientation of different body sectors” signals “engagement with multiple courses of action and interactional involvements, and differential ranking of those courses of action and involvements” (Schegloff 1998:536). In this case, T1’s arms and torso remain oriented towards INT and M (main activity), while her head and gaze are turned towards T2 (side activity).
The way in which the artefact is multimodally managed by INT (through verbal and embodied actions) is, once again, conducive to a change in the participation framework, which clearly splits into two units, thus providing further evidence of the fact that the artefact is not accessory to the event, but rather constitutive of it and constantly influencing the ongoing interaction. The triadic format of the interaction is therefore, at least momentarily, suspended, with the teachers showing disengagement and lack of monitoring of the main activity.

5.3 Multimodal attempt towards self-inclusion (teacher)

The participatory configuration changes again after approximately 9 seconds when, as shown in Figure 8, INT slightly leans towards the report, while maintaining her gaze oriented to the artefact, and utters some words (*convivenza democratica*, line 8) in an interrogative tone. This clearly signals a comprehension problem related to the understanding of the report during her familiarisation phase. As a result, the teachers suspend their parallel conversation but two different reactions are triggered: T2 responds only verbally to INT’s question (*si*, line 10), without engaging visually or reorienting her head. Conversely, T1 reorients her gaze towards INT without however being visually reciprocated. While continuing to gaze at INT, T1 utters a question (*c’è scritto così?*, line 9) complemented with gesture which engages the whole body (i.e. bending upper body, stretching arm towards the document being held by INT).

![Figure 8: Sequence 1, part IV](image)

T1’s gesture and posture can be interpreted as an offer to look at the document herself to clarify its content. These embodied moves reinforce T1’s attempt to re-integrate into the participation
framework, which was not successfully implemented via gaze reorientation alone. Such attempt mirrors T1’s gesture at the beginning of the sequence: while Figures 4 and 5 showed how T1 yielded interactional control to INT by passing over the report, in Figure 8 her gesture seems to be performed to regain such interactional control. In both cases, the gesture is oriented towards the artefact. Nevertheless, her attempt does not prove successful this time, given that INT continues to perform her familiarisation phase by gazing down at the report and reading its content aloud. By doing so, INT also retains control over the ongoing activity. As a result, T1’s offer of help is not pursued any further, but is reflexively adjusted to INT’s (lack of) response: T1 retracts her arm and body, falling back into her own side of the interactional space, without nevertheless engaging in any other parallel conversation.

5.4 Multimodal attempt towards self-inclusion (mother)

Figure 9 shows that, while INT continues to read out the report, M self-selects. The self-selection is achieved via an indexical gesture pointing to the report, combined with a question uttered in a broken Italian (scrivere in italiano, line 12); the multimodal formatting of the action is similar to the one used by T1 in Figure 8, line 9 (i.e. a pointing gesture coupled with talk and gaze reorientation). Nevertheless, unlike T1’s, M’s move succeeds in grabbing INT’s attention. This can be due to two main factors: on the one hand, the ecology of action, on the other hand, the broader goals of the interaction.

With regard to the ecology of action, M’s position next to INT gives her direct visual and physical access to the artefact, which is within reach; this enables her to touch the school report while pointing at it, thus attracting INT’s attention. INT interrupts her activity to answer M’s question and follows it up by signalling that she is ready to start the rendition of the report (lines 13-15). Consequently, INT proceeds to re-including M not only verbally (i.e. by responding to her question) but also multimodally (i.e. by interrupting her familiarisation process with the content of the report, reorienting the sheet and making it accessible to M). Once again, the artefact is used as a means through which participation undergoes some structural changes: in this case, in particular, it is used to reintegrate one of the parties, namely M, in the participation framework.

M’s question is treated by INT as a complaint, hinting at M’s scant competence in Italian, and at her consequent exclusion from the ongoing interaction. INT’s response (i.e. reassuring M that the content of the report will now be translated into English) is therefore in line with her role in the interaction, i.e. coordinating and ensuring a triadic framework.
Between lines 17-28 (Figure 10), the actual rendition of the report starts, with INT engaging in a sight-translation of the document which is interspersed with M’s acknowledgment tokens (okay, hm lines 21, 23, 28) signalling understanding despite being uttered in a very low tone of voice and without any visible embodied confirmation (e.g. nodding). Throughout this chunk, the teachers maintain a certain degree of monitoring of the interaction by gazing towards INT without engaging in any parallel activities.
5.5 Multimodal attempt towards other-inclusion (mother)

Another key point in the interaction is what follows INT’s turn signalling that her rendition has been completed (that’s it, figure 12 - line 29). At this moment M keeps her gaze down to the report, thus generating a 4.28s gap in the conversation (line 30): the lack of a clear uptake from her is dispreferred in conversation⁶ and needs to be repaired by other-selection. During this gap, T1 and T2 are oriented towards INT while INT is oriented towards M.

---

⁶ The terms “preferred” and “dispreferred” are to be considered technical conversation analytic terms used to characterise “sequential properties of turn and sequence construction, not participant desires or motivations” (Schegloff 1988:445, see also Heritage 1984). As explained by Heritage (1984:267), “actions which are characteristically performed straightforwardly and without delay are termed ‘preferred’ actions, while those which are delayed, qualified and accounted for are termed ‘dispreferred’.”
Mutual gaze between T1 and INT is not established; consequently, the triadic format is not restored (see Pasquandrea 2011, 2012 for an analysis of how lack of mutual alignment may lead to maintaining dyadic sequences in interpreter-mediated encounters). Instead, INT takes the floor at this point and makes the report entirely available to M while inviting her to read it independently (line 31). Once again, the artefact is used as a means to change the participation format and introduce a new activity. This is complemented by an offer to help in case of lack of understanding (if you don’t understand I can help you, line 34).
It is worth noting that, based on her partial understanding of English and monitoring of the interaction, T1 could also have self-selected to respond to M’s lack of response, considering that teachers are supposed to steer the development of these meetings. Nevertheless, this does not happen in sequence 1; conversely, teachers respond to INT’s autonomous turn (suggesting that M would be able to read the report herself) with a light giggle (lines 32-33), as suggested by the proximity of the two actions and the emphasis on *yourself* in INT’s turn (line 31). As a result, the strong coordinating role played by INT, which emerges through her autonomous handling of the artefact, seems to be tacitly endorsed by the teachers.

### 5.6 Further split in the participation format

A further split in the participation framework is produced at this point (Figure 13): INT and M start to collaboratively work towards clarifying the meaning of specific sentences in the report (lines 35-50, not reported here) with M becoming more involved by physically holding the artefact while teachers gradually self-exclude by engaging in another parallel conversation. The gradual decrease in monitoring on the part of the teachers, coupled with the starting of a concurrent activity linked to the artefact, leads again to the yielding of interactional control to INT, who continues to be tacitly legitimised as a coordinator and interactional pivot.

![Figure 13: Sequence 1, part VIII](image-url)
5.7 Closing the sequence

Another long gap in conversation occurs at the end of INT’s second attempt at rendering the content of the school report (line 51, Figure 14), during which M does not produce a clear uptake and does not engage multimodally with INT. Differently from the previous instance (Figure 12, line 30), though, teachers are not monitoring what is going on at this point, as they are completely absorbed in a concurrent activity. At lines 52-54, INT once again self-selects: her turn shows an alternation of English and Italian, seemingly using code-switching as a resource to reinclude the teachers into the participation framework while bringing the sequence to a close, without attempting to pursue a clear uptake on the part of M.

Similarly to part VII (Figure 12), while uttering the turn, INT does not visually engage with the teachers, who have in turn reoriented towards her, and answer the question through embodied deixis/identification, i.e. verbal explanation of where to sign (quassù, Figure 15 - line 55) combined with a gesture pointing to the exact part in the report. Despite the triadic format of the interaction being restored, and the purpose of the sequence being accomplished (i.e. signing the report, which eventually happens at line 57), no checking of M’s understanding of the content of the report is performed. While M is signing the report, T1 self-selects to round off the encounter and explicitly
bring it to a close (line 58). Once again, the artefact helps to re-establish contact among the parties, thus contributing to producing another shift in the participatory framework and, in this case, to bringing the meeting to a close.

Analysis of Sequence 1 has shown that the combination and sequential organisation of specific verbal and embodied resources on a moment-by-moment basis to handle the artefact is conducive to participation shifts and contributes to materialising the transfer of interactional control. In this specific case, interactional control is shifted from the teachers, who are supposedly in charge of the development of the event, to the interpreter. Section 6 will examine an instance of completely
different multimodal handling of the same activity, highlighting in particular its implications in terms of participation and distribution of tasks and responsibilities.

6. Sequence 2: a case of stable participation framework

Sequence 2 exhibits how, within the framework of the same activity, the multimodal handling of the same artefact leads to a much more stable participation framework, i.e. not subject to as many shifts and reconfigurations as in sequence 1. For reasons of space, the sequence will be presented in a contrastive manner, mostly to point out similarities and differences in terms of multimodal practices and impact on participation dynamics.

Similarly to sequence 1, the artefact is introduced at the beginning of the sequence by T1, who makes it relevant through verbal deixis/identification (allora guarda questa è la scheda..., Figure 16 - line 1) and embodied action, i.e. by sliding it towards the centre of the table and reorienting it so that everyone (INT in particular) has access to its content. This is a recurrent practice in the data: in both sequences, the artefact is introduced by one of the teachers, and multimodal resources are mobilised to make it accessible to all parties while explaining its function.

Also similarly to sequence 1, talk is addressed to INT, as evidenced by the use of pronouns (poi lo leggi..., Figure 17 - line 8). Nevertheless, differently from sequence 1, T1 maintains her control over the artefact: the explanation of the structure and features of the document is entirely taken up by T1,
and complemented with a pointing gesture to each relevant part (cf. screenshots on Figures 17 and 18a/b), which are conducive towards channelling the participants’ attention towards the artefact.

**Figure 17: Sequence 2, part II**

T1’s active role relieves INT from the task of producing an autonomous explanation of the document, without nevertheless preventing her from fulfilling her role as intercultural mediator and as coordinator. Regarding the former, at lines 9-14 (Figure 18a), INT produces an autonomous expansion to provide M with background information necessary to make sense of T1’s turns (in particular, INT explains to M how the Italian marking system works). This is performed without INT taking the lead over the interaction (which is retained by T1). Nevertheless, the artefact remains the main point of collective attention throughout INT’s rendition too, which maintains not only the verbal component of T1’s original utterances, but is also accompanied by gestures pointing to the relevant parts in the report, thus ‘replaying’ the source turn in all its multimodal facets. Turn-taking is also handled multimodally, with change of speakership being strongly projected by INT’s redirection of gaze towards T1 while uttering the acknowledgment token *ok* (line 21), thus indicating that the latter can resume the turn.
The remainder of the sequence is characterised by T1 alternating the reading of excerpts from the report with her own comments; this verbal alternation is complemented by gaze shifts from the report (when reading) to M and INT (when expanding), thus creating a sense of inclusion of all parties throughout the sequence. INT’s renditions are seamlessly inserted at relevant transition spaces (e.g. Figure 20, line 28; Figure 21, line 43), without disrupting the flow of the conversation and ultimately providing for a more stable participation framework.
This is shown by the sequence of screenshots below, which exemplifies the recurrent multimodal handling of the sequence by T1.

Figure 19: Screenshots showing the multimodal handling of sequence 2

The way interaction is handled in sequence 2 presents a marked difference when compared to sequence 1: instead of leaving INT with the task of familiarising herself with the content of the report, in sequence 2 the actual content is read out by the teacher herself and interspersed with her own expansions to clarify the mere written content (introduced, for instance, by *cioè*, line 24). One downside of this practice is that it often results in long chunks which may be difficult for INT to manage, as shown by the tendency to provide a summarised rendition, where the information is often reordered and/or explicitated with respect to the source turn uttered by T1 (e.g. Figure 19, lines 28-33). This way of managing turns on the part of T1 also causes an increase in the multitasking activity of INT, who has to sight-translate chunks of text while keeping track of the additional information produced by T1.
Figures 21 and 23 represent the final part of the sequence, where the same pattern described in this section can be found, i.e. fairly long chunk produced by T1 (alternating reading aloud from the report with personal remarks) followed by INT’s rendition. The overlap at lines 41-42 between T1 and T2 shows an instance of other-completion of the turn, indicating that T2, although mostly silent, is actually monitoring what T1 is saying, thus remaining engaged in the interaction.
Another point of difference with respect to sequence 1 is that the teachers monitor the interaction during INT’s rendition as well; there are no instances of disengagement on their part, and the participatory framework maintains a triadic configuration throughout the sequence. For instance, line 50 (Figure 21) shows a short gap in the interaction, during which INT is looking at the report, possibly retrieving what to say next; throughout this long rendition, the teachers do not engage in any parallel activity or conversation, but keep (silently) monitoring the interaction, as per Figure 22 below.
Figure 23 shows what follows this short gap (line 50): INT resumes her rendition of T1’s turn, while splitting her gaze between M and the report, thus providing multimodal evidence of her multitasking activity (sight translating the relevant chunk of the report while integrating it with T1’s additional information). One instance of autonomous, coordinating intervention on INT’s part is found in lines 62-63, where INT autonomously asks M whether she understood everything or had any questions. This is followed by a brief remark on the part of M, who thanks the teachers for what they have done for her child (not reported here), and by the final signature of the report, which signals the accomplishment of the main goal of the sequence.

What seems to emerge from the two extracts analysed are two completely different ways of handling the interaction, with different repercussions on the distribution of tasks and responsibilities: in sequence 1, INT seems almost tacitly ‘legitimised’ to take the lead and handle what is ultimately a complex action while teachers detach themselves from this process. While this is not negotiated a priori, it seems to be the result of a more complex, yet fine-grained, multimodal negotiation, where artefact manipulation plays a key role. What emerges from sequence 2 is, instead, what we could call a form of “shared coordination” of the interaction, where the interpreter intervenes without nevertheless taking the lead. Far from being an evaluative judgement, the point here is also to show how participatory frameworks are highly sensitive to multimodal, embodied behaviour displayed by all participants on a moment-by-moment basis throughout the encounter, which therefore need to be accounted for in any fine-grained analysis of interaction.

7. Discussion and concluding remarks
To recap, analysis has shown that participation is to be regarded as situated, co-constructed via the joint efforts of the participants in the interpreter-mediated encounter. Moreover, it is influenced not only by their verbal resources, but rather by a complex interplay of different semiotic resources, which also include the manipulation of physical objects present in the environment. This paper deals
in particular with the management of a specific object, i.e. the school report, which becomes the focus of attention for the participants due to its central role in the activity being performed. Two comparable sequences, in which two similar constellations of participants perform the same activity (i.e., reading and signing the report) have been presented with a view to showing how different ways of managing the object may be more or less conducive towards the inclusion/exclusion of some of the primary parties, and generate more (sequence 1) or fewer (sequence 2) shifts in the participatory framework.

7.1. A recurring practice: pointing at the report

Some recurring practices emerge from the analysis of the two sequences. For instance, in both cases participants frequently recur to pointing as a way of dealing with the teacher’s report. This resonates with the findings from other studies on the local management of inscribed artefacts in interaction (see, for instance, Goodwin 2000b; Mondada 2012a), which show that pointing at a document is an orderly organised activity, through which interactants perform a number of social actions, such as introducing new topics, focusing participants’ attention, claiming epistemic responsibility, prefacing other actions, and more generally mobilising the artefact as a meaningful semiotic resource.

In our data, pointing at the report not only serves the purpose of helping to explain and clarify specific points of it, but also exerts other functions. In particular, during the extensive analysis of Sequence 1, pointing has been shown to contribute towards:

- making the object relevant for the ongoing interaction and creating a joint focus of attention for the participants (5.1);
- projecting the next action, thus making the report progressively relevant to it (5.1., 5.3.);
- projecting next speakers and allowing them to self-select (5.3., 5.4.) (cf. Mondada 2007 on pointing as a way of projecting next speakers);
- triggering and facilitating progression through different phases of the encounter (5.7.).

In Sequence 2, beside the functions highlighted above, pointing is also used by the teacher in order to retain control over the interaction via a constant focus on the report, which is shared and collectively managed. By so doing, T1 generates a more stable participation format, without the fractures and schisms observed in Sequence 1.

The pointing gesture is employed by the participants in accordance with the different local configuration of participation, thus displaying their constant monitoring of the different, concurrent courses of actions. At the same time, analysis shows how this semiotic resource is implemented simultaneously with others (posture, gaze, speech, proxemics) in a multimodal ensemble that needs to be analysed in all its complexity to be able to fully understand how participation dynamics unfold.
In particular, it has been noticed that monitoring is performed by the interactants not only through speech, but also (and mostly) via other available semiotic resources (particularly gaze). As a consequence, participation can also be displayed when a participant is silent and seemingly disengaged from the ongoing interaction.

7.2. Ecology of objects and participation

Analysis has highlighted a very different development of sequence 1 and sequence 2, particularly with regard to what participants keep or yield control of during the interaction. The differences that have emerged can be related to a different ecology of objects, which involves, among other things, the participants' positioning, their ability to directly interact with the report, and more broadly the objects' affordances.

The multimodal resources mobilised during the communicative event under investigation are related to and constrained by the peculiar ecology of action of the event, which is shaped by the spatial arrangement of the participants and their limited workspace (seated around a table), the artefact they manipulate (school report) and the specific action carried out (reading and signing the report). A very clear example is M’s gesture of touching the report, thus attracting INT’s attention, as opposed to the similar (and failed) attempt on the part of T1 (see 6.4). We argued that the two different outcomes can be (at least partly) related to M having the report physically accessible and being located in a more proximal position with respect to INT than T1.

During the interaction, the report is far from being a “static” and “neutral” object: rather, it is actively used by the participants as a means for triggering changes in the participation format. In this respect, analysis of sequence 1 shows that each participation shift is marked by the use of the report as a semiotic resource. In terms of participation formats, one of the main differences between the two sequences analysed lies in the fact that, in sequence 1, the report is mostly managed by INT, whereas in sequence 2 it is mostly handled by the teachers: in both cases, physical access to the report embodies the participants' possibility of managing the unfolding of the interaction. This provides evidence of the fact that coordination and participation are negotiated by all the parties-talk on a moment-by-moment basis.

7.3. Exploring alternative patterns of behaviour

A holistic approach to the analysis of interaction, going beyond the verbal level, also helps identify the impact that different sequential displays and combinations of multimodal resources of all parties may (or may not) have in terms of participation. For instance, during the first gap in conversation (sequence 1, line 31), INT self-selects at a time when M’s response is missing but T1 is monitoring the interaction. At that sequential point, INT could have tried to pursue T1’s contribution via other-selection through gaze (Rossano 2012), without having recourse to any verbal resource; if taken up by T1, this move could have given the latter the opportunity to steer the unfolding of the interaction,
as is normally expected in this kind of meeting. Furthermore, evidence has also been provided of how a different management of the artefact can lead to different participation formats (dyadic vs triadic as visually represented in Figure 24) and, particularly, of its impact on the retaining or yielding of interactional control as well as on the distribution of tasks and responsibilities.

![Figure 24: instances of dyadic (left) and triadic (right) participation formats](image)

An empirical and micro-analytical look at different ways of handling the interaction also provides for a more nuanced view of how interaction is co-constructed, which is often the result of tacit agreement between the parties, particularly in inherently collaborative settings such as the one considered here. Identifying and comparing micro-discourse multimodal routines can contribute to raising awareness of specific phenomena as well as of the complexity of interpreter-mediated interaction. Sequence 1 is a particularly striking example of underestimation of the complexity of the ongoing action; sequence 2 also shows lack of awareness of key principles when working with an interpreter, such as of the importance of chunking long turns. More studies of this nature across different institutional settings could help identify empirical strategies for interactional management which could enhance interactional competence of all parties involved in a triadic exchange, with implications for interpreter education and for the efficiency of institutional communication.

7.4. Directions for future research

This paper’s contribution to the goals of the Special Issue can be placed at the intersection between the embodied-disembodied and the collaborative-conflictual continuums, as participation is subject to shifts towards more or less collaborative configurations and affected by a range of multimodal and multilingual resources mobilised by participants.

By contributing to our understanding of how participation is co-constructed and finely-tuned, multimodal analysis can also give us insights into how and, most importantly, to what extent interpreters can intervene in the ongoing encounter without substituting any of the primary parties. Increasing such awareness can support the moment-by-moment decision-making process, which is an integral part of an interpreter’s work as a coordinator.

More broadly, the case study presented in this paper addresses the call for “expanding
territories” (Bezemer 2015) to advance the field of multimodal studies through the exploration of an ‘a-typical’ communicative event such as an interpreter-mediated interaction. This paper also strengthens the call for a multimodal turn in interpreting studies, with research accounting systematically not only for interpreter-mediated interaction as a situated social activity, but also integrating the very ecology of such activity, which includes the spatial arrangement of the participants, their embodied actions and their use of various material resources (such as objects, documents and technologies), with a view to gaining further insights into the interactional dynamics of this complex practice.

**TRANSCRIPTION CONVENTIONS**

| (1.5) | pause (in seconds and tenths of seconds) | (.) | micro-pause (shorter than 0.5 seconds) |
| ? | ascending intonation | - | abrupt interruption of talk |
| ((text)) | inaudible | (text) | dubious text |
| °text° | quiet volume | °°text°° | very quiet volume |
| text | emphasis | : | lengthened sound |
| [ | start of overlapping talk | @ | laughter |
| - - - - - | continuation of embodied action | >> | start of embodied action |
| >text< | faster pace of speech |

**Acknowledgments**

Our special thanks go to Vassilis Korkas for his invaluable help and support throughout the process.

**References**


