The Interpreters' Newsletter 2021 (26), 137-155
DOI: 10.13137/2421-714X/33268 ISSN 1591-4127 (print) ISSN 2421-714X (online)
https://www.openstarts.units.it/dspace/handle/10077/2119

Management of overlapping speech in remote healthcare interpreting

Esther de Boe
University of Antwerp

Abstract

Dialogue interpreting is a highly complex interactional process that requires close coordination of linguistic content and embodied semiotic resources between the participants. Such coordination greatly depends on how participants relate to the "ecology of action" (Mondada 2016), i.e. their immediate environment. In remote interpreting (RI), the reduced visual access or lack hereof makes that not all participants share the same visual ecology. This compromises the efficiency of using embodied resources such as gaze and gesture in turn-taking. As a result, common interactional issues, such as overlapping speech, can disrupt the communication and may even lead to communication breakdown (De Boe 2020). Yet, the ways in which turn-taking is managed multimodally in RI remains underexplored. Therefore, this paper investigates the effect of overlapping speech on the progressivity of the communication in face-to-face interpreting, telephone interpreting and video interpreting in three simulated doctor-patient consultations. With visual access between the interpreter and the primary participants ranging from no access (by telephone) to limited access (by video) and full access (face-to-face), these interpreting methods provide a fruitful ground to analyse the interplay between the different resources used for turn-taking. The analysis shows how the management of overlapping speech is determined by the specific ecologies of action and contributes to our knowledge of multimodal processes at work in RI.

Keywords

Dialogue interpreting, remote interpreting, telephone interpreting, video interpreting, healthcare interpreting, turn-taking management, overlapping speech.

Introduction

Dialogue interpreting (DI) refers to interpreter-mediated communication involving triadic exchanges between primary participants and an interpreter (Mason 2001). Within the study of DI, turn-taking has received much attention, particularly regarding the ways in which its management differs from non-mediated dialogue. Turn-taking analysis offers insights into the unique and complex features of interaction management in interpreting and the role of the interpreter in this process (Roy 2000: 4). By investigating talk turn by turn, researchers show how participants organise communication among themselves and provide insights into the development of the co-construction of meaning and miscommunication (Wadensjö 1998: 202).

One of the most obvious characteristics of DI is the indirectness of the communication: the primary participants do not address each other directly, but take turns with the interpreter, who manages the overall interaction (Wadensjö 1998). In this process, temporality, i.e. the appropriate timing of verbal as well as embodied responses between participants (Mondada 2016), is essential to ensure synchronisation of interaction (Beukeleers *et al.* 2020). However, in DI, the presence of the interpreter as an intermediary between the primary participants makes timing particularly complex (Englund Dimitrova 1997: 162).

Especially since the Covid-19 health crisis, DI increasingly takes place remotely by means of telephone interpreting (TI) and video interpreting (VI), as well as by video relay service (VRS) combining TI and VI. Research on TI (Wadensjö 1999; Amato 2018; Castagnoli/Niemants 2018; Spinolo *et al.* 2018; De Boe 2020), VI (Braun 2004; Balogh/Hertog 2012; Braun/Taylor 2012; Licoppe/Veyrier 2017; Davitti 2018; De Boe 2020; Hansen 2020) and VRS (Napier/Leneham 2011; Warnicke/Plejert 2012) indicates that coordination of interaction is more complex in RI. This has been linked to the participants' compromised or altogether lacking visual access to each other (Braun 2017; De Boe 2020) which has consequences for the use of embodied resources. For example, gaze plays an important role in coordinating interaction (e.g. Kendon 1967; Goodwin 1981; Heath 1986). How embodied resources are used depends on the conditions in which the interaction takes place, which determine how participants interact with their immediate environment and with each other. These immediate environments are also referred to as "ecologies" (e.g. Mondada 2016). They can be spatial, when participants are physically present in one place, or visual, when they are not physically together but have visual access to (some features of) each other (Licoppe/Veyrier 2017; Davitti 2018; Hansen 2020). According to Luff *et al.* (2003: 53), "the interpretation and production of action are inextricably embedded within the immediate environment". Therefore, when participants communicate from different physical locations, the environment of action is "fractured" into separate local ecologies, which may undermine their ability to coordinate actions (Luff *et al.* 2003). In VI, examples of fractured ecologies were provided by Hansen (2020), who confirms Luff *et al.*'s (2003) conclusions that in video-mediated communication, participants' utterances can become disconnected from the ecology in which they were produced, which may pose problems for turn-taking. Moreover, as Hansen

(2020) argues, interactional issues in VI can occur as a result of participants' lack of awareness that the visual access provided by audiovisual media is limited and unequally divided between the participants, particularly in video remote interpreting, where the primary participants are located at the same place and the interpreter is present at a remote location (Braun/Taylor 2012: 39-41).

Although research focusing on interactional aspects of RI is gaining ground, much remains to be explored in this domain, especially concerning the ways in which embodied resources are used in turn-taking management. Moreover, few studies have engaged in micro-analytic investigations of the nature of RI methods (Napier *et al*. 2018: 236). Finally, most studies on remote DI have been conducted in legal settings (e.g. Braun/Taylor 2012), whereas other important contexts, such as healthcare interpreting, have so far remained under-researched.

Against this backdrop, this paper presents a micro-analysis of three simulated doctor-patient consultations. Following Mondada (2016: 340), a multimodal approach was adopted to investigate how overlapping speech (OS) was collaboratively managed by the participants. In order to investigate the interplay between the various semiotic resources at work in turn-taking, the simulations were designed around three different interpreting methods, in which visual access between the interpreter and the primary participants ranged from full access (F2FI) to no access (TI) and limited access (VI).

In what follows, I discuss key concepts concerning turn-taking (Section 1), the research methodology underpinning this paper (Section 2) and the results of the analysis (Section 3), followed by a discussion and conclusion of its outcomes (Section 4).

1. Turn-taking

1.1. Discourse-based interactionist approaches

Turn-taking is a basic feature of conversation that demonstrates its interactional character (Sacks *et al*. 1974: 728). By systematically investigating naturally occurring conversations, Sacks *et al*. (1974) exposed general rules underlying the coordination of conversation, including the principle that generally one participant talks at a time and that turn transitions may show gaps and overlaps (Sacks *et al*. 1974). Over the course of time, this theory has been further refined by explorations into the multimodal character of conversation, e.g. by studies on the role of mutual gaze in the coordination of interaction (e.g. Kendon 1967; Heath 1986; Rossano *et al*. 2009).

In Interpreting Studies, sociolinguistic approaches have been frequently applied to the study of DI by, amongst others, Wadensjö (1998), Roy (2000), Mason (2001) and Bot (2005). Within this "dialogic discourse-based interactionist paradigm" (Pöchhacker 2016: 75), researchers made elaborate reconstructions of turn-taking in DI to provide insights into the ways in which understanding in mediated discourse is achieved interactionally and conditioned by the sociocultural settings in which it takes place (Wadensjö 1998: 154). Their notions

of turn-taking were based on the work of, amongst others, Goffman (1981) and Goodwin (1981), and led to the observation that, in addition to accurately conveying messages at both a linguistic and pragmatic level, interpreters also coordinate interaction. This idea has in the meantime become widely accepted. According to Wadensjö (1998), coordination is both implicit and explicit (Wadensjö 1998). Implicit coordination is accomplished when the interpreter has his/ her turn immediately after one of the primary participants, which is also referred to as "regular turn-taking" (Roy 2000). Explicit coordination takes place when participants deviate from regular turn-taking in the form of "discursive 'moves' by the interpreter" (Pöchhacker 2016: 147), for example, a non-rendition in case of clarification of misunderstanding. Baraldi and Gavioli (2012) pointed out that coordination is in fact much more complex and introduced the notions of "basic" and "reflexive" coordination that emphasise the intertwined character of the two types of coordination.

Explicit or reflexive coordination is needed frequently in DI, for example, in case of overlapping speech (OS). First, OS can be the result of troubles to negotiate or time a transition-relevance place (TRP), a possible point of turn-transition. This usually occurs after a shorter or longer discontinuity in talk (Sacks *et al*. 1974), when an intra-turn space by a speaker is mistaken for an inter-turn space by the next speaker (Sacks *et al*. 1974). In interpreter-mediated conversation, this is usually the interpreter. A second type of OS is a direct turn: a verbal reaction by one primary speaker to the other, during the interpreter's turn or immediately following the other speaker's turn completion, before the interpreter has had a chance to start the rendition. Such a turn can be a spontaneous sign of active listening or acknowledgement (Bot 2005: 128), not aimed to take the floor, or an expression of miscomprehension by means of a request for clarification (Wadensjö 1998).

How interpreters cope with OS depends not only on their own skills, but also on the behaviour, expectations and goals of the primary participants, who, together with the interpreter, constitute a "communicative radius" (Wadensjö 1998). According to Roy (2000: 68), "[…] the participants, the discourse, and the moment combine […] to create interactional harmony whereby a turn happens successfully and comfortably". As conversation analysts have demonstrated, participants are concerned with advancing the progress of talk in interaction and therefore maximise "cooperation and affiliation" and "minimise conflict in conversational activities" (Atkinson/Heritage 1984: 55). This preference for progressivity rather than delay (Stivers/Robinson 2006: 386) also comes to the fore when a participant is searching for a word and other participants suggest options in order to allow the turn to progress (Goodwin/Goodwin 1986); their aim being to avoid "interactional difficulties" (Stivers/Robinson 2006: 368) and thus achieve trouble-free turn-taking. In this paper, I will refer to this as "smooth turn-taking", which is made possible through the participants' knowledge of how to signal and recognise potential TRPs. Speakers design their turns in such way that the other participants can recognise potential turn transitions by means of pauses, lexical choices, intonation and shifts in gaze direction (Sacks *et al*. 1974). Interpreters are actively involved in achieving smooth turn transitions by monitoring such signals (Roy 2000; Bot 2005). Since these signals are verbal as well

as non-verbal and "accessible to and actively used by all participants" (Bot 2005: 122), turn-taking can be considered a cooperative, interactional activity, which is achieved multimodally.

1.2. Multimodal dimensions of turn-taking

Although discourse-based interactionist studies certainly did not previously neglect multimodal aspects of DI, their research framework has more recently been further refined and complemented by research focusing specifically on multimodality (e.g. Pasquandrea 2011; Mason 2012; Krystallidou 2014; Davitti/Pasquandrea 2017; Davitti 2018). In addition, the introduction of technology-mediated research methods such as mobile eye-tracking (e.g. Vranjes 2018) has led to more fine-grained accounts of the role of gaze, gesture, body posture, proxemics, the handling of artefacts and spatial arrangement in interaction (Davitti/Pasquandrea 2017). In multimodal approaches, face-to-face interaction is defined as

> multimodal interaction in which participants encounter a steady stream of meaningful facial expressions, gestures, body postures, head movements, words, grammatical constructions, and prosodic contours. (Stivers/Sidnell 2005: 2)

From this approach, no semiotic resources are prioritised over others. Rather, the various modes are considered intertwined layers in the complex process of participants interacting with one another (Pasquandrea 2011: 457) by using a multitude of semiotic resources to convey meaning and monitor comprehension (Mondada 2016). Gaze direction and mutual gaze between participants play an especially important role in both the coordination of the interaction and the expression of involvement in the communication (Kendon 1967; Rossano *et al.* 2009). Since DI constitutes a "complex participation format" (Pasquandrea 2011: 456), it lends itself particularly well to multimodal analysis, especially for comparing the role and functioning of the different resources across interpreting methods involving different visual ecologies.

2. Methodology

2.1. Research design

In order to allow for a comparative, multimodal analysis of remote and face-to-face DI, a data set was designed[1], consisting of the video recordings of 9 semi-scripted simulated interpreter-mediated consultations. Each simulation involved an experienced interpreter, a gynaecologist and a simulation patient

1 The data set was originally designed for the PhD project *Remote interpreting in healthcare settings: A comparative study on the influence of telephone and video link use on the quality of interpreter-mediated communication*, defended in 2020.

and was performed using 3 different interpreting methods: F2FI, TI and VI. The current paper is based on three simulations, with a total duration of 25,52 minutes (F2FI), 26,60 minutes (VI) and 32,42 minutes (TI).

In spite of their artificial character, I chose to work with simulations, because they allow for a better control of the conditions by keeping variables (e.g. doctor and patient, themes and structure of the consultations) as constant as possible (see also Napier/Leneham 2011). The simulations took place in French-Dutch: the doctor was a native speaker of Dutch with limited notions of medical French and the simulation patient a native speaker of French with little knowledge of Dutch, whereas the interpreter was a near-native speaker of both. The simulations were designed around three closely related themes taken from real-life gynaecology practice. To encourage authentic interaction between the participants, the scenarios were non-scripted; the doctor and the interpreter were informed only about the consultations' themes, whereas the simulation patient was given a list of complaints that she had to present for each theme. The interpreter was allowed to take notes, which she did during all three interpreting methods.

The configurations used for RI were TI and VI: the doctor and patient were located in the same room, while the interpreter was present at a distant location (Braun/Taylor 2012: 39-41). In TI, a telephone on speaker mode was used; in VI, Skype video calling (for images see Sections 3.3.1-3.3.3).

All simulations were video-recorded using three different camera angles to maximise the capture of triadic and dyadic interaction in the doctor's room and one separate camera in the interpreter's room. To allow for an extensive view of the communicative situation, the recordings were synchronised by means of the transcription software Elan[2], providing a simultaneous view of the participants from different angles.

2.2. Data analysis

The video data were transcribed, provided with a gloss in English and annotated for different categories related to interpreting quality, including message equivalence issues, interactional issues (including OS), technological issues and instances of repair.[3] Based on Heritage and Maynard (2006) and Amato (2018), each simulation was divided into three parts: (1) introduction (welcome/greeting), (2) body (presentation of complaints, examination and treatment) and (3) closing phase (thanking/goodbyes). For the purpose of this paper, focusing on the management of OS – which is potentially more problematic in the body of the consultations, where it may lead to omission of important content – only the bodies of the simulations were analysed.

The analyses take into account that, within the discourse framework of doctor-patient consultations, the communication is structured sequentially, consisting mainly of three-part sequences of question-response-acknowledgement

2 <https://archive.mpi.nl/tla/elan>.

3 For a detailed overview of the annotation categories, see De Boe (2020).

(Mishler 1984). These sequences were taken as units of analysis. Each part of a sequence consists of turns, which are themselves formed by one or more turn-constructional units: sentences, clauses, phrases or single words.

To investigate turn-taking management and assess its effect on turn-taking, first, all sequences involving OS were identified and divided into two subcategories: (1) non-concurring OS: instances of OS that occurred in the absence of other issues and/or repair strategies, and (2) concurring OS: instances of OS that entailed other issues (at the level of interaction or message equivalence) and/or repair strategies. The analyses focus on the latter category, since the occurrence of issues at several levels indicates possible communication breakdown (Bot 2005). In conversations, repair comprises all actions that deal with problems of hearing, speaking and understanding talk (Schegloff *et al*. 1977), ranging from interactional to content-related issues.

Finally, the video data of the instances of concurring OS were scrutinised to establish how the various communication modes were used and combined in the management of OS across the three different interpreting methods. From these instances, the sequences illustrating most saliently the effect of the management of OS on the communication flow were selected for the purpose of this paper.

3. Results

3.1. Types of OS

The most frequently observed type of OS resulted from the participants' (especially the interpreter's) interactional behaviour of leaving no or little time in between turns. This complies with the conversational rule identified by Sacks *et al.* (1974) that participants in a conversation attempt to keep spaces in between turns to a minimum in order to ensure a smooth flow of talk. This "rapid turn-taking", taking the turn "as soon as the opportunity arises" (Englund Dimitrova 1997: 149-150) often caused OS, which was associated with participants' trouble with signalling and recognising TRPs, usually after a silence. When a participant wants to cede the floor to the next speaker, s/he usually slows down his/her speaking, drops his/her intonation, hesitates or pauses (Sacks *et al*. 1974), often combined with gazing at the next speaker (Kendon 1967). With all three interpreting methods, it occurred frequently that a pause by the current speaker was taken for a TRP by the next speaker, while in fact, it turned out to be an intra-turn pause. In the data observed, it was usually the interpreter who overlapped and dropped out by abandoning the turn, while the current speaker continued, thereby re-establishing the "one-at-a-time" principle of turn-taking (Sacks *et al*. 1974). However, in both types of RI, it turned out to be problematic for the interpreter to find the right moment to regain the floor (see Sections 3.2.2 and 3.2.3). The second type of OS that pervaded the data consisted of direct turns. Although the majority of both types of OS concurred with other issues and/or necessitated further repair, these overlaps were not necessarily disruptive. Rather, their impact on the communication flow depended on how OS was managed multimodally (see Section 3.2.3).

3.2. Management of potentially disruptive overlapping speech

3.2.1. Face-to-face interpreting

In the F2FI session, OS hardly posed a threat to smooth turn-taking. This is illustrated by Excerpt 1[4]. In this simulation, the patient consults the gynaecologist concerning pelvic floor issues.

1	DO	oké (.) euh:m (.) gebruikt u op dit moment anticonceptie/ (.) want u hebt drie kinderen/ *oké (.) eh:m (.) are you using at this moment contraception/ (.) because you have three children/*
2		(.) de jongste is vijf/ (.) wat doet u om niet meer zwanger te worden/ (.) *the youngest is five/* (.) *what do you do not to get pregnant anymore/*
3	IN	est-ce que vous utilisez la contraceptio:n (.) une forme quelconque euhm (.) vous avez des enfants *do you use contraceptio:n* (.) *a form of any kind? ehm* (.) *you have children*
	in	(gazes at PA ----------|) (mutual gaze PA/IN ----------|)
4		(.) le plus jeune a cinq ans (.) qu'est-ce que vous utilisez pour protéger/ (.) *the youngest is five years* (.) *what do you do to protect /* (1.1)
	in	(gazes down at her notes-------| gazes at PA ----------->>)
	pa	(gazes at IN -----| nods & gazes down ----------|)
5	PA	pour [l'instant nous] *for [the moment we]*
	comm	(mutual gaze PA/IN->>)
6	IN	[pour ne pas être enceinte/] [*not to be pregnant/*]
	comm	(mutual gaze PA/IN---------->>)
7	PA	(.) on utilise des préservatifs\ (..) j'ai pas repris la pilule (.) *we use condoms* (..) *I did not restart with the pill*
	pa	(gazes at IN ----------)
	in	(gazes at P----------|gazes at DO-----|gazes at PA---------->>)
	in	(facial expression----------|)
8	IN	après/ *afterwards/*
	comm	(mutual gaze PA/IN-->>))
9	PA	j'ai pas repris après *I didn't start again afterwards*
	comm	(mutual gaze PA/IN---------|)
10	IN	ik ben niet herbegonnen met mijn pil (.) wij gebruiken enkel preservatieven\ *I have not restarted with my pill* (.) *we are only using condoms*
	int	(gazes at DO---------->>)
	do	(gazes at IN----------|gazes down---------->>)
	do	(nodding)

Excerpt 1 (F2FI-1)

4 See appendix for transcription conventions. To maintain the readability of the transcripts, only multimodal comments relevant to the analysis are included in the transcripts.

The doctor (DO) asks the patient (PA) what kind of contraception she uses (Line 2). The interpreter (IN) renders this as "what do you use to protect", her intonation going slightly up (Line 4). During Lines 3 and 4, PA gazes at IN, then halfway through Line 4, PA gazes down and slightly nods her head as an acknowledgement that she has understood the question. This is an anticipation by PA to take the next turn, since participants who want to take the floor tend to look way from the speaker before initiating their turn (Kendon 1967). As IN pauses (1.1) at the end of Line 4, PA interprets IN's intra-turn pause – in spite of the slightly rising intonation – for an inter-turn pause, possibly because IN's rendition of DO's question can easily be understood without IN's addition afterwards. PA starts a turn (Line 5), thereby causing OS with IN's self-repair in the form of the addition "not to be pregnant" (Line 6). During the OS, PA shifts her gaze back to IN, engaging in mutual gaze. IN makes a facial expression indicating embarrassment (possibly for overlapping with PA) and, towards the end of PA's turn, briefly shifts her gaze to DO, apparently showing readiness to provide the rendition. However, PA is still finishing her turn (Line 7). IN first opens and then closes her mouth and shifts her gaze back to PA, leading again to mutual gaze with the patient, and then quickly alternates her gaze between DO and PA. In other words, IN temporarily suspends her turn, but as soon as PA stops talking, IN briefly requests clarification from her (Line 8). While PA provides a clarification by repeating her phrase and adding "afterwards" to it (Line 9), IN already shifts her gaze back to DO and leaves no gap between the end of PA's turn and her own rendition, in which she summarises the two previous turns by PA in one rendition (Line 10). During the entire sequence, DO's gaze pattern is very stable. She gazes at PA, except for during Line 10, where she briefly engages in mutual gaze with IN, before looking down and acknowledging the information by means of head nodding. This is in line with findings by Bot (2005: 137) that therapists usually gaze at patients, also during the interpreter's turn, to monitor the patient's reactions.

To sum up, Excerpt 1 shows that during the OS, the participants relied heavily on gaze to regulate turn-taking after overlap occurred. In this way, they verified whether the other person continued her turn or not. As a matter of fact, the interpreter did not drop out completely from the overlap, but remained visibly in stand-by mode, looking up from her note-book and shifting gaze continuously (see Figure 1). This confirms findings from Oloff (2013: 139) describing a continuous monitoring of the availability of the co-participant and of the next possible occasion to resume the suspended turn. Moreover, the interpreter monitors her comprehension by requesting clarification. As a result, the OS causes only a slight disruption of the communication flow.

Figure 1: Face-to-face interpreting

3.2.2. Telephone interpreting

In TI, the turn-taking rhythm was slowed down as the participants left more space in between the turns, especially the interpreter. This corroborates Wadensjö's (1999) findings that, in TI, participants have a more cautious way of communicating, which results in longer moments of turn transition.

Excerpt 2 (on the theme of pregnancy) shows how a pause by the interpreter leads to an overlap, while repair is made difficult due to noise outside the room.

1	IN	euhm (.) moi je vais vous expliquer qu'est-ce qu'un cycle de in vitro fait (.) *ehm (.) I am going to explain to you what an in vitro cycle does* (.)
2		qu'est-ce que ça signifie et les chances que vous avez alors de de réussir (.) *what it means and the chances you have well to succeed* (.)
3		et c'est pour vous et pour votre partenaire de 25 pour cent\ *and it's for you and for your partner 25 per cent*\
4		(1.2) [par cycle] [*per cycle*]
5	PA	[25 pour cent de chance] de plus D'AVOIR UN ENFANT/ [*25 per cent of chance*] *more TO HAVE A CHILD/*
	pa	(frowning---------->)
	pa	(gazing at telephone---------->)
	do	(gaze at PA ---| gaze at telephone ---------->)
6	IN	20:::25 per cent kans meer om zwanger te worden (.) om een kind te krijgen/ *20:::25 per cent chance of getting pregnant* (.) *to have a child/*
	env	(noise/bad sound quality---------->)
	do	(leaning forward---------->)
	do	(gazing at telephone---------->)
	pa	(gazing at telephone---------- | gazes at DO ---------->)
7	PA	ou bien dans 25 pour cent des cas je tombe enceinte/ *or in 25 per cent of the cases I get pregnant*
	pa	(gazes at telephone---------->)
	env	(noise/bad sound quality---------->)
	do	(leaning further forward---------->)
8	IN	of is het dat ik in 5:25 per cent gevallen zwanger word/ *or is it that in 5:25 per cent of the cases I get pregnant/*
	do	(leaning forward---------->)
	pa	(gazes at DO---------->)
9		(2.1)
	DO	ja, dat dat laatste (.) per cyclus/ (..) *yes, the the latter* (.) *per cycle/* (..)
	do	(gazes at telephone---------->)
	pa	(gazes at DO ---------->)
10		we moeten vier vrouwen behandelen om een iemand zwanger te krijgen\ *we have to treat four women to get one person pregnant*\
	do	(gazes at telephone---------->)
	pa	(gazes at DO ---------->)

Excerpt 2 (TI-1)

In Excerpt 2, DO has just explained PA's chances of conceiving. This is rendered by IN in Lines 1-3, after which she takes a pause (1.2), before adding "per cycle" (Line 4). During IN's turns, DO gazes down at the desk, while PA stares at the telephone. When IN stops speaking, which afterwards turns out to be an intra-turn pause (since she continues her turn after the pause), PA takes the silence as an inter-turn pause, possibly also because of IN's dropping intonation and because the turn can be considered complete in terms of syntax. PA takes the floor to request clarification (Line 5), leading to OS, and does not cede the turn. She increases the volume of her voice, which is a way of holding the floor (Schegloff 2000) and continues to gaze at the telephone, frowning. DO follows PA's gaze at the telephone. While IN takes the next turn to render PA's

question (Line 6), there is noise in the hallway outside the room, which compromises the sound produced by the speaker mode in the doctor's room. DO moves her head closer to the telephone, especially her ear, indicating difficulty hearing. PA shifts her gaze from the telephone to DO (see Figure 2). Immediately when IN stops speaking, PA shifts her gaze back to the telephone and asks another question (Line 7), while DO still holds her head close to the telephone. DO remains in this position until IN has rendered the second request (Line 8), while PA gazes again at DO. After IN has finished talking, DO leaves a long silence (2.1) before answering the questions (Turn 10). During her answer, DO gazes at the telephone, whereas PA continues to gaze at DO.

Figure 2: Telephone interpreting – Interpreter (left) / Doctor and patient (right)

Excerpt 2 illustrates that, due to the low sound quality of the speaker mode, the doctor and patient seemed to have to concentrate hard on understanding the interpreter, as expressed by their frequent frowning and leaning forward towards the telephone. Both doctor and patient consequently gazed more frequently at the telephone than at each other. Moreover, at times, background noise (e.g. in the hallway) compromised the sound quality even further, which also impacted on the gaze pattern. This pattern differed completely from the one observed in Excerpt 1. Whereas in F2FI, the primary participants continuously shifted their gaze between themselves and the interpreter, in TI, the primary participant who was being addressed by the interpreter's renditions stared at the telephone, while the other primary participant gazed at the participant who was being addressed. Meanwhile, mutual gaze was established only sporadically.

3.2.3. Video interpreting

In VI, pauses leading to OS frequently disrupted the interaction, confirming findings by Braun (2004: 85) indicating that interactional phenomena such as pauses, OS and listener responses function differently because of the difference in timing due to delay in the transmission of sound and image in VI. This type of OS, as well as OS caused by a direct turn, led to communication breakdown on several occasions. Apart from difficulties with timing, OS additionally caused sound quality problems in VI, which occur in this particular type of videoconferencing software when two people speak at a time. This further complicated the timing of turn-taking and rendered repair more complex. Moreover, as came up

in the post-simulation interviews, the camera angle of the laptop in the doctor's room made the view of the interpreter insufficient for the primary participants.

In Excerpt 3, the patient consults the doctor for complaints of abundant menstruation.

1	IN	euhm si jamais on donnait euh une *spirale* (.) une forme de stérilet sans hormones (.) *ehm in case we give euh a spirale* (.) *a form of IUD without hormones* (.)
	pa	(gazes at screen----------|)
	do	(gazes at screen----------|)
2		les saignements peuvent augmenter (..) [et (..)] *the bleeding can increase* [*and* (..)]
	int	(gazes down at notes---------->)
	com	(mutual gaze/ laughter DO/PA---------->)
	env	(bad sound quality---------->)
3	PA	[oh non] (.) là non [*oh no*] (.) *that no*
	int	(gazes down ------| gazes at screen-->)
	com	(mutual gaze/laughter DO/PA)
	env	(bad sound quality----------|)
4	IN	et les douleurs aussi (.) of heb ik het niet goed gehoord/ *and the pain too* (.) *or didn't I hear that right/*
	in	(frowning ----------|)
	in	(bending forward to the PC---------->)
5		(3.2)
	DO	hoe bedoelt u/ *what do you mean/*
	do	(raising eyebrows --|)
	do	(smiling---------- |)
6	IN	u zei op het einde dat het zonder hormonen (.) *you said in the end that without hormones it* (.)
	In	(frowning----------|)
7		een spiraal zonder hormonen of heb ik het niet goed gehoord/ *an IUD without hormones or didn't I hear that right/*
8	DO	nee (.) dat klopt *no* (.) *that is right*
9	IN	oké *okay*
10	DO	een spiraal zonder hormonen geeft meer bloedverlies en meer pijn *an IUD without hormones causes more loss of blood and more pain*
11	INT	tout à fait (.) j'ai j'avais bien compris (..) donc une spirale ou une sorte de stérilet sans hormones (.) *precisely* (.) *I have I had understood it right* (..) *so a spiral or a kind of IUD without hormones* (.)
12		aurait notamment fait plus de saignements et causerait plus de saignements (.) *would actually have made more bleeding and would cause more bleeding* (.)
13		et plus de douleurs abdominales *and more abdominal pain*

Excerpt 3 (VI-3)

In this sequence, DO has just explained the pros and cons of placing an IUD without hormones as a possible treatment. While IN is speaking (Lines 1-2), PA and DO are both gazing at the screen. When IN comes to the part of the information saying that this type of IUD may increase the bleeding (Line 2), PA quickly shifts her gaze to DO, who gazes back at her, and reacts immediately in an emotional way in a direct turn ("oh no, that, no!", Line 3). PA and DO continue their mutual gaze and laugh together. However, IN has not finished her rendition. At the moment PA utters her reactive expression (Line 3), IN is just gazing down at her notes, as a result of which she does not see that PA addressed DO directly, but only hears a distorted sound, which is caused by the mutual laughter between DO and PA as a result of PA's reaction to DO (See Figure 3). IN bends forward to the screen and frowns, inquiring "or didn't I hear that right?" (Line 4). During IN's request for clarification, PA and DO shift their gaze from each other to the screen, still smiling. However, in the doctor's room, no sound distortion was audible. Therefore, the question by IN seems to come as a surprise to DO, who leaves a gap of 3.2 seconds after IN's question, which may be a way to avoid further OS (in case IN would resume her turn), and then asks IN what she means (Line 5). IN verifies with DO if she has provided the correct information to PA (Lines 6-7), which DO confirms (Line 8) and which is followed by an acknowledgement by IN (Line 9). DO then repeats the information she provided earlier (Line 10). Subsequently, IN also repeats the information she has rendered before in French, preceded by a meta comment confirming her correct understanding (Lines 11-13).

Figure 3: Video interpreting – interpreter (left) / doctor and patient (right)

Although in the end, there is no loss of message content, the direct turns and laughter caused technological issues, which inhibited smooth turn-taking and complicated repair, rendering the communication altogether inefficient. In addition, the communication breakdown obviously led to feelings of insecurity on the part of the interpreter, as apparent by her request for clarification, facial expression (frowning), posture (moving closer to the computer) and two self-repairs ("I have I had" and "would actually have made more bleeding and have caused more bleeding", Lines 10-11). Excerpt 3 also illustrates the role of gaze. Due to the lack of mutual gaze while the interpreter was looking at her notes instead of at the screen, she did not notice the direct turn and did not yield the turn earlier, which could have limited the impact of the OS. It also demonstrates that the

fragmented ecologies caused by the reduced visual access and delay in transfer of image and sound can lead to interactional issues.

4. Discussion and conclusion

In this paper, I investigated the effect of overlapping speech on the progressivity of the communication in three different interpreting methods: F2FI, TI and VI. By means of multimodal analyses of potentially problematic instances of OS in three simulated doctor-patient consultations, I demonstrated differences in the accomplishment of turn-taking, that were linked to the specific ecologies of action created by the remote conditions.

In F2FI, where the participants had full visual access to each other, the occurrence of OS was not necessarily problematic. Even when combined with other issues and/or repair, OS hardly caused disruption of the communication flow. When OS occurred, the interpreter immediately reacted by withdrawing from the turn, and when repair was needed, it was carried out efficiently, supported by close monitoring, especially gaze. Moreover, as opposed to both RI methods, the participants could still hear what was said by both participants involved in the overlap, which also facilitated smooth turn-taking.

In TI, OS did not appear to be extremely disruptive either, due to the participants' more careful ways of communicating, leaving longer moments of silence in between the utterances and the renditions. Nevertheless, sound quality issues in the form of background noise deteriorated the already weak speakerphone volume, which complicated the management of OS. In addition, mutual gaze between the doctor and the patient was reduced. Since mutual gaze plays an important role in establishing rapport (Krystallidou 2014), i.e. a relationship of trust and mutual responsiveness between healthcare provider and patient, fostering the therapeutic process[5], this seems a negative side effect of TI. The same applies to involvement in the interaction, which is also expressed by means of gaze (Kendon 1967; Rossano *et al*. 2009; Pasquandrea 2012: 150).

In VI, however, OS frequently disrupted the communication flow, mostly due to the delay of sound and image, which caused sound quality problems and rendered the timing of turns more complex, leading to 'hitches' in the communication flow. Moreover, the interpreter's difficulties with the timing of a renewed attempt to take the floor after her initial withdrawal indicated that the use of embodied resources for repair was less efficient than in F2FI. This was also related to the interpreter's note-taking. In VI, the interpreter missed important visual cues because she was gazing down at her notes, whereas in F2FI, the interpreter could easily combine reading her notes with gazing up frequently to monitor turn-taking.

The analyses also illustrate that in RI, the limited visual access and delay in sound and image (especially in VI) made it extremely difficult for the participants to project the end of a turn and identify TRPs. Moreover, the access to embodied

5 *Farlex Partner Medical Dictionary* (2012) <https://medical-dictionary.thefreedictionary.com>.

communication modes between the participants was unequal in both remote settings. Since in TI and VI, the doctor and patient were physically together, they were able to fully employ embodied resources towards each other, whereas the interpreter only had partly access to them. As a result, it was particularly difficult for the interpreter to monitor the achievement of mutual understanding and rapport-building between the primary participants. The fractured ecologies that were thus created had implications for the achievement of smooth turn-taking.

Another observation is related to awareness. Rather surprisingly, OS was less disruptive in TI than in VI. This seemed to be related to participants' experience with this medium, which makes them conscious of the fact that the participants on the other side only have auditory access and makes them slow down the pace of communication. This was a clear difference with VI, in which the participants hardly seemed aware of the visual and auditory constraints of the medium and behaved similarly as in F2FI. This made them react in a more spontaneous, face-to-face manner, causing OS that in turn disrupted the communication flow and rendered both turn-taking and repair less fluent, confirming previous results by Hansen (2020).

Overall, this paper confirms the central role of embodied resources in turn-taking in DI by demonstrating how overlap management is determined by the specific ecologies of action created by remote conditions. It also shows the complexity of these ecologies and illustrates the usefulness of multimodal analysis as a methodology to investigate remote dialogue interpreting.

Transcription conventions adapted from Davitti (2018), drawing on Jefferson (2004) and Mondada (2016)

(1.5)	silence expressed in seconds
(.)	micropause of less than 1 second
:	sound elongation
CAPITALS	raised voice
text/	rising intonation
text\	falling intonation
[text]	onset and end of OS
-->	action described continues across subsequent lines
--|	action described ends
in/pa/do	lower case for embodied behaviour of participant
comm	commentary on participants' behaviours in brackets (text)
env	environmental issue, (e.g. sound quality)

References

Amato A. (2018) "Challenges and solutions: some paradigmatic examples", in A. Amato / N. Spinolo / M. J. González Rodríguez (eds) *Handbook of Remote Interpreting: Research Report Shift in Orality Erasmus + project: Shaping the Interpreters of the Future and of Today*, 79-98, <http://shiftinorality.eu/es/resources/2018/05/11/shift-handbook-remote-interpreting>.

Atkinson J. M. / Heritage J. (1984) (eds) *Structures of Social Action: Studies in Conversation Analysis*, Cambridge, Cambridge University Press.

Balogh K. / Hertog E. (2012) "AVIDICUS comparative studies – part II: Traditional, videoconference and remote interpreting in police interviews", in S. Braun / J. Taylor (eds) *Videoconference and Remote Interpreting in Criminal Proceedings*, Guildford, University of Surrey, 101-116.

Baraldi C. / Gavioli L. (eds) (2012) *Coordinating Participation in Dialogue Interpreting*, Amsterdam/Philadelphia, John Benjamins.

Beukeleers I. / Brône G. / Vermeerbergen M. (2020) "Unaddressed participants' gaze behavior in Flemish Sign Language interactions: planning gaze shifts after recognizing an upcoming (possible) turn completion", *Journal of Pragmatics* 162, 62-83.

Bot H. (2005) *Dialogue Interpreting in Mental Health*, Amsterdam/New York, Rodopi.

Braun S. (2004) *Kommunikation unter widrigen Umständen?: Fallstudien zu einsprachigen und gedolmetschten Videokonferenzen* (in English, *Case Studies on Monolingual and Interpreted Video Conferences*), Tübingen, Narr.

Braun S. (2017) "What a micro-analytic investigation of additions and expansions in remote interpreting can tell us about interpreters' participation in a shared virtual space", *Journal of Pragmatics* 107, 165-177.

Braun S. / J. L. Taylor (2012) "AVIDICUS comparative studies – part I: Traditional interpreting and remote interpreting in police interviews", in S. Braun / J. Taylor (eds) *Videoconference and Remote Interpreting in Criminal Proceedings*, Guildford, University of Surrey, 85-100.

Castagnoli S. / Niemants, N. (2018) "Corpora worth creating: a pilot study on telephone interpreting", Special Issue of *InTRAlinea*, <http://www.intralinea.org/ specials/article/corpora_worth_creating_a_pilot_study_on_telephone_interpreting>.

Davitti E. (2018) "Methodological explorations of interpreter-mediated interaction: novel insights from multimodal analysis", *Qualitative Research* 19/1, 7-29.

Davitti E. / Pasquandrea S. (2017) "Embodied participation: what multimodal analysis can tell us about interpreter-mediated encounters in pedagogical settings", *Journal of Pragmatics* 107, 105-28.

De Boe E. (2020) *Remote interpreting in healthcare settings: A comparative study on the influence of telephone and video link use on the quality of interpreter-mediated communication*, unpublished PhD thesis, University of Antwerp at Antwerp.

Englund Dimitrova B. (1997) "Degree of interpreter responsibility in the interaction process in community interpreting", in S. E. Carr (ed.), *The Critical*

Link: Interpreters in the Community, Amsterdam/Philadelphia, John Benjamins, 147-164.
Goffman E. (1981) *Forms of Talk*, Philadelphia, Pennsylvania Press.
Goodwin C. (1981) *Conversational Organization Interaction between Speakers and Hearers*, New York, Academic Press.
Goodwin M. H. / Goodwin C. (1986) "Gesture and coparticipation in the activity of searching for a word", *Semiotica* 62, 51-75.
Hansen J. P. B. (2020) "Invisible participants in a visual ecology: visual space as a resource for organizing video-mediated interpreting in hospital encounters", *Social Interaction. Video-based Studies of Human Sociality* 3/3, 1-25.
Heath C. (1986) *Body Movement and Speech in Medical Interaction*, Cambridge, Cambridge University Press.
Heritage J. / Maynard D. W. (2006) *Communication in Medical Care: Interaction between Primary Care Physicians and Patients*, Cambridge, Cambridge University Press.
Jefferson G. (2004) "Glossary of transcript symbols with an introduction", in G. H. Lerner (ed.)
Conversation Analysis: Studies from the First Generation, Amsterdam/Philadelphia, John Benjamins, 13-31.
Kendon A. (1967) "Some functions of gaze direction in social interaction", *Acta Psychologica* 26, 22-63.
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.
Licoppe C. / Veyrier C. (2017) "How to show the interpreter on screen? The normative organization of visual ecologies in multilingual courtrooms with video links", *Journal of Pragmatics* 107, 147-164.
Luff P. / Heath C. / Kuzuoka H. / Hindmarsh J. / Yamazaki K. / Oyama S. (2003) "Fractured ecologies: creating environments for collaboration", *Human-Computer Interaction* 18, 51-84.
Mason I. (2001) *Triadic Exchanges: Studies in Dialogue Interpreting*, Manchester, St Jerome Publishing.
Mason I. (2012) "Gaze, positioning and identity in interpreter-mediated dialogues", in C. Baraldi / L. Gavioli (eds) *Coordinating Participation in Dialogue Interpreting*, Amsterdam/Philadelphia, John Benjamins, 177-199.
Mishler E. G. (1984) *The Discourse of Medicine: Dialectics of Medical Interviews*, Norwood, Ablex Publishing Corporation.
Mondada L. (2016) "Challenges of multimodality: language and the body in social interaction", *Journal of Sociolinguistics* 20/3, 336-366.
Napier J. / Leneham M. (2011) ""It was difficult to manage the communication": testing the feasibility of video remote signed language interpreting in court", *Journal of Interpretation* 21/1, <https://digitalcommons.unf.edu/joi/vol21/iss1/5>.
Napier J. / Skinner R. / Turner G. H. (2018) "Enabling political participation through remote interpreting: a case study", in J. Napier / R. Skinner /

S. Braun (eds) *Here or There: Research on Interpreting via Video Link*, Washington, DC, Gallaudet University Press, 230-236.
Oloff F. (2013) "Embodied withdrawal after overlap resolution", *Journal of Pragmatics* 46/1, 139-156.
Pasquandrea S. (2011) "Managing multiple actions through multimodality: doctor's 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.
Pöchhacker F. (2016) *Introducing Interpreting Studies*, London/New York, Routledge.
Rossano F. / Brown P. / Levinson S. C. (2009) "Gaze, questioning and culture", in J. Sidnell (ed.) *Conversation Analysis: Comparative Perspectives*, Cambridge, Cambridge University Press, 187-249.
Roy C. (2000) *Interpreting as a Discourse Process*, New York and Oxford, Oxford University Press.
Sacks H. / Shegloff, E. A. / Jefferson, G. (1974) "A simplest systematics for the organization of turn-taking in conversation", *Language* 50, 696-736.
Schegloff E. (2000) "Overlapping talk and the organization of turn-taking for conversation", *Language in Society* 29/1, 1-63.
Schegloff E. A. /Jefferson G. / Sacks H. (1977). "The preference for self-correction in the organization of repair in conversation", *Language* 53/2, 361-382.
Spinolo N. / Bertozzi M. / Russo M. (2018) "Shaping the interpreters of the future and of today: preliminary results of the SHIFT Project", *The Interpreters' Newsletter* 23, 45-61.
Stivers T. / Robinson J. D. (2006) "A preference for progressivity in interaction.", *Language in Society*, *35*(3), 367-392.
Stivers T. / Sidnell J. (2005) "Introduction: multi-modal interaction", *Semiotica* 156/1-4, 1-20.
Vranjes J. (2018). *On the role of eye gaze in the coordination of interpreter-mediated interactions an eye-tracking study*, unpublished PhD thesis, University of Leuven at Leuven. <https://lirias.kuleuven.be/retrieve/525192>.
Wadensjö C. (1998) *Interpreting as Interaction*, London/New York, Longman.
Wadensjö C. (1999) "Telephone interpreting and the synchronization of talk in social interaction", *The Translator* 5/2, 247-264.
Warnicke C. / Plejert C. (2012) "Turn-organisation in mediated phone interaction using Video Relay Service (VRS)", *Journal of Pragmatics* 44, 1313-1334.