Video-mediated interpreting

The interactional accomplishment of interpreting in video-mediated environments

Jessica Pedersen Belisle Hansen

MultiLing Center for Multilingualism in Society across the Lifespan
Department of Linguistics and Scandinavian Studies
Faculty of Humanities

University of Oslo
2020
til Vegard

takk for at du går sammen med meg
Abstract

The thesis explores the interactional accomplishment of interpreting in video-mediated environments. Three articles explore video recordings of video-mediated interpreting in hospital encounters through conversation analysis. The final article employs discourse analysis to explore interviews and government documents.

Employing multimodal conversation analysis, the first article explores the temporary suspension of medical professionals’ turns in order to let the interpreter interpret. The second explores interpreters’ embodied displays of trouble caused by trouble hearing and understanding in the video-mediated environment. The third article explores how participants’ orient to the visual affordance of the media in the organization of interpreting. The thesis demonstrates how interpreting is an interactional activity accomplished collaboratively by participants in situ. The participants have different access to linguistic content, to the visual ecology and to background knowledge. Features of the video-mediated environment, like delay and lack of mutual visual access, may delimit which resources participants have available to organize interaction, for instance pre-beginning signals and embodied actions. Participants can overcome some challenges through adapting their actions to the setting, like the temporary suspension of medical professionals’ longer turns, and by creating an interactional space that is appropriate for the activities.

The fourth article discusses the comparison of video-mediated interpreting to telephone and onsite interpreting in government documents and interviews with practitioners. While government documents use the comparison between media to construct a rationale for increased use of video-technology to provide interpreting, practitioners’ narratives demonstrate how technology is relevant for the accomplishment of their work. The combination of analytical approaches demonstrates how technology as a workspace is not just a matter of efficient service provision but fundamentally alters resources participants have available to establish understanding in interaction.
Sammendrag


Ved bruk av multimodal samtaleanalyse utforsker den første artikkelen hvordan tolk og helsepersonell forhandler om midlertidige opphold i helsepersonells lengre turer for at tolken skal komme til for å tolke. Den andre utforsker tolkers uttrykk for problemer med å høre eller å forstå i det video-medierte rommet. Den tredje viser hvordan deltakerne bruker mediets visuelle affordanse i utførelsen av tolking. Avhandlinga viser hvordan tolking er en interaksjonell aktivitet som organiseres av deltakere i og gjennom samhandlingen. Deltakerne har ulik tilgang til språklig innhold, ulik tilgang til den visuelle økologien og forskjellig bakgrunnskunnskap. Trekk ved det video-medierte rommet som forsinkelse og redusert visuell tilgang, kan begrense ressursene deltakerne har til rådighet for å få til interaksjonen, som pre-begynnelsessignaler og gester. Deltakere kan løse noen av utfordringene ved å tilpasse handlingene til situasjonen, for eksempel kan helsepersonell skape foreløpige opphold i sine lengre taleturer slik at tolken kommer til for å tolke, og ved å tilpasse det interaksjonelle rommet til aktivitetene de skal utføre.

Den diskursanalytiske artikkelen drøfter sammenlikningen av video-mediert tolking med telefon- og fremmøtetolking i offentlige dokumenter og i intervjuer med tolker og helsepersonell. Ved å sammenligne skjermtolkning med telefon- og fremmøtetolking konstruerer offentlige dokumenter et argument for økt bruk av videoteknologi til tolking. Mens dokumentene i liten grad forholder seg til hvordan interaksjonen i slike samtaler utfolder seg, viser praktikeres fortellinger om skjermtolkede samtaler hvordan teknologien er av betydning for utførelsen av arbeidet deres. Kombinasjonen av analysemetodene viser hvordan teknologien som arbeidsplass ikke bare er et spørsmål om effektiv tjenesteleveranse, men endrer ressursene deltakere har til rådighet for å etablere forståelse.
Acknowledgements

This thesis would never have come into existence without the help and support of so many wonderful people. I am deeply grateful to all those who I cannot mention by name for the risk of identifying participants in the study: Patients and next-of-kin who have kindly allowed me to video record their medical appointments. Practitioners, both medical professionals and interpreters, who have let me observe, video-record and analyze their work practices. Practitioners and other stakeholders who have met with me, shared their views, and helped me gain new perspectives. Interpreters who have helped me with linguistic expertise. All those who have helped me coordinate the project, identify relevant meetings to record, and organize the recordings in a secure manner. This has been important for the accomplishment of the project. I am truly thankful!

I want to thank my supervisors, Jan Svennevig and Jannis Androutsopoulos. They have supported, inspired, advised, enlightened and challenged me throughout the process. I have appreciated Svennevig’s guidance and supervision since he first taught me about adjacency pairs in 2013. I have had the pleasure of getting to know Androutsopoulos through the work with this thesis. Through his supervision, I have had the opportunity to explore new directions and to see my work in a new context. I am grateful for their willingness to follow me down rabbit holes and, perhaps more importantly, to help me back on track again.

For collaboration, I would also like to thank Tolkesentralen, Oslo University Hospital, Akershus University Hospital and Sunnaas Rehabilitation Hospital.

For linguistic expertise, I would like to thank Pawel Urbanik, Maria Obojska, Visnja Cicin-Sain, Solvej Helleshøj Sørensen and Annika Lisa Belisle. I would also like to thank also Mari J. Wikhaug Andersen and Marit Johanne Furunes for help with transcribing.

For inspiration and support through different forms and channels during the process, I would like to thank several scholars: Marja Etelämäki, Jennifer Gerwing, Hanne Skaaden, Elizabeth Stokoe, and Johan Tønnesson. I would especially like to thank Christian Licoppe for highly appreciated insights and comments for my midway evaluation.

I have had the pleasure of participating in various research groups during my work, all of which have inspired me, helped ground my work and provided an opportunity to discuss and develop my studies. I would like to thank the members of the research groups Conversation analysis and interactional linguistics at the University of Oslo, Network for interaction analysis at the University of Southern-Eastern Norway, Text and Rhetoric at the University of Oslo, Interpreting, language and communication at Oslo Metropolitan for their help and support, for inspiration and stimulating discussions.

I would like to thank the Centre for Interaction Research and Communication Design at the University of Copenhagen for inviting me to stay with them in October and November 2019, and for including me in their inspiring research environment during the stay. A special thanks goes to Mie Femø Nielsen, Brian Lystgaard Due, Simon Bierring Lange, Thomas Toft, Jeanette Landgrebe, Anne Merrit Rikke Nielsen and Marta Karolina Olsen for making me feel welcome, including me in their daily activities and helping me develop my work during my stay.
In Copenhagen, I was invited to spend time with another research group, INTERPRETING. This collaboration resulted in new insights and further development of my work. I would like to thank Martha Sif Karrebæk, Paulina Bala, Marta Kirilova and Solvej Helleshøj Sørenson for inspiring discussions, seminars and fieldtrips.

I would like to express my gratitude to MultiLing Center for Multilingualism in Society across the Lifespan. The center has provided an inspiring and vibrant research environment with rich access to scholars studying multilingualism from a range of different disciplines. I would like to thank especially Elizabeth Lanza and Mari C. R. Otnes for making difficult times during this process easier. Also, I would like to thank Elisabeth Maria Neuhaus for pointing me in the right direction and helping me sort out all kinds of matters on so many occasions.

Thank you to all my colleagues and fellow fellows at the Department of Linguistics and Scandinavian Studies and MultiLing. Many thanks, especially to Hanna Solberg Andresen, Rafael Lomeu Gomes, Jenny Gudmunsen, Pernille Hansen, Helene Kilmer-Rumpf, Elina Kangas, Anne Marie Dalby Landmark, Kristin Vold Lexander, Kristin Torjesen Marti, Maria Obojska, Ingeborg Sophie Bjønnes Ribu, Ida Seljeseth, Yesim Sevinc, Jorunn Simonsen Thingnes, Pawel Urbanik and Ingvild Badhwar Valen-Sendstad.

For inspiration, support and help during different stages of the process, I would like to thank Hilde Fiva Buzungu, Hanne M. Løfsnes, Marthe Ødegård Olsen, Camilla Sandrud, Maria Wattne, Kristian Skedsmo and Randi Havnen.

I am grateful to my parents, my siblings and their families for patience and care. Without all of you, I would not be me. A special thanks goes to Mari Bjerkli and Annika L. Belisle for their helpful advice as readers.

Last, but not least, without the love and support of my partner, I would not have been able to complete this thesis. Thank you for your patience. Thank you for your support. Thank you for walking with me.
# Contents

1 Introduction ......................................................................................................................... 1

1.1 Good communication in healthcare ................................................................................. 2

1.2 Video-mediated interpreting as a solution to a range of problems ................................. 3

1.3 Aim and scope of the study ............................................................................................. 6

1.4 Outline of the thesis ......................................................................................................... 8

2 Research status .................................................................................................................... 9

2.1 Interpreting as interaction ............................................................................................... 9

2.2 Approaches to video-mediated interpreting .................................................................. 10

2.3 Video-mediated interaction ............................................................................................ 12

2.4 Accomplishing interpreting in the video-mediated environment ................................ 13

3 Theoretical framework ...................................................................................................... 15

3.1 Embodying institutional interaction ............................................................................... 15

3.2 Multimodality as a point of departure ............................................................................ 16

3.3 Specifying mediated interaction .................................................................................... 17

3.4 Reciprocity of perspectives and intersubjectivity in conversation analysis .................. 18

3.5 Media ideologies ............................................................................................................ 19

4 Methodologies .................................................................................................................. 21

4.1 Collecting data ................................................................................................................ 21

4.1.1 Video recordings as data .......................................................................................... 22

4.1.2 Interviews as data .................................................................................................... 28

4.1.3 Ethnographic data and processes .......................................................................... 29

4.2 Analysis ............................................................................................................................ 34

4.2.1 Conversation analysis .............................................................................................. 34

4.2.2 Discourse analysis .................................................................................................... 36

4.2.3 Combining conversation analysis and discourse analysis ...................................... 37

4.2.4 Transcribing for analysis ......................................................................................... 39

4.3 Ethical considerations ...................................................................................................... 46

5 Summary ............................................................................................................................ 49

5.1 Article 1 ........................................................................................................................ 51

5.2 Article 2 ........................................................................................................................ 52

5.3 Article 3 ........................................................................................................................ 53

5.4 Article 4 ........................................................................................................................ 54
6 Discussion and conclusion........................................................................................................55
6.1 Contributions ......................................................................................................................56
6.2 Theoretical implications .....................................................................................................59
6.3 Implications for practice ....................................................................................................61
6.4 Limitations and future research .......................................................................................64
7 References ..........................................................................................................................65
8 List of appendices ................................................................................................................73
1 Introduction

This thesis addresses a tension between perspectives on practices of video-mediated interpreting in medical encounters in Norway. This tension was recently asserted in an article in Vårt Oslo, a local Oslo-situated newspaper. The article (Velle, 2020) reported that worried healthcare professionals had contacted OMOD Center for Social Justice about their experiences with video-mediated interpreting. Video-mediated interpreting, the medical professionals found, was difficult to use when treating some psychiatric patients with whom they did not share a common language. In the article, a statement by the manager of an interpreting unit that provides interpreting services for hospitals in the Oslo area opposed the medical professionals’ concerns. The manager explained how video-mediated interpreting and telephone interpreting were ways to prevent infection during the COVID-19 pandemic. The interpreting unit had not received complaints or reports of deviations regarding interpreting assignments carried out through screen or telephone, he reported. Furthermore, he suggested that this practice would increase in the future since the interpreting service considered this to be a good alternative. However, according to OMOD, problems with video-mediated interpreting in patient treatment had resulted in patients losing access to medical treatment for several months. This news article contrasts the more common positive spin on predictions of how video technology can solve a range of problems in the provision of interpreting in Norway (e.g. Nesvik, 2018; Nilsen, 2017). Furthermore, the newspaper article asserts an important distinction, one that is not always that clear in discourse about video-mediated interpreting. The article demonstrates how video-mediated interpreting may refer to a way of providing services or to the actual accomplishment of video-mediated interpreting within specific settings, which, as such, has consequences for the accomplishment of the medical appointment. Combining conversation analysis and ethnography, the thesis considers the interactional accomplishment of interpreting in video-mediated settings. This is discussed in contrast to some of the media ideologies of video-mediated interpreting drawn upon by practitioners within a medical context and in government documents within the broader societal context. The thesis is article based: Three of the articles employ conversation analysis to explore how participants in hospital encounters accomplish interpreting in the video-mediated environment. The fourth article employs the concept of remediation to explore media ideologies in government documents and interviews with practitioners. In the following introduction to the thesis, I first address the ideal of good communication in healthcare and interpreting as a way to secure good healthcare services. Thereafter I topicalize some of the issues related to interpreting in healthcare, before I discuss how video-mediated interpreting is thought to solve some of these problems. Finally, I outline the aim and scope of the undergone study.
1.1 Good communication in healthcare

Good communication is important for quality of healthcare and patient safety (Gulbrandsen, Jensen, Finset, & Blanch-Hartigan, 2012) and good communication between patient or next-of-kin and medical professional is an ideal in medical practice (see for instance Gulbrandsen & Finset, 2014; Heyn, 2018). Combined with trust, good communication is considered to be crucial for the successful meeting between a patient and their general physician (Helse- og omsorgsdepartementet, 2013) and fundamental for the meeting between patients and healthcare professionals (Helsedirektoratet, 2011).

The right to information in healthcare services is determined by law. For instance, central paragraphs in the Patient and Users’ Act determine patients’ right to information (Pasient- og brukerrettighetsloven, 1999): § 3-1 determines that the patients have the right to participate in the implementation of healthcare services. Furthermore, § 3-2 states that the patient has the right to receive information that is necessary to gain insights into their health situation and the content of the healthcare. According to § 3-5 this information should be adapted to the recipient’s individual condition, such as age, maturity, experience and cultural and linguistic background, and healthcare personnel should, as far as possible, ensure that the recipient has understood the content and the meaning of the information. Moreover, § 4-1 determines that patients must consent to the healthcare, and in order for consent to be legally given, the patient has to receive necessary information about their health condition and the content of the proposed healthcare. According to The Norwegian Directorate of Health, this means in some cases that interpreting may be necessary in order to fulfil these requirements (Helsedirektoratet, 2015).

Corresponding to § 3-5, the Healthcare Professional Act (Helsepersonelloven, 1999) § 10 states that healthcare professionals are required to give information to those who have the right to receive information. As far as possible, this involves that the healthcare professional sees that the person receiving the information has understood the information and its consequences. The information must be adapted to the recipient’s individual condition regarding age, maturity, experience and cultural and linguistic background. This may involve using an interpreter or technical aids (Helsedirektoratet, 2018).

Addressing matters of interpreting in hospital encounters, Linnestad and Buzungu (2012) suggest that interpreting is a prerequisite for equal access to healthcare:

To provide qualitatively good interpreting services is a prerequisite for patients and next of kin with limited Norwegian proficiency to gain access to equal health services. Inadequate
Informing and shared decision making are realized through interaction. In institutional settings, participants engage in different types of interaction within different frames and purposes. For instance, in physiotherapy, talk between the patient and a physiotherapist can realize medical assessments, motivation, instruction, learning and much more. In a meeting, a medical professional and patient may carry out a range of different activities, for instance presenting, gathering information about, diagnosing and treating a concern (Robinson, 2003, 2012) or shared decision making (Landmark, Svennevig, Gerwing, & Gulbrandsen, 2017), to mention just a few. To return to the newspaper article, in some cases medical treatment is realized through talk. As such, interpreting is thought to facilitate medical activities in settings where medical professionals and patients do not speak the same language. The healthcare professionals in the newspaper article were concerned that they had not been able to realize the medical treatment with video-mediated interpreting. Before I return to the matter of interpreting as interaction, I will give a brief introduction to some of the complexities regarding providing interpreting within medical settings and how video-mediated interpreting is considered to solve several of these.

1.2 Video-mediated interpreting as a solution to a range of problems

The hearing note to a draft law about interpreting (Det Kongelige Kunnskapsdepartement, 2019) suggests that there are two main conditions that need to be taken into account when considering whether there is need for interpreting in a specific situation within the public services. One concern is the participants’ linguistic competencies: do they share a common language? The other is the matter of importance: is the encounter of such an importance that interpreting is necessary? In the healthcare setting, legislation and guidelines give some indications as to what is required. However, the documents provide few clues as to which criteria should apply when deciding whether to book an interpreter (Sagli, 2015). In an interview-based study carried out in 2013-2014 among Polish migrants in Norway, Czapka, Gerwing, and Sagbakken (2019) found that participants often received information regarding their health condition and treatment in a language they did not fully understand. Furthermore, they found that access to interpreting was limited or denied for instance because of reluctance of health personnel to book interpreting and overestimating the patient’s language skills (Czapka et al., 2019). Findings from The Norwegian Government’s Official
Norwegian Report (ONR) on interpreting in the public sector are in line with this, and suggest that there is under-usage of interpreting in healthcare services (NOU 2014: 8, 2014).

The linguistic situation is also complex. The ONR estimates that there is need for interpreting in more than 100 languages in Norway (NOU 2014: 8, 2014, p. 16). The need for interpreting in different languages follows fluctuations of migration. The diffusion of a language in the Norwegian context will change over time with migration flow. For instance, while Tigrinya and Amharic were among the five most common languages for interpreting at the Norwegian Directorate of Immigration in 2011 due to the large number of asylum interviews in these languages at that time, the two were the 12th and 20th most common languages in healthcare services at the same time (Linnestad & Buzungu, 2012, p. 8). More recent numbers from an in-house interpreting unit that provides interpreting for Oslo University Hospital, Akershus University Hospital, and Sunnaas Rehabilitation Hospital, show how Tigrinya was the fifth largest interpreting language in these hospitals during the years 2015-2017 (Buzungu, Nilsen, & Løfsnes, 2018, p. 15). This demonstrates how Tigrinya’s position in Norwegian society has changed during recent years. Polish is also a larger language in the hospital context, the second largest after Arabic (Buzungu et al., 2018), however, since immigration from Poland is mostly work-related, the high numbers of interpreting are not reflected within the institution for seeking asylum (Linnestad & Buzungu, 2012, p. 8). Structures such as migration flow and changes in the life situations of migrants after arrival will affect which languages are relevant in different sectors of public services.

General requirements regarding tendering and the increased need for interpreting services in the Norwegian public sector have caused several public offices to tender interpreting, some ranking price over quality. Some offices in the Norwegian public sector have difficulties controlling the quality of interpreting services under these conditions (NOU 2014: 8, 2014). This has been highly relevant within healthcare services, and has resulted in a lack of oversight of the quality of interpreting services in that sector (Linnestad & Buzungu, 2012; NOU 2014:8, 2014). In 2014, as a measure to improve the quality of interpreting, Oslo University Hospital established an in-house interpreting unit in collaboration with Akershus University Hospital and Sunnaas Rehabilitation Hospital to ensure quality in interpreting services. They found that one of the challenges in organizing interpreting services for hospitals is the high number of short assignments and how these are dispersed across different locations (Linnestad & Buzungu, 2012). Another difficulty is to get a hold of interpreters at short notice, which is often one of the needs in medical care. The goal was therefore also to develop a regional plan for screen interpreting and establish a competence center for interpreting in the health sector, and to improve emergency response plans so they could provide interpreting 24/7 (Løfsnes & Nilsen, 2017).
In the political, societal and linguistic setting described above, video-mediated interpreting is suggested to solve a range of problems. Use of video technology, similar to the telephone, affords interaction at a distance and as such provides a means or a tool to organize and provide interpreting services. Video technology is considered to be a way to provide interpreting to a linguistically diverse and scattered population and is suggested to be a better option to telephone interpreting (NOU 2014:8, 2014). It is no wonder that public servants – be it medical professionals, police officers or social workers – find the idea of getting access to qualified interpreters situated across the entire country, and, depending on how the interpreting services are organized, to do so on short notice, appealing. Early descriptions of the in-house unit indicated that as many as 40% of all assignments should be carried out as screen interpreting (Løfsnes, Buzungu, Buzungu & Hansen, 2016). In 2017, 93% of the assignments were onsite interpreting, 6% were telephone interpreting, and approximately 1% were on-screen interpreting (Buzungu et al., 2018). In the 2018 report, the interpreting unit reflected upon the initial ambitious goal made that 40% of all the assignments should be interpreted via video:

Low access to units for screen interpreting, little support in the videoconference field from Sykehuspartner [service provider of ICT and HR] and the lack of an integrated call-center solution from NHN [Norwegian Health Net, provider of ICT infrastructure] has resulted in the share of on-screen assignments to be 1% (…). The Interpreting Unit has of the fall 2017 been asked by the Southern-Eastern Norway Regional Health Authority to develop a regional strategy for on-screen interpreting. (Buzungu et al., 2018, my translation)

Several Norwegian hospitals have trialed screen interpreting during the last decade: In their project trialing screen interpreting, researchers at Health Finnmark aimed to see if this technology could improve the availability of interpreting services for Sámi-speaking patients (Furskognes, Eliassen, Molund, & Christiansen, 2013). Oslo University Hospital carried out a project exploring technological and communicative aspects of video-mediated interpreting (Hansen & Løfsnes, 2016). Haukeland University Hospital set out to improve the quality of interpreting services by considering the viability of video-mediated interpreting as an alternative method (Haukeland universitetssjukehus, 2017). Helse Førde (2018) also established a project to explore video-mediated interpreting as a way to gain access to qualified interpreters. All but the Oslo project aimed to test and possibly implement screen interpreting as a way of providing interpreting services. Since Oslo University Hospital already had some experience with screen interpreting, their aim was to explore financial, technological and interactional aspects.
Arguments for increased use of video technology often build on ideas of efficient service provision and video being better than the telephone due to the visual affordance of the media (see article 4). However, interpreting is not just a product easily delivered to the person who has ordered interpreting for a meeting with someone with whom they do not share a language. Participants in specific situations accomplish interpreting collaboratively in situ. Naturally, interpreting requires skills and competencies from the interpreter and requires the interpreter to carry out certain actions. The interpreters in this study, all have formal qualifications such as interpreter education. They need to have a professional skillset, they need a high level of proficiency in the relevant languages and, as a professional interpreter, adhere to ethical guidelines. However, in order for interpreting to be accomplished in interaction, other participants need to let the interpreter get the floor every now and then and given time to interpret, they also need to respond to the interpreter’s actions, such as repair initiation. As such, interpreting is interactionally achieved and, subsequently, collaboratively achieved by participants in interaction. The concerned medical professionals in the news article address an important issue: if the interpreting does not work, the treatment does not work. While interpreting is a service that is booked and billed, it is at the very same time an interactional activity that is collaboratively achieved by participants in interaction.

1.3 Aim and scope of the study

This thesis explores the accomplishment of interpreting within a specific interactional setting, hospital encounters, and a specific interactional space, the video-mediated environment. The study combines conversation analysis and ethnography. Three of the articles in this thesis employ conversation analysis, and examine the interactional organization of video-mediated interpreting in hospital encounters. The final article builds on the analysis of ethnographic data, documents and interviews, to explore media ideologies drawn upon by stakeholders in discourse about video-mediated interpreting in society and about experiences from the accomplishment of medical encounters.

Exploring video recordings of hospital encounters, the thesis explores the following problems:

1. How do participants in video-mediated, multilingual hospital interaction accomplish interpreting?
   - How do the participants organize and orient to their interactional space?
   - How do the participants organize actions relevant to interpreting in this interactional space?
In what ways is interpreting enabled, constrained or inhibited by participants’ configurations of the socio-material setting?

2. How do discourses on video-mediated interpreting, expressed in governmental documents and interviews with practitioners, relate to the interactional accomplishment of video-mediated interpreting?

The study builds on video recordings of video-mediated interpreting in hospital encounters and ethnographic data. The thesis’ four articles are listed below including information about their current state in the publication process.

Article 1

Creating space for interpreting within extended turns at talk

The article is written together with Professor Jan Svennevig. The article employs multimodal conversation analysis in the exploration of turn-organization in medical encounters, specifically tending to the temporary suspension of medical professionals’ longer turns at talk for the organization of interpreting. The article has been submitted to Journal of Pragmatics and has been accepted with revisions. Not all of the revisions have been made at the time of submission of the thesis.

Article 2

Recruiting repair: Interpreters’ displays of trouble in video-mediated environments

The article employs multimodal conversation analysis in the investigation of interpreters’ displays of trouble in video-mediated environments, and discusses the phenomenon in light of studies of repair and recruitments. The manuscript has not been submitted to a journal at the time of thesis submission.

Article 3

Invisible participants in a visual ecology: Visual space as a resource for organizing video-mediated interpreting in hospital encounters
The article employs multimodal conversation analysis in the investigation of how participants orient to the visual materiality of the setting, and how they use the visual ecology they create in and through the interaction to achieve the multilingual activity of interpreting in hospital encounters. The manuscript has been accepted by the journal *Social Interaction. Video-Based Studies of Human Sociality* for a special issue on the accomplishment of video-mediated interaction and is “in press”.

**Article 4:**

*Remediating the mediator: Media ideologies in policies and practices of medical interpreting*

The article employs discourse analysis and operationalizes remediation as an analytical concept to explore media ideologies drawn upon by stakeholders in government documents and interviews. The article supplements the interactional studies with a different approach. The manuscript has not been submitted to a journal at the time of thesis submission.

### 1.4 Outline of the thesis

The thesis is structured into six chapters. This first chapter gives background information and motivations for the study, before presenting the research questions. In chapter 2, I situate the thesis within the broader academic discussions to which it contributes. In chapter 3, I discuss theoretical themes and underpinning theoretical assumptions that guide that analysis. In chapter 4, I sketch out research design, methods and data collection processes as well as the analytical process and ethical considerations. In chapter 5, I present findings from the four articles. In chapter 6, I discuss the findings from each of the articles in light of each other, as well as the theoretical and practical implications of these findings. Finally, I review limitations of the study and make suggestions as to future research.
2 Research status

This chapter serves to place my work within broader academic discussions. The thesis explores encounters in Norwegian hospitals where medical professionals, patients and next-of-kin, do not speak the same language, and interpreting is provided via video. The thesis contains four articles addressing four distinct and clearly separate topics. As such, each article draws on and outlines previous research relevant to each of the separate topics. Three of the articles employ conversation analysis to explore video recordings of video-mediated interpreting, and the final article uses discourse analysis and the theoretical concept of media ideologies to explore discourse about video-mediated interpreting in Norwegian society. The three conversation analytic articles explore the interactional organization and accomplishment of interpreting within this specific mediated environment. The fourth explores how different stakeholders, through government documents and interviews with practitioners, medical professionals and interpreters, address several of these same topics related to video-mediated interpreting. This section of the thesis gives context to the undergone work and motivations of the study. In this chapter, I place my work within academic lines of discussion on interpreting, video-mediated interpreting and video-mediated interaction.

2.1 Interpreting as interaction

In language discordant meetings between medical professionals and patient, interpreting serves to make interaction possible between the speakers of different languages. Interpreting and translation studies have “enjoyed an eclectic tradition, in line with [their] interdisciplinary nature” (Angelelli & Baer, 2016a, p. 5), and have borrowed freely from related disciplines in the humanities and social sciences, drawing on for instance cultural studies, sociology, anthropology and applied linguistics (Angelelli & Baer, 2016b). Referring to signed language interpreting, Wilcox and Shaffer (2005), suggest that we have witnessed a “panoply of models” attempting to describe the interpreter and interpreting in a sufficient manner, for instance as a helper, a conduit or a bilingual-bicultural specialist, including various approaches such as sociolinguistics and interaction (p. 28). While one might disregard such models as the internal understanding of interpreting from within the field, perspectives on interpreting and interpreters’ work building on the models or conceptualizations occur also outside the field of interpreting. For instance, guidelines found in different arenas of the public sector for working with interpreters might reflect especially the idea of the interpreter being a form of conduit for the transformation of information between languages (Hansen, 2018; Li et al., 2017). Even interpreters’ codes of ethics may reflect the same ideals (Gavioli, 2016). The interpreter’s ethical guidelines suggest that the interpreter in their work should interpret “everything
that is expressed”, and not omit, change or add anything (Integrerings- og mangfoldsdirektoratet [IMDi], 2020). An interactional approach gives important insights into the accomplishment of the interaction, and can inform such understandings and guidelines (Hansen, 2018; Li et al., 2017).

Wadensjö’s (1998) seminal book *Interpreting as interaction* explored interpreting as just this, interaction. Interpreter utterances, according to Wadensjö (1998), bridge a linguistic gap and a social or interactional gap between two language users (p. 109). As such, interpreting is a matter of both translating and coordinating other participants’ utterances. If we consider interpreting to be an interactional activity (Robinson, 2013), this implies that interpreting is interactionally achieved by participants in interaction in situ. Conversation analysis, with its detailed study of small phenomenon and attention to the fine order of people’s actions, can give insights into how such activities are assembled (Sacks, 1984).

2.2 Approaches to video-mediated interpreting

In the introduction to one of the most recent books collecting studies on video-mediated interpreting, Brône and Salaets (2020) suggest that the field has already dealt with the pros and cons of video-based interaction for the interpreting profession. A range of different methodological approaches have provided insights into different aspects of video-mediated interpreting. For instance, questionnaires, surveys and interviews have given insight into aspects of video-mediated interpreting, such as stress factors for interpreters (Braun & Taylor, 2012b; Roziner & Shlesinger, 2010), interpreters’ onset of fatigue (Moser-Mercer, 2005) and interpreters’ views on their work (Brunson, 2018) and on their roles (Devaux, 2018), to mention just a few.

For some years now, studies on video-mediated interpreting have compared interpreting via video, including various constellations of participants and technologies, with interpreting via telephone and onsite interpreting. The various ways of providing interpreting, have been compared along different dimensions, for instance satisfaction (Locatis et al., 2010; Price, Pérez-Stable, Nickleach, López, & Karliner, 2012) and quality (Balogh & Hertog, 2012; Braun & Taylor, 2012a; De Boe, 2020). A study exploring interpreting in hospital encounters, for instance, found that when comparing telephone interpreting and video-mediated interpreting to onsite interpreting, patients rated all three methods the same, while doctors and interpreters rated onsite interpreting higher, and video-mediated interpreting higher than telephone interpreting (Locatis et al., 2010). Based on questionnaires that were filled out after video-mediated interpreting sessions in asylum interviews, Skaaden (2001) found that interpreters experienced more problems in the interaction than the other participants did. This suggests that moving the work to a video-mediated environment might affect
the work of the interpreter in a different way than it affects the work and activities of the other participants in the interaction.

In a survey comparing onsite interpreting, telephone interpreting and video-mediated interpreting, Price et al. (2012) found that interpreters considered all three methods satisfactory for conveying information, while scenarios with substantial educational or psychosocial dimensions (such as family meetings, consent for complex medical procedures, inpatient nursing teaching, hospital discharge instructions, physical or occupational therapy and case management/social work) had lower scores. In scenarios where the telephone did not suffice, interpreters’ considered video-mediated interpreting to offer improved communication.

A different dimension of comparison is quality of interpreting in the different environments focusing on the interaction. In their work on video-mediated interpreting in judicial settings, Braun and Taylor (2012a), Balogh and Hertog (2012) and Miller-Casino and Rybinska (2012) developed comparative studies at three different sites to explore the quality of the interaction in video-mediated interpreting compared to onsite interpreting. In order to assess and compare the quality of interpreting through different channels, the studies were designed as experiments using simulations in order to control factors. Problems in the interaction were labelled as inaccuracies, omissions, additions, linguistic problems, paralinguistic problems and synchronization problems (Braun & Taylor, 2012a). The labels were based on criteria for assessment of interpreting and categories for the analysis of non-verbal and visual communication (Braun & Taylor, 2012a, p. 102). Braun and Taylor (2012a) found that there were problems with turn-taking more than three times more often when the interpreting was carried out through video compared to face-to-face. They also found that omissions (159%) and additions (290%) were more frequent in the video-mediated settings (Braun & Taylor, 2012a, p. 92).

Based on a similar comparative design, De Boe (2020) compared video-mediated interpreting, telephone interpreting and onsite interpreting in medical encounters. The study was based on simulations carried out with video-mediated interpreting, telephone interpreting and onsite interpreting. The categories were partially pre-defined and were adjusted during work with analysis. The most frequently observed and problematic finding resulting from interactional behavior by all participants in this study was overlapping speech. While this occurred in all of the settings, problems were more easily fixed in onsite interpreting. Overlapping speech was also found to cause sound problems in the mediated settings, something that required more interactional work to solve (De Boe, 2020).
These studies, while providing novel insights into the interaction, do give rise to some important discussions. Braun and Taylor (2012a) suggest that it was a limitation to their study that it was based on simulations. Problems in the interaction did not unfold as they would in authentic interaction, as participants in the simulation would return to the storyline even in cases where something was altered in the translation. Similarly, in their study of mystery shoppers and real shoppers in interaction, Stokoe, Sikveland, Albert, Hamann, and Housley (2020) find that there are differences between the role-plays and authentic interaction both on the micro level and on the macro level. The behavior of the participants in the role-plays was unlike that of people in a real-life setting, meaning that the outcomes of the interactions are different.

Based on data from the simulations, Braun (2017) later explored one of the problem categories found during the quantitative analysis, additions and expansions, using conversation analysis. This analysis demonstrated how conversation analysis could give insights into the local situation in which a phenomenon occurred as well as a deeper understanding of the specific actions involved. The interactional perspective on video-mediated interpreting provides important insights into the local organization of the interaction. Prior studies on video-mediated interpreting that employ conversation analysis, are for instance studies from interpreted French courtroom hearings with remote defendants, which explore camera actions as interactive moves (Licoppe & Veyrier, 2017) and ‘chunking’ of longer turns at talk (Licoppe, Verdier, & Veyrier, 2018; Licoppe & Veyrier, 2020). Within a specific sign language interpreting service, Wannicke and Pleijert (2012, 2018) provide insights into the turn management of the video-relay services, and interpreters’ use of the headset as a resource in interaction.

This study sets out to investigate how participants in video-mediated, interpreted, hospital encounters, achieve the mediated activity – interpreting – in this specific interactional environment. This is achieved by exploring turn-taking, interpreters’ embodied displays of trouble and how participants orient to the visual affordance of the setting. While the investigation of how the participants achieve the medical appointment would be equally relevant, the thesis explores the achievement of video-mediated interpreting.

2.3 Video-mediated interaction

In video-mediated interaction, different technological devices with different affordances allow the participants to interact with each other at a distance (Arminen, Licoppe, & Spagnolli, 2016; Due, forthcoming; Hutchby, 2001). The technology serves not only as a tool that gives participants access to each other, but serves as a resource for the accomplishment of activities (Mondada, 2007).
Different perspectives on research on video-mediated interaction may address video-mediation as a premise for the interaction and the accomplishment of certain activities, or focus on how participants use the specificities of the setting as resources in interaction.

Some features of the mediated setting that might affect the effectiveness of participants’ use of semiotic resources are, for instance, delay and visual affordances. Delay is a fundamental part of the video-mediated environment, and can be addressed in different ways in the study of video-mediated interaction: as a premise for interaction (Schmitt, Gunkel, Cesar, & Bulterman, 2014); its occurrence in certain activities such as in second language learning (Rusk & Pörn, 2019); as a trouble source, for instance, in turn-taking (Ruhleder & Jordan, 2001); or as a resource in interaction (Rintel, 2013). Previous research has found that mediation may alter the affordances of resources in interaction. For instance, gaze and movement which are ordinarily used as attention-getting devices in interaction have been found weaker in the video-mediated setting (Heath & Luff, 1991, 1993; Hutchby, 2001). Similarly, referential activity has been found difficult in the video-mediated environment (Luff et al., 2003; Luff, Heath, Yamashita, Kuzuoka, & Jirotka, 2016) and participants in interaction have been found to develop new practices for the specific interactional space, such as in sign language teaching (Hjulstad, 2016).

Studies of video-mediated interaction have shown how participants accomplish certain activities within the video-mediated environment, such as meetings (Nielsen, 2019; Oittinen, 2018), medical consultations (Due, forthcoming; Pappas & Seale, 2009, 2010) surgical procedures (Mondada, 2007), teaching (Hjulstad, 2016; Mondada, 2007), second language learning (Rusk & Pörn, 2019), sign language interpreting in video-relay services (Warnicke & Plejert, 2012), and interpreting in judicial settings (Licoppe & Verdier, 2013; Licoppe et al., 2018; Licoppe & Veyrier, 2017, 2020). Studies within video-mediated interaction have also found that participants develop new practices in the accomplishment of activities, such as showings, noticings and instructions (Due, Lange, Nielsen, & Jarlskov, 2019; Licoppe et al., 2017; Rosenbaun & Licoppe, 2017; Zouinar & Velkovska, 2017). The technology affords participants with a device for creating an interactional space (Licoppe, 2015, 2017; Mondada, 2007) while at the same time being a technology used within a specific physical setting (Due, forthcoming; Licoppe & Veyrier, 2017; Mondada, 2007; Pappas & Seale, 2009).

2.4 Accomplishing interpreting in the video-mediated environment

The thesis assumes the position that interpreting is interaction and, as such, consults literature on the social organization of interaction in the analysis of video-mediated interpreting. Article 1
explores turn-taking, and literature on turn-taking creates the backdrop for this article. Firstly, turn-taking is considered in a general perspective (e.g. Sacks, Schegloff, & Jefferson, 1974). Then, specific for the consecutively interpreted interaction, we examine turns in consecutive interpreting (e.g. Angermeyer, 2007; Gavioli & Baraldi, 2011; Li, 2015). Interpreters’ turns in interaction are often found to respond to other participants’ turns, either by providing a rendition of other participants’ turns, or by producing other actions, such as asking for clarification (Gavioli & Baraldi, 2011, p. 211). The article is situated within conversation analytic literature exploring the production of turns in installments (Svennevig, 2018), or chunking of turns – as specific for the interpreter-mediated encounter (Licoppe et al., 2018; Licoppe & Veyrier, 2020).

Similarly, the second article is based within the conversation analytic tradition, exploring interpreters’ embodied displays of trouble with literature on the organization of repair (Schegloff, Jefferson, & Sacks, 1977) and recruitments (Drew & Couper-Kuhlen, 2014; Enfield, 2014; Kendrick & Drew, 2016). The third and final article within the interactional framework demonstrates how the visual socio-materiality of the setting is used as a resource in interaction. This understanding of the interaction displays both how interpreting is an activity with a specific interactional order and how this activity is organized in situ by participants drawing on resources they assume they have available for the organization of interaction. The mediated setting is understood as an interactional space or environment for the organization of interaction while the technology is at the same time a resources the participants can use and organize.

The fourth article explores media ideologies (Gershon, 2017) drawn upon by stakeholders based upon a different dataset than the three prior articles. While the three previous articles explore what participants in these specific settings do, focusing on the sequential organization of interaction, the fourth and final article examines how stakeholders talk about what they do. While this shifts the theoretical and methodological focus, the interaction is still considered to be interactionally achieved – both the interaction the participants talk about and the interaction between myself as the researcher and the participants.
3 Theoretical framework

This study explores the achievement of interpreting in a video-mediated environment situated within hospital encounters. Three of the articles explore this from an interactional perspective and therefore assume theoretical underpinnings of ethnomethodology and (multimodal) conversation analysis. The fourth article employs discourse analysis to explore media ideologies drawn upon by stakeholders in documents and in interviews. The interactional articles and the media ideologies article are based on different datasets; they draw upon different methods and build on different theoretical assumptions. However, they complement each other within the study as a whole.

The structure of this chapter follows the order of the analytical work, where the conversation analyses preceded the work with the analysis of media ideologies. As such, the chapter first addresses theoretical assumptions that the three interactional articles build on, the topics being: institutional interaction, multimodality, mediation, and reciprocity and intersubjectivity. Finally, the chapter addresses media ideologies as a theoretical concept.

3.1 Embodying institutional interaction

The institutionality of interaction is not determined by the setting alone; interaction is institutional depending on whether the “participants’ institutional or professional identities are somehow made relevant to the work activities in which they are engaged” (Drew & Heritage, 1992, pp. 3-4). The institutional character of talk is embodied through for instance turn-taking systems that are clearly distinct from ways that turn-taking is ordinarily managed in talk (Drew & Heritage, 1992, p. 25). Other systematic differences of institutional interaction may emerge when participants’ talk is organized within the constraints of a specialized turn-taking system (Drew & Heritage, 1992, p. 26). For instance, the range of options and opportunities for actions in interaction may be limited and the activities that are left may be specialized and re-specified. According to Drew & Heritage, institutionalized reductions and specializations are conventional (p. 26).

Heritage and Clayman (2010, p. 34) suggest that there are three dimensions to institutional interaction: 1) participants in institutional interaction often orient to certain goals tied to their identities relevant to the institution; 2) there are certain constraints on the interaction considering what is treated as acceptable contributions to the business at hand; 3) there are inferential frameworks and procedures particular to the institutional context that are associated with the interaction. Drew and Heritage (1992) suggest that these institutional conventions are associated with various participation frameworks with associated rights and obligations as well as with different patterns of opportunities to initiate and sanction interactional activities.
3.2 Multimodality as a point of departure

Within the inquiry of social interaction, interest has been paid not only to the verbal organization of interaction but also to the embodied organization of interaction. Within conversation analysis (CA), multimodality includes all relevant resources that are mobilized by participants to build and interpret the public intelligibility and accountability of their situated action: grammar, lexicon, prosody, gesture, gaze, body postures, movements, manipulations of artifacts, etc. (Mondada, 2018, p. 86). Mondada (2018, p. 86) lists a number of features that characterize multimodal resources. First, multimodal resources relate to the organization of action but do not make sense of it. Second, multimodality as a notion includes linguistic and embodied resources, in principle treating them equally, without supposing the priority of one over the other. Third, multimodal resources are not limited to conventional resources, such as grammar and types of gesture, but include situatedly occasioned resources depending on the local characteristics of the ecology of an activity – which both enables and constrains what participants treat as meaningful resources. Fourth, multimodal resources are characterized by a temporality that combines both successive and simultaneous lines of conduct. Fifth, they are combined in various configurations, depending on the activity, its ecology and its material constraints.

In the study of interaction, researchers have long accounted for gaze and bodily orientation in the organization of interaction (e.g. Goodwin, 1981; Heath, 1986). Studies of multimodality within Ethnomethodology and Conversation Analysis (EMCA) have taken on different perspectives during the last decade (Mondada, 2018, p. 87), and Mondada distinguishes between different approaches to multimodal analyses. Some analyses focus on the organization of specific settings, often institutional, in order to understand complex spatio-material contexts of actions (e.g. Heath & Luff, 2000). Some focus on multimodality in relation to the organization of turns, sequences and actions in order to understand how action is made intelligible and accountable (e.g. Goodwin, 2000). Finally, she suggests that some approaches integrate studies of grammar, syntax and lexicon in interaction (e.g. Keevallik, 2013). This move toward multimodality and the subsequent consequences for conversation analysis have been discussed extensively in special issues of the *Journal of Pragmatics* edited by Deppermann (2013) and Hazel, Mortensen, & Rasmussen (2014), raising important issues for further discussion.
3.3 Specifying mediated interaction

This study explores the achievement of interpreting in a video-mediated environment, and as such defines the interaction as video-mediated. For Arminen et al. (2016) “[m]ediation refers to the way the particular organization and unfolding of activities in definite material settings might constrain or enable the production of particular forms of accountable responses and shape the criteria to assess their relevance” (p. 293). In their discussion on mediated interaction, Arminen et al. (2016) approach different configurations of technology in use. At one extreme, “technology seems to create the venue in which and through which interaction appears to take place”, while at the other, it is classed as a medium or device to be a resource or tool for interaction “in a complex ecology involving a variety of interactional resources” (Arminen et al., 2016, p. 290).

Arminen et al. (2016) suggest moving beyond an approach to technology-as-context and to show how, on the basis of data-driven empirical studies, technologies and media can be shown to be both relevant and consequential with respect to the sequential organization of interaction (Arminen et al., 2016, p. 292). As a point of departure for the evaluation of mediated communication, Dourish, Adler, Bellotti, and Henderson (1996, p. 34) suggest that face-to-face communicative behaviors are not always the appropriate baselines. Moving away from this perspective, they claim, allows the exploration of important, intrinsic properties of video as a communicative medium in its own right. They furthermore suggest that a range of communicative practices adapted to the medium arise over time as familiarity with the medium increases. The practices are related to the people and work practices involved and must be studied in long-term use. Finally, they suggest that the influence and importance of the technology extend beyond the persons who engage directly with it, and beyond the immediate context and environments. Corresponding with the first suggestion, Arminen et al. (2016, p. 297) suggest that direct comparisons of patterns of sequential organization across settings risk leading to simplified, flawed insights that fail to grasp the participants’ lived sense of action. Interaction practices that are very recent or even not yet established, tend to be contrasted with similar practices observed in more familiar settings, such as face-to-face interaction, in order to establish some positive or negative consequences (Arminen et al., 2016, p. 291).

Video technology provides participants in interaction with the possibility of mutual visual access to one another. Unlike interaction with co-present participants, video-mediated interaction is asymmetric (Arminen et al., 2016, p. 297). Only parts of the participants’ bodies and surroundings are made visually accessible to others, which in turn constrains the participants’ possibility to visually monitor co-participants. In addition to participants having trouble with references in video-mediated interaction, orientational shifts and gestures – while perhaps being captured by the camera
and thus being displayed on the screen – may be disconnected from relevant features of the environment to the remote participant (Luff et al., 2003). Referring to how participants are unable to design conduct so it is sensible and recognizable to the remote co-participant who has only limited access to the environment in which the action is produced, Luff et al. (2003) describe the conduct as fractured – both from the environment where it is produced and from the environment in which it is perceived (p. 55).

This study explores the achievement of interpreting in a video-mediated environment. While video technologies emulate co-presence at a distance and allow distant parties to see each other, video-mediated environments can be both enabling and constraining to the accomplishment of recognizable interactional moves (Arminen, Licoppe, & Spagnolli, 2016).

### 3.4 Reciprocity of perspectives and intersubjectivity in conversation analysis

In their treatment of embodied actions such as references in the video-mediated environment as fractured from the environment in which the references occur, Luff et al. (2003) refer to Schutz and the reciprocity of perspectives. Luff et al. (2003) found that participants in video-mediated interaction often presupposed that what they saw corresponded with how the co-participant saw and viewed the environment (p. 55). In Schutz’s terms the participants presupposed a “reciprocity of perspectives” and an “interchangeability of standpoints” only to discover that the scene was not available to the co-participants in the way that they assumed (Luff et al., 2003, p. 55). Schutz’ (1953) concept of *reciprocity of perspectives* suggests that although people have individual perspectives, in commonsense thinking “the world taken for granted by me is also the world taken for granted by you” (p. 8). In Knoblauch’s (2013) words, reciprocity is the form of intersubjectivity that allows understanding:

Reciprocity cannot be reduced to reciprocating in the sense of exchange theories of action. It must, instead, be considered as a basic principle implied in more specific forms of interaction; such as exchange, transaction, and reciprocation, without being reduced to these forms. Rather, it is already implied in communicative action. Instead of being only “oriented” towards someone else, communicative action also implies some kind of anticipation of the other’s understanding of one’s action. It is the interpretation of the course of action by another actor (who is or may be oneself) which lies at the root of Schutz’ notion of reciprocity. (p. 331)
Commonsense thinking overcomes differences in individual perspectives as a result of two basic idealizations: 1) the idealization of the interchangeability of standpoints; and 2) the idealization of the congruency of the system of relevances (Schutz, 1953, p. 8).

In conversation analysis, the current utterance displays a hearing or analysis of the utterance to which it responds, and this is “publicly available as the means by which previous speakers can determine how they were understood” (Heritage, 1984, pp. 254-255):

> Conversational interaction is structured by an organization of action which is implemented on a turn-by-turn basis. By means of this organization, a context of publicly displayed and continuously updated intersubjective understandings is systematically sustained. It is through this ‘turn-by-turn’ character of talk that the participants display their understandings of ‘the state of the talk’ for one another. It is important to note that, because these displayed understandings arise as a kind of byproduct or indirect outcome of the sequentially organized activities of the participants, the issue of ‘understanding’ per se is only rarely topicalized at the conversational ‘surface’. (Heritage, 1984, p. 259)

The same sequential structure that is the grounds for understanding in interaction is the foundation for the researcher’s understanding when conducting analyses. I will return to this in the methods section.

### 3.5 Media ideologies

Media ideologies are a part of broader focus on semiotic ideologies (Gershon, 2010b; Keane, 2003), and focus on the semiotic ideologies of voice, body, image and sound (Gershon, 2010b, p. 284). According to Gershon (2010b), media ideologies function in a similar way to language ideologies; she understands the concept as people’s beliefs, attitudes and strategies about the media they use. Media ideologies involve assumptions that people hold about how the choice of a medium over another impacts the accomplishment of communication (Gershon, 2010a). Media ideologies as a term can sharpen the focus on how people understand both communicative possibilities and material limitations of a channel, and how they conceive of channels in general (Gershon, 2010b). Within Gershon’s concept of media ideologies, remediation is one of several themes. There exists a comparison, tacit or explicit, between a particular medium and all others available within a media ecology (Gershon, 2017, p. 20).

The concept of remediation addresses how people put media ideologies and media practices associated with a channel of communication in dialogue with ideologies and practices of other
channels (Bolter & Grusin, 1999; Gershon, 2017). Within Bolter and Crusin’s concept of remediation, all technologies lie on a continuum between immediacy and hyper-mediation, that is, between a perceived lack of mediation and excessive attention to mediation (Gershon, 2010b, p. 288). In this thesis, media ideologies have been operationalized as an analytical device through discourse analysis. This is elaborated on in the methods section.
4 Methodologies

This thesis draws on conversation analysis and ethnography. Conversation analysts have drawn upon ethnography in different ways and conversation analysis has been applied in ethnographies. For instance, Mondada (2012, p. 38) suggests that ethnography can provide a proto-analysis for the making of video recordings, which makes it possible to choose what and how to record. Before even placing a recording device in the room, the researcher has made a number of choices regarding, for instance, perspective, recording device and more (Mondada, 2013, pp. 39-41). The initial idea for this project was to collect video recordings of video-mediated interpreting in hospital encounters. These would undergo multimodal conversation analysis, and in addition the materials would be augmented with ethnography (Heath, 2004, p. 273). During the work with the project, the ethnographic leg has grown. In this chapter, I describe the methodological work in the PhD project, addressing data collection, data sources and ethnographic processes, analysis, transcription, and finally ethical considerations in the project.

4.1 Collecting data

In this section, I discuss the data sources and processes of collecting data. The project has been carried out in collaboration with an in-house interpreting unit providing interpreting for three hospitals. The data sources in this study include video recordings of video-mediated interpreting in hospital encounters and ethnographic data sources such as fieldwork, texts and interviews. My engagement with the field has been spread over a long period of time where collecting video recordings of video-mediated interpreting has been central to the process.

Table 1

<table>
<thead>
<tr>
<th>Activity</th>
<th>Time</th>
<th>2015</th>
<th>2017</th>
<th>2018</th>
<th>2019</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fieldwork engagement</td>
<td>Q4</td>
<td>Q2</td>
<td>Q3</td>
<td>Q4</td>
<td>Q1</td>
</tr>
<tr>
<td>Interviews</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recordings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When I began the work, I already had video recordings of five hospital encounters from a prior project (Hansen, 2016, 2018). Two of the recordings are of onsite interpreting, while three are with video-mediated interpreting. During the PhD project, I collected new video recordings making adjustments according to prior experiences, accumulating video recordings of eight new encounters with video-mediated interpreting. In addition, I conducted 30 interviews with 20 interpreters, 14 medical professionals and 4 other stakeholders. Some of the interviews were conducted with several
participants present. The table below shows the number of interviews and the number of participants. The following sections address video recordings as data and interviews as data before I discuss some of the ethnographic processes involved.

<table>
<thead>
<tr>
<th>Interpreters</th>
<th>Medical professionals</th>
<th>Other stakeholders</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 interviews</td>
<td>12 interviews</td>
<td>4 interviews</td>
</tr>
<tr>
<td>20 participants</td>
<td>14 participants</td>
<td>4 participants</td>
</tr>
</tbody>
</table>

**4.1.1 Video recordings as data**

As the project set out as a conversation analytic project, video recordings of video-interpreted hospital encounters have been treated as a core data source throughout the process. This part of the dataset consists of video recordings of 13 hospital encounters: 2 with the interpreter present and 11 with the interpreter participating via video technology. The interpreting is carried out in seven different languages in addition to Norwegian (see Table 3).

**Table 3**

<table>
<thead>
<tr>
<th>Language</th>
<th>Onsite</th>
<th>Video-mediated</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albanian</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Arabic</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Bosnian/Croatian/Serbian</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Mandarin</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Polish</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Thai</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td>11</td>
<td>13</td>
<td></td>
</tr>
</tbody>
</table>

Video recordings provide the researcher with a medium for preserving the data’s relevant features for later analysis, while also functioning as a *configuring* device (Mondada, 2009, p. 2). As such, video recordings not only document what happened at a specific event, but are reflexively produced, causing the recordings to structure and arrange the data of the analysis (Mondada, 2009). According to Mondada, the recordings thus shape the data and give them particular orderliness and meaning. These data are constructed with consideration to their purpose: to investigate the interactional accomplishment of video-mediated interpreting in hospital encounters. As such, the video recordings are not just recordings of data ready to be harvested, but through choice of perspectives, angles and placement of cameras, the researcher creates specific data.
The perspectives from which the video recordings were made have analytical consequences. This round of data collection built on the experiences from my prior work (Hansen, 2016, 2018). The first recordings I made were from the interpreter’s perspective, with a camera placed in the interpreting studio capturing the interpreter and the screen where the other participants were visible. While working with this data, I experienced that the data such as it was constructed challenged some conversation analysis’ analytical assumptions. For instance, silence at the change of speakership between the two sites in the interaction, for example between the interpreter’s rendition of a question and another participant’s response, might be unusually long (Hansen, 2016, p. 47). Conversation analysis may recognize that a long silence, a gap, before the response to a question indicates something about the responder’s stance to the question or to the response itself (Lee, 2013; Stivers & Robinson, 2006). However, the systematicity of the increased length of silence gave rise to analytical questions about how the interaction unfolded on each side of the technology and consequently to the participants at each site. Was the long gap at the interpreter’s studio shorter at the ward? Similarly, finding that the interpreter’s embodied displays of trouble did not receive displays of uptake from participants at the ward (Hansen, 2016, p. 65) made me curious of how the interpreter was represented in the setting and how the screen representing the interpreter was used as a resource in the interaction. Data from only the interpreter’s perspective would not suffice if I were to further investigate these or similar matters. For the PhD project, my aim was to get video recordings from both the interpreter and the hospital ward’s perspectives. This accounted for the fact that the interaction was not just something happening within the space the participants created using video technology, but that it occurs within a more complex setting drawing on resources in the mutually available space as well as the participants’ physical surroundings. Table 4 gives an overview of recordings according to camera placements and sum of recorded minutes. The interpreting was either onsite or video-mediated, and the recording cameras were placed either in the ward, in the interpreting studio or both.

<table>
<thead>
<tr>
<th>Camera placement</th>
<th>Number of recordings</th>
<th>Sum minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Onsite, ward</td>
<td>2</td>
<td>84</td>
</tr>
<tr>
<td>Video-mediated, interpreting studio</td>
<td>3</td>
<td>124</td>
</tr>
<tr>
<td>Video-mediated, ward</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>Video-mediated, ward and interpreting studio</td>
<td>7</td>
<td>153</td>
</tr>
<tr>
<td>Sum</td>
<td>13</td>
<td>392</td>
</tr>
</tbody>
</table>

The recordings include meetings from different wards, both meetings with inpatients who may have just arrived at the hospital or who have perhaps spent several weeks there, and polyclinical appointments where outpatients come in for brief consultations. The meetings with onsite
interpreting were carried out in different physical environments than the ones with video-mediated interpreting. One was in the patient’s room, the other in an examination setting with special equipment. Video-mediated interpreting as it was organized in the hospitals while I was collecting data was often done in meeting rooms or examination rooms with specific videoconference equipment. The medical professionals had varying experience using the technology and with video-mediated interpreting. Some of the interpreters in the study were carrying out video-mediated interpreting for the very first time, while some had done this many times before. Below are illustrated examples of what the video-mediated situations looked like.

Figure 1: The interpreter’s perspective

The interpreter participates in the interaction from a remote location, an interpreting studio equipped with a desktop system for videoconferencing. The setup uses about as much space as a personal computer and is designed for videoconferencing from an individual workspace. It has a screen, camera, loudspeakers, microphone, and a control panel which allows the interpreter to make adjustments to sound and image settings. The room is also equipped with a telephone for telephone interpreting. In the setting illustrated above, the interpreter has set up her workspace with a personal computer. The interpreters also have a pen and a notepad available and take notes during the session.
In the hospital meeting room illustrated in Figure 2, the large screen at the end of the meeting room table serves as both a screen for the computer in the room and as a screen for videoconferencing. In addition to a wide-angle camera above the screen, a multidirectional microphone is connected to the system by cable and can be placed on the table, closer to the participants. The room is furnished for videoconferencing and the camera can capture all the participants around the table depending on chosen settings. Adjustments to the videoconference system and technical settings can be made using a control panel on a small table next to the screen. Some wards are equipped with smaller systems (see Figure 3) similar to the desktop unit in the interpreter’s studio.
Figure 3: A polyclinical setting

The polyclinical setting illustrated in Figure 3 is in a multipurpose room used both for specific types of examinations and appointments requiring interpreting. A computer for access to journal systems and a desktop videoconferencing unit are available on the desk. The system is designed for videoconferencing from an individual workspace, and has a narrower camera frame than the more advanced system above.

Making recordings at two sites simultaneously required careful planning and logistics (some of which are discussed under ethical considerations). In the interpreter’s studio, I would position the camera so it would capture both the screen and the interpreter’s upper body as shown in Figure 4. Figure 4 is from a session where I was testing camera placements before a recording. It shows how the recording captures me both directly by the video camera and by the camera on the desktop which projects my image onto the screen. At the wards, I would have to make choices regarding camera position quickly, often as the participants were taking their seats. I would not always have access to the rooms in advance, and in cases where I did, I would not necessarily know where the participants would be seated. In order to make recordings from both sites, I would set up the cameras in the interpreter’s studio in advance and get help to start the recording and to collect and take care of the camera until I could collect it after the meeting.
A third option would be to work with screen captures from each site. This would require involvement from others, as the infrastructure is not readily available to make recordings. More importantly, with initial insights into how the participants organize the interaction, I knew it would be interesting to see not only what was captured by their videoconference systems and displayed on their screens, but their immediate physical surroundings as well.

I was present at the ward during all the sessions that were recorded for the doctoral research project. Being present is not just a matter of observing but, furthermore, turns me into a participant. Although I attempted to assume the role of an observer, participants included me in jokes and I am visible on some of the recordings. This was a methodological issue during the initial work on an article on opening phases. My presence would affect the opening phases, as I would need to ask consent and have a physical presence in the room. I had to reframe the article for other reasons, but was happy to avoid the problem of my own presence. On the other hand, my presence allowed me to make adjustments to camera settings during the session if participants shifted positions; this reminded the participants that the interaction was being recorded. Some of the recordings I made in my work are with the interpreters onsite. In these cases, I placed the camera in the room and then left. In these scenarios, the participants would glance at the camera and even express uncertainty of how or when to end the meeting as if they were perhaps waiting for me to return and stop recording. Either way, the camera is not neutral and unnoticed but a part of the setting whether I was present or not. While we might attempt to and wish to become invisible as researchers, in some ways I find it sincere to be present and remind the participants that they are being recorded. I will return to
matters of identifying meetings relevant for recording and obtaining consent in 4.4 Ethical considerations.

4.1.2 Interviews as data

During the scope of this study, I have carried out 30 interviews with 20 interpreters, 14 medical professionals and 4 representatives from hospitals, government authorities and the interpreter education. The interviews with the practitioners are semi-structured and follow a short guide with five questions\(^1\) (see appendix B). The first interview I made was an open conversation with 5-6 interpreters. It was an informal conversation in an open space and interpreters would come and go during the talk.

\[\text{Table 5: Interviews}\]

A total of 37 participants engaged in the 30 interviews I carried out. During my field engagements as a researcher, I would ask practitioners I met if they were willing to participate in interviews. In some cases, I would have to ask the organization first if I was permitted to ask employees to interview them. This was not because of the content, but because of time and resource management. Recruiting participants for interviews went quite easily. The interpreters who are interviewed have various interpreting languages, and therefore have different experiences from their work (see further discussion on this in article 4), while the medical professionals work in a range of wards. The only requirement for participation was that the participants had some experience from video-mediated interpreting. Toward the end of the data collection, process I developed the interview form and interviewed four stakeholders from other parts of society (see appendix C).

---

\(^1\) This was done based on experiences from informal talks and consulting a peer and fellow interpreter, Hilde Fiva Buzungu, and my supervisor, Jannis Androutsopoulos.
The content of interviews is a product not only of what the participants have told me, but are a result of my actions in our encounters as well.

A research interview, for example, usually forms an environment expressly designed to elicit the respondent’s, not the interviewer’s, narrative. Interview circumstances, format, and protocol dictate that the interviewer does the asking, while the respondent provides the story. Narrative topics are predesignated, and storylines at least partially predetermined. (Gubrium & Holstein, 1998)

In the interview form I used in meetings with practitioners, the questions were constructed to elicit narratives. I return to the interviews in section 4.2 Analysis.

### 4.1.3 Ethnographic data and processes

This section addresses the processes that have resulted from my time in the field. Since the repositioning of myself as a researcher has been of importance, I will start by situating myself in order to give an impression of my point of departure. I am a sign language interpreter, although I have not been a practitioner for a while. I have held different positions within the field of interpreting, within both spoken language and sign language interpreting. I have taught interpreting students as well as medical practitioners and students. I have practical experience from video-mediated interpreting. I conducted my master’s thesis on video-mediated interpreting while I was working as an advisor in a hospital. Furthermore, through my previous work at the Oslo University Hospital, I was project manager for a project on video-mediated interpreting (Hansen & Løfsnes, 2016). On February 28, 2017, I left my job as an advisor to begin my work as a doctoral fellow at the University of Oslo. Shortly after, I returned to the field with the purpose of being a researcher. While my time in the field generated data points such as video recordings, interviews, documents, field notes, observations and reflections, it has also resulted in a process of repositioning myself as a researcher. This has allowed me to gain access to wards and to gain new insights and perspectives.

Blommaert and Jie (2010) describe ethnographic fieldwork as not only data collection, but also as a learning process. The researcher arrives as an outsider with limited knowledge and “gradually moves from the margins of the social environment to a more central position” (Blommaert & Jie, 2010, p. 27). In some cases, the researcher has experience from the field. Blommaert and Jie (2010) describe how in educational research, for instance, researchers might have a significant amount of experience as teachers. However, “when that teacher turns into a researcher s/he stops being a teacher” (Blommaert & Jie, 2010, p. 27). In the following, I will
describe the process of gaining access to video recordings of video-mediated interpreting and how this process resulted in new insights.

**Gaining access**

The field I was trying to gain access to is distributed over several organizational and physical sites. As such, this is a multisite project. My aim was to establish contact with enough wards so I could be certain that I was able to acquire video recordings of video-mediated interpreting in hospital encounters. During my first round of data collection carried out in 2015 (Hansen, 2016), I collaborated with one hospital to make the recordings. The employees were informed about the collaboration, and I got the chance to inform them about the project. This initial contact with the hospital was crucial for data collection. In the second round of data collection, carried out in 2017-2019, I began by identifying and contacting wards within several hospitals that had the technology for video-mediated interpreting, in order to identify which might be willing to collaborate. I had meetings with, for instance, the division for equity in healthcare in one hospital in order to identify wards and gain contacts within the hospital. I was interested in identifying wards which had the technological equipment, and if they were using it and therefore might have occurrences of the type of events that I was interested in recording. Furthermore, they would have to be willing and interested in collaborating with me. Across several hospitals, I met with staff at several of the wards where they had technology for interpreting, informing them about the project and asking about their experiences. Not all the wards were willing to or had the capacity to meet with me. Fortunately, several were interested in collaborating.

During the process, I would meet with persons from different wards. Some would tell me that they hardly used the technology for interpreting. This was due to a range of different reasons. For instance, some had experienced technical issues and never got the equipment up and running again. One ward told me that they had stopped using the videoconference unit after a reorganization at the ward. They now used two computer screens in their daily work and in this new setting, the videoconference unit simply took too much space on their desk. Another ward let me know that they did not currently use the technology but were interested in getting started. They asked if I could perhaps come and teach them how to do this. This to me was a dilemma. I was interested in gaining contact with them and learning about their situation, however, I would prefer to avoid teaching them something and later making recordings of them doing what I had instructed them to do. Following is a short extract from my field notes during this period.

*To get to the assignments, I follow the technology. When I talk to those who have the technology, they tell me that they don't use it, or that it doesn't work, that they want to use it, that some are*
hesitant but that the hospital wants them to use it. We end up talking about what they should do to use it and if they can have more training. I ask if I can come to the training. They ask if I can do the training. I want to help out. I have to think a bit more about this. (...) I have a strong sense that I am changing the field just by being present.

Field notes, January 2018

On the one hand, I wanted to contribute to the field. On the other hand, I wanted to avoid positioning myself as an expert and consequently affecting the data. I wanted to avoid setting restrictions or giving advice at this point, which might at a later point become relevant in analysis of video recordings. In this case, I asked representatives from the hospital who worked with interpreting to join us. I facilitated a workshop, and attempted to let the hospital representative be the authority.

I have knowledge and experience, and when I am in the field, they ask me about this. Some things I know something about, other things I know little about. This is what I have to realize and accept.

Field notes, January 2018

I wished to limit how I affected the field during my time in it. Though, seeing as I have worked in the field and I was now revisiting with the lens of a researcher, this was not entirely avoidable. However, I could try to limit the consequences. While this situation did require some reflection, in the end it was not relevant to make recordings at this ward. The wards where I have made the recordings are wards that had already used video-mediated interpreting for a while and who had developed a way of doing it.

**Being an insider and an outsider**

While I was re-entering the field, I attempted to be attentive of possible consequences of being there as a researcher so soon after having left as an advisor. Having insights and some knowledge about the involved organizations was useful in finding contacts and in accomplishing the project. Furthermore, this gave valuable insights for the accomplishment of the project, which would otherwise require extensive ethnographic research to achieve. However, this raised some concern regarding my position. Some of the wards I was visiting were even wards I had visited during my earlier work as an advisor. Through interviews with some of the employees in these wards, I could see that my previous role as an advisor there was not necessarily as prominent to the healthcare professionals as it was to me. Although the medical professionals did not necessarily recognize me as a representative from the unit, I am well aware that several of the interpreters did. Some of my work when gaining access to situations and asking for their consent has been to try to dissociate myself from both my previous work and from any normative perspectives to interpreters’ practice which I might be assumed to hold as an advisor in the hospital. I believed that this would be
important not only to gain the interpreters’ trust so they might let me make recordings, but also in
the interviews I would be making with them.

One challenge is to balance being an insider and an outsider for ethical reasons, another is
balancing the insiders’ perspectives with the explorative stance of the researcher. Trying to balance
knowledge from my previous work as an interpreter, an advisor and an educator, with the
investigative and curious stance of a researcher.

And although I cannot know of any other perspective than my own, I am an interpreter at heart. So,
then it might be easier to see the interpreter’s perspective and the organizational aspects from the
interpreting unit’s side. Because I recognize them. I think the most important part of my future work,
is to avoid assuming that I know something about the other participants’ perspectives on the basis
of my own experience. While my experiences might help me gain access to the field, my experiences
do not make out the foundation of the study. They give me a basis for where to start my work.

Field notes, January 2018

During the process of repositioning myself in this field, I became increasingly aware of the
perspective I brought into the work. One goal was to gain enough distance to be able to observe
what is behind the intuitive actions of a practitioner. The insider’s perspective was useful for gaining
contact and trust with the participants, while assuming an outsider’s perspective was crucial to
assuming a more exploring and open position as a researcher.

**Presence in the field**

Similar to being present when making video recordings, my presence during fieldwork was not
neutral. When spending time in the interpreting unit and in the hospitals, I would talk to people.
When I talked to people, they would offer their point of view on video-mediated interpreting and
the use of technology for interpreting. I got the impression that my mere presence would make the
topic of video-mediated interpreting relevant and invite people to share their thoughts on the subject.

I sat talking with several interpreters first. About everything and nothing. One asked me what I was
working on. I told him about the project. He told me about his experiences. He felt onsite interpreting
was better. Something was lost when he was not present. I asked what was lost. He said it was
difficult to put into words what was lost.

February 2018

I was aware that my presence affected the sites I was visiting. However, importantly, being present
at the wards, not just through formal meetings informing staff about the project and pursuing
collaboration, but spending time there, gave me new insights. Entering a building, walking through
corridors, sitting in a waiting room, absorbing the atmosphere and realizing how incredibly different
it is to be present somewhere with everything that entails rather than gazing at a part of a room and
some people on a screen. There is no smell, the sounds are distorted by technology and there are no locked doors. When working through the technology, you do not need to walk through the long corridors or ask someone for directions.

During a site visit in March 2018, I was beginning to become aware of this other perspective. After having met with the head of the ward and talked about my project and whether they used the technology or not – they had used it a little, not much – I asked if I could see the room where the unit was placed. Until I began my work on the PhD, I did not know what things looked like “on the other side”. Until this point, I had seen all the interaction from the interpreter’s perspective, that is, through technology. This might be a perspective natural and simple for me to assume since I have a background from interpreting. Discovering that what the interpreter has access to through the technology is different than what the participants have access to in the ward, I realized that I had emphasized the interpreter’s perspective in previous recordings. This realization, among others, motivated my choice of methods for the project by factoring in the importance of looking at not only what the interpreter does in the interaction, but also what the other participants do. This was important in order to gain an understanding of the nature of video-mediated, interpreted, hospital interaction. I had not really understood how important this understanding of the wards’ perspectives was up until my work with the PhD. The following is an extract of my notes from when I asked the head of the ward if I could see the unit:

Where is the unit?
I asked if I could take a picture of their unit in the room if it was vacant. She said that this should be possible. If she found the unit. She wasn’t quite sure where it was placed. I wondered if she knew if it was taken into use. She wasn’t sure about that. They had two units at this ward. We agreed that I should bring my stuff since we were going to an area close to the exit and since we had already talked about everything we needed to talk about. I brought my backpack, notebook and jacket and then we went. We went to the floor where I had entered the building earlier. We went into the emergency wing. We continued to the receptionist’s desk. The receptionist directed us further. We arrived at a hall leading to two rooms. The door to each room had a small window. On the one window there was a note almost covering the entire window. It said something like “functioning interpreting studio”. Behind the door there was a seating arrangement with a couch and a chair, and in the corner there was a table with a videoconference unit. I was about to ask how they used to sit when they used it. I began asking, but remembered that she wouldn’t be able to answer. I told her that that was what I was about to ask. She let me take a picture. She talked about using the equipment more. I felt that I might seem to be an ambassador for on-screen interpreting just by being present and talking about it and that that really wasn’t my intention. But ute av syne, ute av sinn (out of sight, out of mind), or so they say. We finished up, and I left.

Field notes, March 2018
I left the ward, curious of how they use the room. What does the placement of technology and seating arrangement in a room afford for the people using it? Seeing the arrangement of seats, I anticipate some problems will occur both concerning what the interpreter will be able to hear if a participant is seated with their head directed away from the microphone, and concerning the interpreter’s visual access to the interaction. I was curious of how they would organize that and what experiences they might already have had.

I did not make recordings at all the sites I visited. However, spending time at different sites has given me important insights into the complex materialities of the settings at the wards. I have gained insights into the multitude of different perspectives in the organization of video-mediated interpreting in hospitals. Furthermore, during the process, I have experienced a change in my own perspective as a researcher.

4.2 Analysis

The two analytical methods used in this thesis are conversation analysis (CA) and discourse analysis. While I follow the conventions of CA quite strictly in the three CA articles, in the final article, discourse analysis serves to operationalize notions from media ideologies as analytical concepts. In the following section, I will first discuss CA as an analytical method. Finally, I discuss how the discourse analytic study complements the conversation analytic studies and how this contributes to study design as a whole.

4.2.1 Conversation analysis

CA is the systematic analysis of talk the way it is produced in everyday situations of human interaction, talk-in-interaction. With attention to the orderliness of social interaction, conversation analysis provides unique insights into the organization of talk. CA’s aim is to focus on the production and orientation to the production and interpretation of talk-in-interaction as an orderly accomplishment that is oriented to by the participants themselves (Hutchby & Wooffitt, 2008, pp. 12-13).

It is possible that the detailed study of small phenomenon may give an enormous understanding of the way people do things and the kinds of objects they use to construct and order their affairs. It may very well be that the things are very finely ordered; that there are collections of social objects […] that persons assemble to do their activities; that the way they assemble them is describable with respect to any of the activities they happen to do, and has to be seen by attempting to analyze particular objects. We would want to name those
objects and see how they work, as we know how verbs and adjectives and sentences work. Thereby we can come to see how an activity is assembled […]. (Sacks, 1984, pp. 24-25)

With the attention to detail, CA not only gives us insight into the unique details of interaction following a participant’s own perspective, but in doing so, it also demonstrates how participants in interaction carry out their activities. In order to assume the participants’ perspective in the analysis of talk, an emic perspective, CA focuses on the sequential order of interaction. Attention to the sequential order of interaction gives insights into how participants understand and respond to each other in their turns at talk (Hutchby & Wooffitt, 2008).

Conversation analysis provides a tool for the analyst, the next-turn-proof procedure:

But while understandings of other turns' talk are displayed to co-participants, they are available as well to professional analysts, who are thereby afforded a proof criterion (and a search procedure) for the analysis of what a turn's talk is occupied with. Since it is the parties' understandings of prior turns' talk that is relevant to their construction of next turns, it is their understandings that are wanted for analysis. The display of those understandings in the talk of subsequent turns affords both a resource for the analysis of prior turns and a proof procedure for professional analyses of prior turns—resources intrinsic to the data themselves. (Sacks et al., 1974, p. 729)

It is through the attention to the sequential order of talk, exploring what happens in the next turn, that conversation analysis aims to take the participant’s perspective, an emic perspective. The ‘next turn’ is the place where speakers display their understanding of prior turn’s possible completion. It demonstrates how participants in interaction actively analyze the ongoing production of talk to negotiate their own, situated participation in it, and furthermore, their understanding of what the prior turn was meant to do. In addition to the sequential order of talk, the inferential (the cultural and interpretive resources participants rely on in order to understand one another in appropriate ways) and the temporal order of talk (how talk unfolds in time) are crucial (Hutchby & Wooffitt, 2008, p. 42).

Orienting to the details of interaction, CA not only gives insights into the fascinating details of the organization of social interaction – through the orientation to what happens in the next turn, the next-turn-proof procedure – but it gives unique insights into the participants’ understandings of the ongoing actions. Not through an analyst’s external assumptions as to what the participants think or assume, but through exploring the participants’ actions and how these display their understandings of ongoing actions. Although the sequential order of consecutively interpreted
interaction is negotiated by participants in situ, and is different from what we recognize from other interaction, attention to the local organization of the interaction can teach us about the participants’ understandings of the ongoing actions and activities. Transcription is an important part of working with interactional data, and I will return to this following the next sections on discourse analysis and the combination of methods.

### 4.2.2 Discourse analysis

The fourth study in the thesis is a product of the PhD project as a process. The idea and conceptualization emerged in the intersection between the societal context and motivation for the study, the ethnographic processes and engagement with practitioners, and importantly, the work with conversation analytic studies. This data was created without having an analytical framework planned in advance and emerged as a part of the ethnographic processes. The conceptual work for the final article developed after having worked with the project, collected data and discovered emerging themes in discourses on video-mediated interpreting.

Shifting the focus of the analytical lens from the conversation analytic object of study to how people communicate about video-mediated interpreting, revealed that participants frequently engaged with the comparison of video-mediated interpreting to telephone interpreting and onsite interpreting. I did the same. This practice of comparison concurs with one of Gershon’s (2010a, 2010b, 2017) themes of media ideologies, remediation. Participants also engaged with topics related to materialities and affordances.

Government documents are clearly distinct from interview data in several ways. The interviews consist of narratives that have been elicited by my questions, and I am a co-participant in the interaction where these narratives occur (I demonstrate this in section 4.2.4 Transcribing for analysis). Furthermore, the interviews are constructed there and then as a result of our conversation. The data sources have different functions and areas of authority. In order to find an analytical approach that would work for both text analysis and the ethnographic data sources, the analysis draws on critical discourse studies (Wodak & Meyer, 2016), aspects of discourse-historical approach (Reisigl & Wodak, 2016) and critical discourse analysis (Fairclough, 2003, 2016). The goal has not been to identify discourses within the field, but to explore the occurrence of these specific themes in interview data and in text. Conversation analysis is operationalized through open, data driven questions and next-turn-proof procedure. While the preliminary analytical work in this study was open and data driven, the next round of analysis involved categorizing the data according to theoretical categories. After having identified remediation, affordances and materialities, I used
NVivo to code transcriptions of interviews and government documents. Later I made excel worksheets to organize and scrutinize propositions in the texts in different ways. Through simple research questions while working with the text, I identified different patterns in the texts and interviews.

A) Does the comparison between media occur in the text or interview? Is this comparison implicit or explicit?

B) Which channels of communication are being compared?

C) What does the comparison do in the text or interview?

I went through the sections of the documents where video-mediated interpreting is mentioned, exploring the sections’ functions within the text. Similarly, I explored modalities of statements made in the texts regarding video-mediated interpreting and clarity of intertextual connections to other texts. The farther a proposition in the text was from the original source, the more it relies on the common doxa and draws on readers’ assumptions. These different intertextual connections are relevant for understanding the construction of the arguments in the text. The analysis in the article outlines some of the arguments in the text and engages in in-depth analysis of some statements. The statements from the texts are juxtaposed to remediation in interviews.

4.2.3 Combining conversation analysis and discourse analysis

The initial idea for this project was to use ethnography to gain a deeper understanding of the mediated event, understanding the single event in context of what happened before and after. Using ethnography to inform for instance placement of cameras (Mondada, 2013) has been done earlier in conversation analytic studies. Also, ethnography has been used in workplace studies to gain a deeper understanding of complex working situations (Luff, Hindmarsh, & Heath, 2000, p. 13). Conversation analysis has also been used in ethnographic research (e.g. Moerman, 1988). However, in this study, I do not limit my engagement with the ethnographic data sources to enhancing understanding in the conversation analytic studies, I have added a fourth article which shifts the object of study from researching talk as interaction to exploring ideologies in talk and text. In this section, I motivate this choice and discuss the consequences of this.

While the PhD project is motivated by a curiosity regarding how participants in interaction make sense of their activities in these multilingual, mediated, hospital encounters, the study was reasoned for with grounds in societal relevance: the articulated political aim to increase the use of video-mediated interpreting in Norwegian society. The conversation analytic articles give insights into to the accomplishment of interpreting in video-mediated environments. Conversation analysis
gives important and unique insights into the organization of video-mediated interpreting in hospital encounters. However, discoveries during my work with conversation analysis has made it difficult even to use statements from the government documents as a context for the study. Conversation analytic studies demonstrate how proposals from the documents are problematic and even fallacious. Coming from the field myself, I have shared the background knowledge, the doxa, upon which the arguments in government documents rest. My prior work and studies have fed into the discourses on video-mediated interpreting in Norwegian society in different ways. Not necessarily as contributions to the knowledge base on video-mediated interpreting such as it is constructed in documents. In these documents, projects on video-mediated interpreting serve to confirm that screen interpreting is a legitimate way to provide services. The organization of interaction in hospital encounters where interpreting is provided through video technology is rich and interesting and challenges even some of conversation analysis’ basic theoretical assumptions. It is more than interesting enough to study in its own right. However, video-mediated interpreting in society is heavily politicized at this point and I feel obliged to let the perspectives meet.

Engaging critically with the construct of video-mediated interaction in discourses in government documents, demonstrates how the government documents reduce interactional matters that have been reported over the years through various sources and decontextualize these from the activities they are situated within. This is problematic as the implementation of video-technology for service provision is not just a matter of simplifying logistics; it is a matter of fundamentally changing participants’ foundation for establishing understanding in the interaction. I have chosen to create a final article addressing something I am not able to do within the framework and epistemological limits of conversation analysis. I let political arguments, practitioners’ narratives and findings from conversation analytic studies meet. The combination of methods allows me to create a study that refers back to societal issues that I used to argument for the study.

With different objects of study, conversation analysis and discourse analysis have different epistemological foundations. In this thesis, I have chosen to keep the analytical methods apart. While the three first articles employ conversation analysis, the fourth and final article builds on a different dataset and employs a different method in the analysis. The three articles employing conversation analysis were conceptualized before the final article. Skogmyr and Balaman (2018) distinguish between studies that (1) draw on exogenous theories and let these theories guide the research agenda already before data analysis, and (2) those that first adopt the data-driven emic approach that is central to CA and only in the post-analytic interpretations of the findings relate these to exogenous literature (p.8). I have conducted the conversation analytic studies according to the aims of unmotivated looking and asked analytical questions of how the participants organize
their interaction. I have only turned to the exogenous theories after the conversation analytic studies were conceptualized. While I realize that this is unconventional, I do believe that the final article provides a valuable contribution to the thesis as a whole. I return to the findings of the articles and the implications of the thesis as a whole in the discussion.

### 4.2.4 Transcribing for analysis

Transcribing is an analytical process. Through the process of transcribing, the researcher makes reductions which might result in some of the features of the recorded material to vanish, while the transcription at the same time will focus on some of the features in the interaction which will be foregrounded (ten Have, 2002, p. 24). In this way, the process of transcribing might be seen as “instrumental in gaining a sharper focus on the phenomena of interest” (ten Have, 2002, p. 24), while at the same time it might no longer be possible to reconstitute all the features of the interaction from the basis of the transcript (ten Have, 2002). In the process of transcribing, although the ideal might be to transcribe all details in the interaction, some aspects of the interaction have come to the foreground while others may no longer be available.

The video and audio recordings I have made document specific events. The video recordings document so-called naturally occurring talk. That is to say that while the video-recorded hospital meetings meet certain requirements in order to be included in the study, these recordings capture institutional events that would have taken place whether the events were recorded or not. The interviews, on the other hand, are audio recorded and are interactional events that take place for the purpose of the study. I will start by describing the work with the video recordings of video-mediated interaction, before I continue with the process of transcribing interview data.

**Transcribing for conversation analysis**

The data consists of multilingual interaction in Norwegian and seven other languages: Albanian, Arabic, Bosnian/Croatian/Serbian, Mandarin, Polish, Thai and Vietnamese. I am not versed in any of these languages excepting Norwegian. While building a study based on a multilingual dataset gives insights into how participants organize their interaction without delimiting the language systems involved, working with a complex dataset like this is an arduous task even before taking into consideration the financial and temporal limitations of a PhD project.

Conversation analysis promotes unmotivated looking. However, having some experience working with this type of data and having already developed an analytical inquiry in the data, I did
not begin from scratch. I began working with the Norwegian parts of the interaction, transcribing these first while identifying interactional phenomenon of interest. While working with the Norwegian parts of the data, I discovered interactional practices I found interesting and examined them closer. When identifying interesting phenomenon, for instance the temporary suspensions of turns (article 1) or embodied displays of trouble (article 2), I would transcribe the Norwegian talk first, leaving slots for talk in the other languages when necessary. Thereafter I worked with either multimodal organization of interaction or the translations, depending on which linguistic resources I had access to at that point in time. I discovered that I could identify turn construction units (TCU) in other languages and could therefore arrange the transcripts accordingly. When I needed to work on embodied resources in the interaction before I had the chance to get help with translations, I would work with the speech I did not understand, structuring the transcripts according to turn construction units and annotating gesture, gaze, torque and other embodied resources depending on the analytical focus of the specific collection I was working on. Although multimodal conversation analysis, in principle, treats linguistic and embodied resources the same way (Mondada, 2018, p. 86), it is necessary to have a linguistic line in the transcription in order to annotate embodied actions.

For instance, in one case, I had been working with a piece of data for a while, working with the sections in Norwegian. Since the study treats the organization of interaction as multimodal, and embodied actions are commonly transcribed relative to talk (Mondada, 2018), I would try to identify turn construction units in the talk that I did not understand in order to annotate embodied resources relevant to what I was exploring. After having worked with the piece of data for a while, I discovered that a word occurred frequently in the final position of what I assumed made a TCU. I sent a message to a friend who knew the language in question trying to figure out what this word I kept hearing might be (M. Wattne, personal communication, April 17 2019).
May I ask you one more thing? Is ni kha something?

Or mi kha?

The latter might be “yes” (as confirmation to “have you/is there” or similar questions). The prior might be “here”. Kha is still a polite particle/confirmation (kha can mean “yes”). I imagine that something is being shown or pointed to when “ni kha” is said, in which case it means “here”, “this” or similar. If it is “mi kha” there would probably be a question first, if something exists kind of thing, and then the answer is yes.

But could it be “mai kha”? Cause that is “no”!

The following section is from a transcript, demonstrating the interactional position of the particle I had observed.

INT: ee:: kawn uen loei na kha, uh:: before anything else prt
    ai nee kaw pen naathee pen larm na kha
    so then these are the interpreter’s duties prt

The multimodal transcripts in the articles follow the conventions of Mondada (2001) (see appendix A). However, transcribing is a selective activity depending on the objectives of the analysis, the granularity of the transcript and more (Mondada, 2018). The transcripts express different granularities depending on the analytical objectives and contexts for presentation (article, data session or conference presentation), and this is also the case in the articles in this thesis. For some of the analyses, it was necessary to work with the transcriptions from both recordings separately in order to analyze the unfolding of the interaction.

For transcribing and translating the data in other languages, I have collaborated with people who could assist in the transcription and translation of these languages. In addition to identifying the words that are spoken, I operated with certain questions to try to identify some of the
conversational features I would be interested in when working with extracts. Since writing up an analysis and revising an article based on feedback gives rise to new questions, the process would in many cases become iterative and reiterative. As an initial part of the process, I might ask if an utterance was syntactically complete or if, based on the prosody, the utterance sounded complete or if it sounded as if the speaker would continue. In some cases, my questions would be completely irrelevant for the language we were working with. Depending on what the extract we were working on was for, I might need extra information about pronouns. In some cases, for instance with Polish, I have had the possibility to get help from a skilled conversation analyst who know both Polish and Norwegian. When working with for instance Arabic, many choices needed to be made regarding how to transcribe this. Again, I have been very fortunate to draw upon the help of people who have experience with transcribing Arabic for a Nordic setting (for instance, with participants in the INTERPRETING project at the University of Copenhagen).

![Figure 5: Notes from transcribing session](image)

Figure 5 is an image of some notes after a session working with Arabic in Copenhagen. The image shows Arabic handwriting to the right, Danish handwriting in the center and Norwegian handwriting at the upper corners. After having worked with an analysis based on a transcript and its accompanying recording, I would realize that I had new questions. For instance, after having worked with the extract below for a while, I started to wonder which of the words in the sentence was “we” and whether the pronoun could include both male and female participants. Could the interpreter be included by the pronoun “we”? 42
The process has been iterative and it has been necessary to revisit both analyses and translations throughout the analytical process and throughout the work with the articles. Working with translation of the data, both from Norwegian and from the other languages to English, continues to demonstrate how translation is an analytical process. The translations present the data in English, but there are a number of considerations to take when translating. For instance, and quite importantly, while many may argue that glossing is required in the presentation of data, for the purposes of the analyses in this thesis, the transcripts are already very tightly packed with information. The extract below demonstrates the richness of the transcription in an extract demonstrating how the interpreter first produces a repair initiator in Norwegian, before producing one in Albanian:

4 *(0.2)¤(0.8)¤#(0.3)
  ale: *turns to interpreter-->
  int: ➡head forw.¤holds pos.----->

5 INT: unn[skyld]
  sor[ry ] (NOR)

6 ALE: [fi ]tness
  int: ➡-holds head position-->
    (0.4)¤(0.1)

7 int: ➡-turns head to the right, holds-->

8 INT: çfarë the?
  what did you say? (ALB)

9 ALE: palestër.
  gym (ALB)

10 (0.5)

While this has been an incredibly interesting process, it has had its limitations. One is that the main point of departure for all analysis is the Norwegian side of the interaction. Naturally, this has affected the choice of phenomenon and topics and how the analyses have been developed. For instance, working with languages I do not know, although receiving help transcribing and translating, I still do not feel I have a deep enough understanding of the inferential order of the talk to make strong analytical claims. In a larger project with a different budget, other topics could have
been explored. Limiting the involved languages to languages I know would not secure enough data for the project.

Transcribing for discourse analysis

The interview data are audio recordings of various lengths with varying numbers of participants. Compared to the video recordings undergoing conversation analysis, the audio recordings were transcribed to undergo analysis of content. I began by transcribing the interviews at the level of detail as if it was for conversation analysis. This provided valuable insights into how the researcher is not only a co-constructor of meaning through the choice of words and questions asked, but even at an interactional level through leaving silence after a participant’s response and as such allowing or even encouraging a participant to elaborate on a topic.

For these transcriptions, conversation analysis and the tools provided by conversation analysis for transcribing initially provided a framework. However, in these transcripts the level of detail is reduced, for instance leaving out demarcation of intonation, emphasis, speed of speech, pauses and more. After some consideration, I chose to transcribe these orthographically so I could conduct word searches in the documents. After having transcribed some of the interviews, I was granted assistance to continue transcribing. For reasons related to time management, we found that leaving out minimal responses would save time transcribing and this would suffice for the analysis that was going to be done.

Example A: Transcript for the interview study

NUR:  
*det er jo alltid sånn at (x) em*
  it is after all always so that (x) uhm
*når man sitter på den ene siden*
  when one sits on the one side
*så syns man kanske at de som snakker det andre språket*
  then one perhaps thinks that those who speak the other language
*bruker veldig lang tid på å-
  use very long time to-

RES:  
*ja*
  yes
*NUR:  
*forklare ting*
  -explain things
*mange ord*
  many words
*som ee tenker ee ja*
  which uh (I) think uh yes
(x)

---

I am grateful to MultiLing for granting me help transcribing, and to Marit Johanne Furskognes and Mari J. Wikhaug Andersen for helping me transcribe.
jeg kan si litt mer jeg bare-
I can say some more I just-

**Example B: The same sequence with CA transcripts**

NUR: >[det er< jo] allti:: (1.1) ts. sånn at (0.8) em
   it is after all always ts. So that uhm
(0.8)
   når man sitter på den ene siden,
   when one sits on the one side
   så syns man kanskje at e (0.4)
   then perhaps one thinks that uh (0.4)
de som snakker det andre språke:
   those who speak the other language
   bruker veldig lang tid
   use very long
(0.2)
>på [å ]< forklare ting=
   to explain things
RES: 
   [ja]
   yes
NUR: ==>mange ord<
   many words
(0.2)
   [som] e: (0.5) tenker (0.7) ee (0.6) ja.
   that uh (0.5) (I) think (0.7) uh (0.6) yes.
RES: 
   [jah]
   yes
(2.2)
>jeg kan si< litt mer jeg bare:m,
   I can say some more I just um

**Example C: The same sequence prepared for the article**

Det er jo alltid sånn at em når man sitter på den ene siden, så synes man kanskje at de som snakker det andre språket bruker veldig lang tid på å forklare ting.

It is after always like that uhm when one sits on the one side, then one perhaps thinks that those who speak the other language use very long time to explain things.

Finally, the choice of transcription is tightly connected to the object of analysis. Different transcripts afford the analysis of different topics. While the work with conversation analysis has influenced my work, I do acknowledge that too many details not relevant to a specific analysis can be distracting. The data such as it is prepared in the transcript has a minimal level of detail to it.
4.3 Ethical considerations

Making video recordings of hospital encounters is not a task to be taken on lightly. In addition to meeting formal requirements, many ethical aspects have been considered in the design and implementation of this project. In this section, I first describe the formal approvals that have been granted before I discuss some of the choices made and concerns throughout the project.

Seeing that the project is based on hospital interaction, the project has been considered by REK Regional Committees for Medical and Health Research Ethics (2015/648 and 2017/1341) and found not to be governed by the Health Research Act due to the nature of the project. The project is registered and approved by the Norwegian Centre for Research Data (project number 43660 and 55153). Seeing that the project is based on hospital interaction, the project is approved by the three hospitals involved and their respective data protection bodies. The project sought wards interested in collaboration, and in addition to gaining approval from the wards’ management, procedures were drawn up with each ward to secure a recruitment process that would minimize the intrusion to patients as much as possible. Similarly, procedures were drawn up with the interpreting unit in order to identify relevant assignments at the relevant wards. Finally, all participants involved in the recorded encounters have given their informed consent. In cases where someone did not want to participate, no recording was made.

In collaboration with each involved ward, I made specific procedures for contact. I would receive a list of scheduled bookings for video-mediated interpreting from the interpreting unit. The lists contained the name of the ward and the language requested in addition to time and date for the scheduled appointment. Scheduled appointments in languages where the interpreter unit had fewer than four affiliated interpreters were not included out of consideration for the interpreters. I had contact persons at the involved wards, and would ask them first if specific assignments might be relevant for the project. Based on this request, the ward would make the first consideration. If the assignment was considered to be too sensitive, or they considered the patient to be especially vulnerable or if the assignment was not relevant for other reasons, no further pursuit would be made. If the ward approved, two main procedures would be followed. If the ward had contact with the patient in advance, they would ask the patient during that communication. If the ward did not meet with the patient before the scheduled appointment, first contact with the patient about the project would be made at the ward in connection with the relevant appointment. In both cases, information about the project was given and informed consent was sought at the ward. Information about the project was distributed to the involved organizations in the time before recordings were going to be made, with the aim to reach medical professionals and interpreters with general information about
the research project. Before scheduled meetings that were considered by the ward to be suited for recordings, the contact person at the ward would contact the medical professionals involved in the specific meeting and I would contact the interpreter who was appointed the assignment. I gave interpreters information about the project and asked for their consent before the meeting. The medical professionals and patients were asked for consent at the ward. The order of events would vary depending on when the patient arrived and whether I had the chance to speak to the medical professionals in advance. A lot of work would go in to coordinating the recordings and preparing for the work, and in cases where patients did not show up, there would be nothing more to do than to pack up and return.

In order to inform them about the project, I would speak to patients and next-of-kin via an interpreter. This afforded me useful, first-hand experience of certain issues addressed in some of the articles. For instance, I found that I would have to structure the information in a different way than I would normally do, and shorten utterances in order to make them a manageable length for the interpreter (see article 1). Another observation was that patients and next-of-kin asked a lot of questions. I was very pleased about this, as these questions demonstrated to me that we had some kind of mutual understanding of the situation. A third observation was that this had to fit into what in some cases was a tight schedule for the medical professionals. Providing information about the project with interpreting and questions took up to 15 minutes in some cases. Being in the way like this in the ward gave me useful insights into everyday life there. I had dialogue with contact persons at the ward and we made adjustments to the procedures when necessary.

In one case, a patient opted out. I was actually happy that someone did, because it gave me the impression that I gave potential participants in the study an actual possibility to opt out. However, I was also a bit distressed by this. It was not because the patient had opted out. I felt unsure if I had given the patient enough information to know what the project was about before giving them the possibility to opt out. I realized very early on, while giving information on the project, that the patient was reluctant to having me there, and I adapted my presentation to what I interpreted as their discomfort with my presence. I am uncertain of whether I gave the patient enough information to make an informed decision (Speer & Stokoe, 2012).

So far, this section has mainly focused on the collection of video recordings. Interviews and field observations have had more of an institutional focus, and have therefore been of a different nature. While I originally intended to interview patients too, I decided rather early that this would be beyond the scope of the study. The interview forms have had an institutional focus on professional opinions rather than on personal experiences. This gives the data a very different
nature. The informants were recruited through different channels, some through direct contact, others via information passed on by their workplace at which point they could contact us. It was important to me not to pressure anyone into contributing to this part of the project.

The ethical considerations do not come to an end at the completion of data collection. Work with anonymization of the data and maintaining a respectful representation of participants has been and will continue to be a major concern. One aspect of this is patients and next-of-kins’ kind participation, making me privy to their personal information. Another aspect of this is portraying medical professionals and interpreters in a respectful manner. They may not have personal information on the line in the meetings or in interviews in the same way as patients do. However, although the research project has set out to be descriptive, the professionals may be evaluated and judged by readers based on their practices. It has been my intention to represent all participants in a respectful manner.
5 Summary
5.1 Article 1

Creating space for interpreting within extended turns at talk

PhD candidate Jessica P B Hansen and Professor Jan Svennevig
Journal of Pragmatics, accepted with revisions

In consecutively interpreted conversations, long multi-unit turns pose an interactional problem, as the interpreter may need to intervene into the turn space of the current speaker to interpret. This paper employs multimodal conversation analysis to explore multimodal practices used by medical professionals and interpreters to manage the temporary suspension of extended turns-in-progress. The study contributes to an understanding of interpreting as interactionally achieved through participants’ collaboration. The article demonstrates how the temporary suspension of medical professionals’ longer turns for the purpose of interpreting is achieved through joint effort by interpreters and medical professionals in collaboration. The temporary suspension of a turn can be occasioned by the medical professionals by designing their turns in shorter installments. The installments may vary in length and may span from the syntactically, grammatically and pragmatically incomplete to seemingly complete utterances. By designing a turn in installments, the medical professional contributes to creating temporary suspension points, temporarily halting the progress of the turn to allow interpreting. In addition to pausing their speech, medical professionals use a range of resources allowing and even inviting the interpreter to speak, such as gazing toward the interpreter, gesturing toward the interpreter, and explicitly addressing the interpreter. Interpreters can contribute to creating suspension points for interpreting by producing pre-beginning signals, such as audible in-breaths, using gestures and explicitly asking for the turn. The interpreter’s signals display and orientation to certain points in the medical professional’s longer contribution as relevant for interpreting. The medical professionals can pre-empt these points by continuing past them and not leaving silence. While the construction and design of installments may be similar in situations where the interpreter is co-present and in mediated settings, video-mediation poses specific challenges to the negotiation of installment lengths. Delay can cause problems for the participants in negotiating the timing of a TSP. Furthermore, the specific environment and participation framework may limit the resources the interpreter has available for signaling attempts to interpret.
5.2 Article 2

Recruiting repair: Interpreters’ displays of trouble in video-mediated environments

Jessica P B Hansen, manuscript ready for submission

This article employs multimodal conversation analysis to explore interpreters’ embodied displays of trouble in video-mediated environments. The embodied display of trouble serves as a versatile device that engenders repair addressing trouble hearing and understanding. The same bodily movement in a different sequential position can be used to identify trouble sources, for instance actions causing auditory disturbances, at the ward. Furthermore, the embodied display of trouble does not require the interpreter to identify or even choose a language in which to produce a verbal repair initiator or ask for help. Since embodied resources afford simultaneity without interrupting ongoing talk, embodied resources run less of a risk of resulting in overlapping talk. With the unclear turn boundaries of consecutively interpreted talk, embodied displays of trouble can be a useful resource to solve problems. Even though the embodied displays of trouble may go unnoticed by co-participants in the interaction, they are systematically produced and interactionally organized, and the following trajectories of the interaction suggest that they are not merely displays of puzzlement. The interactional settings in which embodied displays of troubles are produced are quite complex. While the technology affords the participants with visual and auditory access to each other, visual access depends on the camera positions participants choose and how they use the screen displaying each other in the organization of interaction. Securing the sightline of participants at the ward is not always possible for the interpreter from a remote location, and the embodied display of trouble does not attract attention from a participant gazing in another direction. Embodied actions have different affordances when being displayed on a screen than what they do when participants are present at the same location. Furthermore, transmission delay is a feature of the video-mediated environment that may change the temporal unfolding of actions at each site.
5.3 Article 3

Invisible participants in a visual ecology: Visual space as a resource for organizing video-mediated interpreting in hospital encounters

Jessica P B Hansen

Social Interaction. Video-based Studies of Human Sociality, in press

This article employs multimodal conversation analysis and demonstrates how visual ecologies in video-interpreted hospital encounters serve multiple purposes for accomplishing interpreting. The visual affordance may enable access to information relevant for the accomplishment of the interpreters’ work. The participants presuppose that the media affords efficient use of embodied actions. However, the participants do not always ensure that their views of each other and each other’s surroundings are congruent with the activities and actions they are attempting to accomplish. Due to video-mediation, participants’ utterances may be disconnected from the ecology in which the utterances are produced. Similarly, due to the multilingual nature of the interaction, the linguistic content may become disconnected from the embodied actions that encompass the linguistic content in the original utterance, such as gesture and gaze. This can cause complications for participants when making sense of participation frameworks and co-participants’ actions. Whereas participants’ lack of or incongruent visual access to each other may cause problems in the interaction, they do not attribute the interactional problems to the insufficient visual ecology or make adjustments to the setting. They simply solve the immediate interactional problem and proceed. What is a relevant visual ecology for the collaborative accomplishment of interpreting, and how this visual materiality does in fact inform interpreting and the interpreter’s work, might not be entirely clear to the participants in the interaction. The participants do not readily connect interactional troubles to insufficient visual access for the accomplishment of ongoing activities. This study has provided insights into the organisation of video-mediated interpreting in hospital encounters and how participants in these settings use and orient to a visual ecology in the organisation of interpreting. As such, the study contributes to the body of knowledge describing various professional activities in mediated environments, and specifically to the understanding of interpreted interaction within a mediated environment.
5.4 Article 4

Remediating the mediator: Media ideologies in policies and practices of medical interpreting

Jessica P B Hansen, manuscript ready for submission

Based on the analysis of government documents and interviews with medical professionals and interpreters, this article addresses media ideologies drawn upon by stakeholders in the debate about video-mediated interpreting in society and in medical encounters. Employing discourse analysis in the analysis of government documents and interviews with practitioners, the article explores how stakeholders compare screen interpreting to onsite interpreting and telephone interpreting. The article outlines a media ecology where video-mediated interpreting is compared to onsite interpreting and telephone interpreting. The article finds differences and similarities between remediation in government documents and in practitioners’ talk about their experiences. In government documents the comparison between channels of communication construct a political argument for increased use of video-mediated interpreting in Norway. The government documents shift between referring to the provision of interpreting, basing arguments on matters of access and economy, and referring to the interaction. The documents are intertextually connected to other documents, and reiterate and recontextualize propositions about video-mediated interpreting. To the practitioners, matters of interaction are more frequent in their narratives about video-mediated interpreting, where video-mediated interpreting is not just a matter of logistics, but a matter of accomplishing work. This article has demonstrated how media ideologies as a concept can contribute to an understanding of tensions between positions held by stakeholders regarding implementation of technologies to a media ecology. The concept of remediation has provided new insights and a different perspective on video-mediated interpreting in Norwegian society. Finally, this article demonstrates how ideologies drawn upon by government authorities do not necessarily correspond to the reality of practitioners’ work.
6 Discussion and conclusion

In this final chapter, I discuss how findings from the thesis contribute to the existing literature on video-mediated interpreting specifically and interaction more generally. I discuss theoretical implications of the study before discussing implications for practice. The conversation analytic articles in the thesis have approached multimodality in two ways: The first is as the spatio-material context that participants in interpreted hospital encounters have both created and used as an interactional space for the accomplishment of interpreting. The second is as the multiple resources that participants draw upon in interaction. The articles have explored the multimodal organization of interpreting within these complex interactional spaces. The thesis takes as a point of departure that interpreting is interactionally achieved, and has demonstrated how interpreting as an interactional activity is accomplished collaboratively by participants in interaction. The conversation analytic articles have focused on practices related to the accomplishment of interpreting, such as turn-taking and repair, within the specific spatio-material environment. Furthermore, the articles have shown how affordances of the video technology, such as the visual affordances, contribute to the interpreter’s basis for understanding (and misunderstanding), and hence to the accomplishment of interpreting. The interactional articles have demonstrated how the video-mediated environment can become a complex setting for the accomplishment of interpreting, and have shown how several interactional complications may and do occur. Touching upon topics of turn-taking and repair in addition to the visual affordance in the accomplishment of interpreting, the three conversation analytic articles give rise to a discussion on intersubjectivity and the reciprocity of perspectives in this interaction. I will return to this in 6.2 Theoretical implications.

The fourth article employs discourse analysis to a different dataset, and reveals how stakeholders refer to different aspects of video-mediated interpreting when engaging in processes of remediation; in this case the comparison of video-mediated interpreting to onsite interpreting and telephone interpreting. The article illustrates how the interactional organization of interpreting in the video-mediated environment is not just a matter of isolated interactional issues but impacts practitioners – both medical professionals and interpreters – in the accomplishment of their work. Problems with the coordination of turn-taking may not just be a matter of “wobbly turn-taking” such as is suggested by the hearing to the draft law on interpreting (Det Kongelige Kunnskapsdepartement, 2019, p. 58), but may be a matter of accomplishing medical encounters and hence providing effective medical treatment. The fourth article, while straying from the theoretical and methodological underpinnings of the rest of the thesis, brings the thesis as a whole closer to its societal context and sheds light on how matters of interaction are, in fact, matters relevant to practitioners in reflections about their work. The change of material setting is not just a matter of
providing services through different channels, but may alter the very nature of people’s work. The conversation analytic articles show how assumed affordances of the media, such as visual access to each other in interaction, are not simply there, but require consideration by participants in order to be achieved. Whereas this is observed in the conversation analytic data, analyzed in the interactional studies and present in the practitioners’ narratives, this is absent in government arguments that propose the increased use of video technology. The thesis has shown both the possibilities and the complications in the organization of interaction within the specific interactional space.

6.1 Contributions

The work presented in this thesis contributes to the understanding of interaction within the context of video-mediated interaction and of video-mediated interpreting. The first article contributes to studies of turn-organization, demonstrating how medical professionals’ longer turns may be produced in installments (Svennevig, 2018) and be temporarily suspended for the accomplishment of interpreting. The article shows how the suspension of medical professionals’ longer contributions can be initiated by either the interpreter or the medical professional, and that they may negotiate the length of installments through interactional practices. As such, the article demonstrates how the accomplishment of the interpreter’s turns, although central to the accomplishment of their work (see for instance, Englund Dimitrova, 1997; Frøili, 2001; Wadensjö, 1998), is also a result of other participants’ actions and, accordingly, collaboratively achieved by participants in situ. The second article explores interpreters’ embodied displays of trouble. This article contributes to the understanding of repair organization (e.g. Mortensen, 2016; Oloff, 2018; Schegloff et al., 1977; Seo & Koshik, 2010) by suggesting that other-initiation of repair might, similarly to requests, be produced with various degrees of transparency, and that interpreters’ embodied displays of trouble are formatted as recruitments (Drew & Couper-Kuhlen, 2014; Enfield, 2014; Kendrick & Drew, 2016). Through the use of multimodal conversation analysis, both articles draw attention to the use of embodied resources in the organization of interaction and to the organization of interaction within this specific setting. The third article takes as a point of departure the participants’ orientation to the visual affordance in the interactional organization of interpreting in the video-mediated environment. The article displays how the visual affordance informs interpreters’ linguistic choices in interpreting and actions. The article demonstrates how participants may use gesture to organize their actions, and how these may not be perceived by co-participants. As such, the article demonstrates how the participants assume that they do have visual access to each other and act accordingly although this is not always the case. This article contributes to understandings of interpreting as an activity in a video-mediated environment (e.g. Licoppe & Verdier, 2013; Licoppe
et al., 2018; Licoppe & Veyrier, 2017; Warnicke & Plejert, 2012) and to the understanding of professional conduct in different video-mediated environments (e.g. Hjulstad, 2016; Mondada, 2007; Pappas & Seale, 2009).

All the conversation analytic articles demonstrate how participants’ gestures may go unnoticed by other participants in the video-mediated environment. There are several reasons for the lack of perception of embodied actions in this environment. Video technology serves as a resource in the accomplishment of a common interactional space. While the participants have the possibility to create a visual space in these settings, they do not necessarily create an interactional space where they have congruent views of each other. While the general assumption, an assumption found in government arguments analyzed in article 4, is that video provides participants with visual access to each other, this is only the case if the participants create an interactional space where it is possible to see each other. Furthermore, in order to actually see each other’s gestures, participants will have to ensure congruent views of each other, and in order for the gesture to make sense, the participants may have to secure congruent views of each other’s surroundings as well. Even in cases where the participants have created an interactional space where they have the possibility to see each other, the perception of the interpreter’s embodied actions still depends on where the other participants are gazing. Furthermore, embodied actions are more noticeable when used by co-present participants than when displayed on a screen. Due to the unclear turn boundaries in interpreted interaction and that participants will often direct their utterances to each other and not to the interpreter, interpreters’ embodied actions may very well go unnoticed in the video-mediated environment. While the interaction proves to be asymmetric (Arminen et al., 2016; Heath & Luff, 1993) and this may have consequences for the way actions are accomplished, enabled, constrained or inhibited, the participants do not address this nor make adjustments to the visual configuration of the setting. While video technology may be a resource for the accomplishment of activities and is a resource that can be modified by participants in interaction, the third article demonstrates how participants in interaction may not make adjustments even when problems arise.

The fourth and final article in the thesis applies a different methodology to a different dataset. The article contrasts positions in government texts with the narratives of stakeholders. The analysis finds that matters of interaction are relevant to practitioners as a matter of accomplishing the work they are doing – accomplishing interpreting and accomplishing the medical appointment. The article demonstrates how media ideologies – here, the process of remediation – are used in the construction of knowledge within government documents, and how this knowledge is used to construct a political rationale for increased use of video technology for interpreting. Furthermore, the article demonstrates how the interaction itself is considered to be important to the practitioners
in the accomplishment of their work. Finally, the article, contributes to studies of media ideologies (Gershon, 2010b, 2017), by demonstrating how the concept of media ideologies, in this case remediation, can be used to explore institutional discourses. Practitioners’ narratives about the accomplishment of video-mediated interpreting are connected to the accomplishment of the medical appointment and interpreting. These narratives address topics of interaction and materialities of the media. Topics addressed by practitioners correspond with findings from the conversation analytic studies. While the situated practices are important in the practitioners’ narratives and tightly connected to the accomplishment of their work, in government documents these same issues are decontextualized and abstracted to matters of, for instance “different dynamics” or “wobbly turn taking” (Det Kongelige Kunnskapsdepartementet, 2019, pp. 57-58). The combination of approaches in this study, although unconventional, allows these perspectives to meet. The combination of studies demonstrates how government documents operate with idealized conceptions of the media’s affordance, for instance regarding participants’ mutual visual access to each other. Furthermore, matters of interaction are decontextualized in the documents and treated lightly. Conversation analytic studies show how visual access is not something participants simply have – it has to be created – and how matters of interaction are relevant for the establishment of understanding in interaction. Participants’ narratives demonstrate how matters of interaction are important to the practitioners in their work. The result of the combination of studies suggests that interactional studies can identify matters relevant for practitioners’ accomplishment of their work and that they can inform service design.

Prior studies of video-mediated interpreting that focus on interaction build largely on simulations and provide quantitative analyses. As such, this study has brought forth novel insights to the organization of video-mediated interpreting as it is accomplished by the participants in situ. Assuming an emic perspective to the interaction furthermore emphasizes issues such as they occur from the participants’ perspectives, highlighting what participants treat as problematic in interaction. The combination of the different approaches in this thesis shows how the ideologies that form the basis of government arguments for increased use of video technology are, to some degree, fallacious and problematic.
6.2 Theoretical implications

Addressing fundamental structures of interaction, such as turn-taking and repair, the study has shown how participants engage in actions relevant to the collaborative achievement of interpreting. Furthermore, it demonstrates how the affordances of the media may be of relevance for the accomplishment of the activities and for participants’ perception of actions – such as delay and the organization of the interpreter’s turns and embodied displays of trouble.

With the emic analytical approach, we find that other participants’ expectations of the interpreter’s actions, and as such their displayed understanding of the accomplishment of interpreting, is relevant to the actual accomplishment of interpreting. Building on the reciprocity perspectives, Schutz (1953) operates with the idealization of the reciprocity of motives. He suggests that even the simplest interaction in common life presupposes a series of commonsense constructs, for instance the constructs of others’ anticipated behavior. In commonsense thinking, people merely have a chance at understanding others’ actions sufficiently for the purpose at hand. In order to increase this chance, people have to search for the meaning a particular action has for the actor.

Article 3 demonstrated how the participants’ different access to the interaction at any given time may cause problems in securing progress in the interaction. The interpreter has only limited background information and limited visual access to the other participants and their surroundings at the ward. For the interpreter, this may cause the utterance produced by participants at the ward to become fractured from the environment in which it is produced. As such, the interpreter loses access to references and the participation framework encompassing the utterance. Similarly, for the doctor or the patient, the verbal content of each other’s utterances becomes fractured from the speaker’s embodied actions and the environment, as the verbal content becomes available only when the interpreter has interpreted the utterance. Since the meeting is carried out in two languages at the same time, the ongoing actions are not always available or even transparent to the all the participants. This may require that the participants figure out what is the current purpose at hand through other means. For instance, a doctor may have to ask verbally if the activity carried out in the other language has been completed. This manifests an extremely complex setting for the accomplishment of intersubjectivity.

Interpreting is an activity that is accomplished in and through the interaction. In order to accomplish actions that constitute interpreting, the participants in interaction need to have some kind of a common understanding of the activity in question. For instance, article 1 has shown that it is not only the interpreter who orients to the activity of interpreting at certain points in the interaction. In order for the interpreter to carry out their work and thus make participants’ turns
intelligible to other participants, they need to be able to take the turn. In order to for one participant to take the turn, another speaker might have to abandon a turn.

The articles have demonstrated how delay might cause trouble timing the interpreter’s turns. Similarly, in order for the interpreter’s embodied displays of trouble to recruit assistance from the other participants, interpreting has to be treated by the participants as relevant at the point in time when the embodied display of trouble is produced; the interpreter’s action has to be treated as relevant to the organization of interaction or other matters at hand. The participants organize a setting based on presuppositions regarding the situation. Participants presuppose that resources that would be efficient in a co-present setting will in fact be efficient in the video-mediated setting. For instance, the interpreter’s audible in-breath – which is a common pre-beginning signal indicating that the interpreter is ready to take the floor – may be delayed in transmission to the ward, causing it to become audible to other participants at a point less relevant for the temporary suspension of turns. The technology may even treat the audible in-breath as noise and cancel it altogether through noise cancellation. Similarly, gesture that may be used to organize the interpreting and the interaction as such may be rendered invisible to the participants at the other site due to their asymmetric camera access to each other. As the participants in interaction have various points of access to the ongoing interaction, they also have various insights as to how they are displayed to the other. They seem to presume that their view is reciprocated; they assume the reciprocity of perspectives.

While the technology has certain features that may challenge the effectiveness of resources participants use in interaction, these features should not be mistaken for problems. Delay is, in itself, not a problem. Delay becomes a problem when participants use resources to organize the interaction that are highly sensitive to timing. Similarly, lack of uptake of the interpreter’s embodied displays of trouble are not caused by the two dimensional image of the interpreter on the screen. However, the interpreter’s movement is less prominent on a screen than if the interpreter were present with the other participants. When participants use these resources in the organization of interaction, they do so based on the assumption that these resources are as available to the participants at the other site as they are to themselves. They operate with the reciprocity of perspectives as a basic assumption. The multilingual nature of interpreting makes it extra difficult to identify problems in the interaction.

While conversation analysis aims to investigate social interaction from an emic perspective (e.g. Hazel et al., 2014; Hutchby & Wooffitt, 1998; Robinson, 2013), the analytical issues encountered when working with video-mediated, multimodal and multilingual data raise topics for
discussion regarding the emic perspective. When studying video recordings of only the interpreter’s perspective, I had only access to the interpreter’s “version” of the situation. When adding recordings from the ward, I would be able to explore both versions of the event. While conversation analysis uses the next turn proof procedure as the methodological evidence of emic perspectives, there are still aspects of the interaction that cannot be taken for granted when working with only one participant’s perspective. However, making video recordings from several sites, gives rise to question of whose perspective should be analyzed (Rusk & Pörn, 2019). Similarly, when working with interaction where participants have different levels of access to the interaction, as is the case with multilingual interaction such as this, the question of whose emic perspective we assume in the analysis becomes relevant.

Finally, the combination of conversation analysis and ethnography the way it has been executed in this thesis has shown how topics of interaction, specifically turn-taking and repair, are not just technical details regarding the interpreter’s work, but relevant to how participants perceive the quality of the work they are carrying out.

6.3 Implications for practice

The study has shown how accomplishing interpreting in a video-mediated environment is an activity that not only relies on the skills and qualifications of the interpreter; it relies on the other participants’ actions as well. Different technological devices have different affordances. For instance, article 3 demonstrated how different videoconference systems have different camera angles. Utilizing the affordances in the media in order to create an interactional space appropriate for interpreting requires that the participants have knowledge of the possibilities in the media. Furthermore, it requires that the participants use the technology in appropriate ways. For instance, directing speech away from the microphone may cause problems of hearing for the interpreter. Aiming the camera at one of the participants, while several are left outside the camera angle may cause complications for the interpreter in making sense of the participation framework. The participants at the ward may need to make available information to the interpreter that is not readily available through the technology. Activities at the ward that do not disturb the participants at the ward, such as moving objects near the microphone, may cause problems for the interpreter. Participants’ knowledge and understanding of possibilities and constraints in the technology can promote the activity of interpreting. Using the technology to create an interactional space for interpreting, may prevent some possible problems during the interaction. Furthermore, the interactional organization of interpreting requires knowledge from participants other than the interpreter about interpreting as work and as interaction. The participants involved, especially the
professional participants, need knowledge about: the situation; the technology along with its possibilities and constraints; the organization of interpreted interaction; and about possible pitfalls in the accomplishment of medical encounters in the video-mediated environment. The initial assumption is that there might be a lot at stake, especially for the patient, and that the fractured nature of the interaction, considering languages, purpose of the meeting, access to interaction, makes this an incredibly complex setting. Actions intended to be correct, for instance according to guidelines or professional norms that the participants orient to in their practice, such as directing gaze to the patient and not to the interpreter, may be correct according to guidelines and may be thought to promote the relation between the patient and the medical professional. However, the same actions may prevent the participants at the ward from realizing that the interpreter is encountering a problem. Seeing that the interaction is quite complex and that each of the participants has only limited overview of the situation, it may also be difficult for each of the participants to gain the knowledge necessary to collaboratively accomplish the interaction in the specific situation in a way that promotes the interaction. The lack of explicit collaboration between the medical professional and the interpreter seems to prevent the professional participants from making necessary adjustments to the technology and thus setting up a proper interactional space during the beginning of the meeting and does not seem to be addressed by the participants even when this causes problems during the interaction.

The participants in interaction seem to orient to the interpreter as one who is supposed to be invisible. By this, I mean that beyond the opening phase of the meetings, where the interpreters say some words about their work, the participants rarely orient explicitly to the interpreting. Interpreting studies have for the most part moved away from the conduit model of interpreting. However, guidelines for interpreting (such as Helsedirektoratet, 2011) and the interpreters’ ethical guidelines (IMDi, 2020) do still assume an approach to interpreting building on the conduit metaphor, where information is transferred by the interpreter from one language to the other. Understanding interpreting as interaction, is not only relevant to studies of interpreting, it is relevant to people who need to collaborate with interpreters in their work, such as doctors, a point also made by Li (2015). For instance, medical professionals’ attentiveness to the interpreter and the interpreting throughout the interaction, such as glancing toward the screen or segmenting longer turns in installments, may promote the interpreting as an ongoing activity in the interaction. Teaching practitioners not only how to use technical devices for video-mediated interpreting, but how to configure the technical equipment and create an appropriate interactional space would help the practitioners avoid some problems in the interaction. Similarly, establishing practices regarding how to talk about these issues in the beginning of the conversations might benefit the interaction.
The analysis of government documents has shown how video technology is thought to be an efficient way to provide interpreting services. It also demonstrates how the media is attributed qualities that are supposed to replace the interpreter’s professional judgement. Screen interpreting is treated as a way to commoditize interpreting without taking into consideration the interactional and local accomplishment of the interpreting. However, as the interactional articles and the practitioners’ perspectives demonstrate, it is not that easy. For practitioners, it is a matter of accomplishing work-related activities in and through the interaction. For instance, the materialities of screen interpreting have been relevant in conversations with medical professionals, which draws attention to the importance of having the right technology for the specific setting in which they are engaging. What is important to medical practitioners depends on the unique setting and the work they are conducting. According to Suchman (1995) “[n]ot only do representations of work involve perspectives and interests, but work has a tendency to disappear at a distance, such that the further removed we are from the work of others, the more simplified, often stereotyped, our view of their work becomes” (p. 58). In the ideological representations of video-mediated interpreting, such as in government documents proposing its increased use, there is a risk that the work being carried out has been stereotyped. In the implementation of video technology for the provision of interpreting, in order to find and develop adequate technological solutions and proper procedures for professional collaboration, knowledge about the work being carried out, about activities that practitioners accomplish in their work, and an understanding of the specific settings is crucial.

This dissertation has investigated the accomplishment of interpreting in video-mediated environments. The participants’ actions in this setting are situated within a specific environment where the medical professionals and the patients in many of the meetings, although not all, have met prior to the recorded event and have a some common understanding of the purpose for the meeting. In these meetings, the participants may even to some degree, have common interests and goals regarding the outcome of the meeting. Although complications arose during the interaction, the participants were willing to look beyond this, solve the problems and carry out their activities. However, in matters where the institutional professional and the speaker of a minority language have differing interests, such as in cases where the minority language speaker is being tried in an asylum interview, a police interrogation or in a meeting with child protection services, it is crucial to take into consideration what consequences seemingly minor misunderstandings and trouble achieving intersubjectivity may have.
6.4 Limitations and future research

In this section, I will address some limitations of the study before I outline some suggestions for further research. First of all, this work has been of a relatively limited scope. Although the number of recorded meetings are limited, the data is rich and extensive. The data collected are in several languages that I do not know. Although this has given unique insights into phenomenon that would not easily have been observed working only with only one language, this has been a difficult dataset for a novice to work with. More data would naturally provide deeper and sounder insights. A bigger project might also be a better context for working with such complex data. I could have delimited the involved languages, however, that could have compromised the viability of the project because of the low use of video-mediated interpreting during data collection. Working with multiple sites has been rewarding and crucial to completion of the project. However, this has required a lot of resources for collaboration and coordination. Furthermore, involving several methods and a complex dataset has naturally caused extra labor. The simple way to avoid this would be to conduct the entire study as a conversation analytic study. However, the fourth study has broadened the context and given insights relevant for practice and policies. The study design is institutional, and in this sense, the voice of the patient is lacking. Due to the complexity and scope of the project, it was not relevant to include interviews with patients. However, this is certainly something that should be considered for future research.

This study has shown how conversation analysis gives useful and important insights into the organization of interpreting in a video-mediated environment. In order to learn how participants in interaction accomplish certain activities, studying how they accomplish just these activities within the specific setting is the best way to learn. Further studies should continue where this study ends, exploring the accomplishment of medical consultations or other institutional encounters where interpreting is carried out through video technology. Furthermore, in order to develop professional practices further experimental studies could be conducted based on analysis of interaction with different types of technical equipment. Experimental studies exploring results of training programs for this type of interaction could give useful insights. Training settings could be developed using CARM, Conversation Analytic Role-Play Method (Stokoe, 2014).
References


8 List of appendices

Appendix A: Transcription key for multimodal annotation

Appendix B: Interview guide for interviews with practitioners

Appendix C: Interview guide for other stakeholders
Appendix A Transcription key for multimodal annotation


* * Descriptions of embodied actions are delimited between
□ □ two identical symbols (one symbol per participant and per type of action)
% % that are synchronized with correspondent stretches of talk or time indications.
*---* The action described continues across subsequent lines
----* until the same symbol is reached.
>>> The action described begins before the excerpt’s beginning.
--->> The action described continues after the excerpt’s end.
int Participant doing the embodied action is identified in small caps in the margin.
fig The exact moment at which a screenshot has been taken
# is indicated with a sign (#) showing its position within the turn/a time measure.
Appendix B Interview guide for practitioners

Interview guide for interpreters:

Some introductory/warm up questions just to get started:

Examples: Have you done a lot of on-screen interpreting? How do you like on-screen interpreting? What is on-screen interpreting? The purpose of this talk is mostly to calibrate and to find a way in.

Main questions:

1. Kan du fortelle meg om et tolkeoppdrag der du gjorde skjermtolking som gikk dårlig? Please tell me about an on-screen interpreting assignment that went bad.
2. Kan du fortelle meg om et tolkeoppdrag der du gjorde skjermtolking som gikk bra? Please tell me about an on-screen interpreting assignment that went well.

I am interested in narratives and will follow up both questions with extra questions to get a full description. I am interested in what they feel is bad and what is good. What they feel causes a good situation and what they feel causes a bad situation.


I wonder what you think of this claim: On-screen interpreting makes the interpreting more professional. Other possible claims to discuss: The emotive gets lost, the communication is less complete, it is good to turn the screen off, it is good for patients not to have the interpreter present when they talk about something personal and other claims that occur in interviews.

I would like to present the person I am interviewing with one claim from previous interviews, fieldwork and documents. The point is to generate discussion and reflections on different aspects of ‘on-screen interpreting’.

Final discussion:

4. Hvis du skulle lage en veileder om hvordan gjøre skjermtolking, hvilke tre anbefalinger ville du gi helsepersonell som skulle kommunisere med pasienten uten at tolken var til stede?

If you were to make an instruction on how do on-screen interpreting, which three recommendations would you like to give medical professionals who were going to communicate with the patient without the interpreter present?
**Interview guide for medical professionals**

Some introductory/warm up questions just to get started:

Examples: Have you tried video interpreting? How do you like on-screen interpreting? What is on-screen interpreting? The purpose of this talk is mostly to calibrate and to find a way in.

**Main questions:**

1. Kan du fortelle meg om et møte som du hadde med pasient der tolken tolket via skjerm som gikk dårlig?
   
   Please tell me about a meeting you had with a patient once, where the interpreter interpreted through the screen (through video technology), that went bad.

2. Kan du fortelle meg om et møte som du hadde med pasient der tolken tolket via skjerm som gikk bra?
   
   Please tell me about a meeting you had with a patient once, where the interpreter interpreted through the screen (through video technology), that went well.

I am interested in narratives and will follow up both questions with extra questions to get a full description. I am interested in what they feel is bad and what is good, and what they feel causes a good situation and what they feel causes a bad situation.

3. Jeg lurer på hva du tenker om følgende påstand: Skjermtolking gjør tolkingen mer profesjonell. Andre påstander: Det emosjonelle blir borte, kommunikasjonen er mindre helhetlig, det er fint å kunne skru av skjermen, det er fint for pasienter å slippe å ha tolken der når de snakker om noe intimit og flere som kommer opp i ulike.

   I wonder what you think of this claim: On-screen interpreting makes the interpreting more professional. Other claims: The emotive gets lost, the communication is less complete, it is good to turn the screen off, it is good for patients not to have the interpreter there when they talk about something personal, and other claims that occur in interviews.

I would like to present the person I am interviewing with one claim from previous interviews, fieldwork and documents. The point is to generate discussion and reflections on different aspects of ‘on-screen interpreting’.

**Final discussion:**

4. Hvis du skulle lage en veileder om hvordan gjøre skjermtolking, hvilke tre anbefalinger ville du gi tolken om hvordan dette bør løses?

   If you were to make an instruction on how do on-screen interpreting, which three recommendations would you like to give the interpreter about how this should be done?
Appendix C Interview guide for other stakeholders

1. Hva er de viktigste utfordringene knyttet til tolking og tilgang til offentlige tjenester i dag? Hvordan kan disse utfordringene løses?
   *What are the most important challenges when it comes to interpreting and access to public services today? How can these challenges be resolved?*

2. I medieoppslag om tolking blir det ofte snakk om dårlig tolking. I den anledning har jeg noen spørsmål om god og dårlig tolking.
   *In news reports about interpreting it is often talked about poor interpreting. Therefore, I have a couple of questions about good and poor interpreting.*
   a. Hva er god tolking?
      *What is good interpreting?*
   b. Holder det å gjøre dette for å være en god tolk? (Hva er en god tolk?)
      *Is it enough to do this to be a good interpreter? (what is a good interpreter?)*
   c. Hva er dårlig tolking?
      *What is poor interpreting?*
   d. Er det sann at alle som gjør dette er en dårlig tolk? (Hva er en dårlig tolk?)
      *Is it so that anyone who does this is a poor interpreter? (What is a poor interpreter?)*

3. Hva tenker du er likhetene og forskjellene mellom skjermtolking og hhv telefontolking og fremmøtetolking?
   *What do you think are the similarities and differences between on-screen interpreting and telephone interpreting? How about on-site interpreting?*

4. Hvilke muligheter og begrensninger tenker du at video har sammenlignet med at 1) tolken selv er til stede og 2) telefon?
   *What possibilities and limitations do you think that video has compared to 1) the interpreter being present and 2) telephone?*

5. Hva tror du tolking via video vil kunne gjøre for tolking i offentlig sektor i dag?
   *What do you think interpreting via video can do for interpreting in the public sector today?*
   a. Hvilke av dagens utfordringer tror du kan løses med tolking via video?
      *Which of today’s challenges do you think can be solved with interpreting via video?*
   b. Hvordan bør offentlig sektor gå frem for å få til dette?
      *How should the public sector go ahead to manage this?*

6. Vi snakket litt om utfordringer knyttet til tilgangen til offentlige tjenester, skjermtolking nevnes ofte som en løsning for å sikre tilgang til kvalifiserte tolker:
   *We talked a little about challenges regarding access to public services. Video interpreting is often mentioned as a solution to secure access to qualified interpreters:*
   a. Hva tenker du om det?
      *What are your thoughts on this?*

7. Det finnes en del påstander om tolking via video som dukker opp i ulike tekster og sammenhenger. I tolkeNOU-en står at «Utvalget har med interesse merket seg at bruk av
skjermtolking i Danmark fører til at tolkesituasjonen oppleves mer profesjonell, på grunn av avstanden til tolken». Hva tenker du om dette?

There are many claims about interpreting via video, which emerge in different texts and settings. In the Norwegian official report “NOU 2014:8 Interpreting in the public sector” it says that “The committee has with interest noted that the use of on-screen interpreting in Denmark makes people experience the interpreting setting as more professional because of the distance to the interpreter.” What are your thoughts on this?

8. Dersom vi skulle lage en veileder for ansatte i offentlig sektor om skjermtolking, hva ville dine topp tre råd til dem som skulle bruke det, være?

If we were to make a brochure with recommendations for public servants about video interpreting, what would your top three advice for them on how to carry out a video interpreted consultation be?

9. Dersom vi skulle lage en veileder for tolker som skulle gjøre skjermtolking, hva ville dine topp tre råd til tolkene være?

If we were to make a brochure with recommendations for interpreters who were going to use on-screen interpreting, what would your top three advice for the interpreters on how to carry out a video interpreted consultation be?
ARTICLE 1: Creating space for interpreting within extended turns at talk

Authors: Jessica P B Hansen and Jan Svennevig

Journal of Pragmatics, accepted with revisions
Creating space for interpreting within extended turns at talk

Abstract

In consecutively interpreted conversations, long multi-unit turns pose an interactional problem, as the interpreter may need to intervene into the turn space of the current speaker to interpret. This paper explores multimodal practices employed by medical professionals and interpreters to manage the temporary suspension of extended turns-in-progress. We identify one practice used by medical professionals by which they produce turns in several chunks, or ‘installments’, temporarily suspending the ongoing turn and allowing the interpreter to begin translating. When speakers do not suspend their longer turns, the interpreter more actively signals, for instance through use of pre-beginning signals, the relevance of suspending the turn and leaving space for interpreting. We explore at what point during medical professionals’ ongoing multi-unit turns interpreting is made relevant, how this is done and by whom. Data from video-interpreted hospital encounters shows how video-mediation poses challenges to the fine-tuned coordination involved in creating temporary suspension points.

Key words: conversation analysis, turn-taking, interpreting, video-mediatied interaction, temporary suspension of turns

1 Introduction

Turn-taking models for interpreting suggest that interpreters produce their renditions after each of the participants’ individual turns (Gavioli, 2015). However, when it comes to longer contributions by one party, the interpreter might intervene during an ongoing turn at talk due to limitations in their capacity for memorizing (Wadensjö, 1998, p. 234). If speakers produce and complete longer contributions without allowing interpreting during the ongoing turn, this prevents them from monitoring the reception of the turn along the way and adjusting its course according to the addressee. Both the speaker and the interpreter may thus have an interest in splitting an extended contribution into installments and translating smaller chunks at a time.

This paper investigates multimodal practices used by medical professionals and interpreters for temporarily suspending multi-unit turns in progress and creating opportunity spaces for the interpreter to start interpreting. We propose to call this a temporary suspension point. We show
that medical professionals may design their multi-unit turns as a series of installments, that is, decompose the multi-unit turn into two or more component parts and present them one at a time, leaving room in between for interpreting. Interpreters, on their side, may invite such temporary suspension points by producing pre-beginning signals towards the end of a potentially complete turn constructional unit in the speaker’s multi-unit turn.

Furthermore, the article shows how the creation of temporary suspension points may be affected by the use of video technology. The data for the study is video recordings of hospital encounters in which the interpreter participates via video from a remote location. We show that the fine-tuned, split-second coordination involved in the establishment of temporary suspension points may be hampered by features of the video technology, such as delay and participants’ limited visual access to each other.

1.1 Multi-unit turns in conversation

Multi-unit turns suspend the general turn-taking rule that the completion of a turn-constructional unit (TCU) establishes the opportunity for speaker shift, a transition relevance place (TRP) (Selting, 2000, Clayman, 2013). Participants in interaction may use special practices to introduce the multi-unit turn, for instance presequences such as story prefaces (Sacks, 1992) and ‘pre-pre’s’ in complex questioning turns (Schegloff, 1980). During the production of a multi-unit turn, special practices may be used to signal the continuation of the turn beyond the TCU in progress, for instance leaning forward at the initiation of a multi-unit turn and maintaining the pose across the TCU boundaries (Li, 2013a). In institutional activity types, the turn-taking system may be specialized to license multi-unit turns in the encounter as a whole or in certain phases of it. Questions in news interviews, for instance, generally invite answers in the form of extended accounts rather than just a single TCU (Clayman & Heritage, 2002). The multi-unit turn is an interational achievement, which implies co-construction, recipient design and moment-to-moment recalibration and reorganization (Schegloff, 1996).

A specific practice used for inviting recipient contributions during a multi-unit turn is ‘speaking in installments’ (Svennevig, 2018). This implies that speakers divide their multi-unit turns up into smaller chunks in order to elicit response from the interlocutor along the way. The speaker leaves a pause after each installment and monitors the recipient by means of gaze, thereby creating an opportunity space for the recipient to produce a response, verbal or embodied (such as nodding). The response provided is most commonly an acknowledgement token or some other claim of
understanding, but the slot provided also provides the recipient with the opportunity to initiate repair at an early stage (Robinson 2014). The speaker projects continuation of the turn by producing each installment with slightly rising intonation, or gesturally by holding or continuing a gesture beyond the completion of each installment. Also, pragmatically, the action performed may display incompleteness in that it does not fulfill the sequential projection of a prior turn. In this way, the multi-unit turn is designed for allowing interpolated contributions by the interlocutor while simultaneously securing the right to continue the turn. In the current article, we argue that a similar practice is used in order to create temporary suspension points and allow for interpretation in the course of a multi-unit turn.

1.2 Interpreting and turn-taking

Interpreters’ turns in interaction have been found to respond to other participants’ turns either by providing renditions of other participants’ turns or by producing other actions, such as requesting clarification (Gavioli & Baraldi, 2011, p. 211). Interpreter’s utterances bridge a linguistic gap (between two languages) and a social or interactional gap (between two or more language users) (Wadensjö, 1998, p. 109). Just by taking the turn every now and then, Wadensjö suggests, the interpreter coordinates the other participants’ utterances. Accordingly, the interpreter is generally described as central to the process of turn-taking in interpreter-mediated encounters (Englund Dimitrova, 1997), and turn-taking is considered to be a basic part of the interpreter’s coordinating task (Frøili, 2001). It has also been noted that other parties may contribute to this process of coordination, such as when other participants encourage the interpreter to take action (Baraldi and Gavioli, 2012). During participants’ longer turns of talk, the interpreter may need to intervene into the turn space of the current speaker. Studies of forensic interpreting have found that in the process of giving their testimonies, minority speakers were deprived of the possibility to complete their account by other speakers self-selecting after the interpreters’ rendition (Licoppe, Verdier, & Veyrier, 2018; Licoppe & Veyrier, 2020). Interpreters’ intervention into another speaker’s multi-unit turn may have consequences for the speakers’ possibilities to resume the floor and to complete their narrative. The resources the interpreter has available in the management of these turns are influenced by the distribution of participants in video-mediated environments (Licoppe et al., 2018; Licoppe & Veyrier, 2020).

While the interactional organization of turn-taking in interpreter-mediated interaction has been described and explored in various ways (e.g. Davitti, 2019; Gavioli & Baraldi, 2011;
Wadensjö, 1998, 1999), the collaborative achievement of the interpreter’s intervention into other participants’ turn spaces has to our knowledge not been explored. Although studies of interpreting recognize interpreting as interaction, few studies address the collaborative achievement of interpreting (Li, 2013b). In this study, we explore how other participants may be involved in the coordination of the interpreter’s utterances down to the most basic interactional organization, the organization of turn-taking. We explore how medical professionals and interpreters negotiate the length of installments and the temporary suspension of medical professionals’ multi-unit turns. How do speakers construct their talk in ways that facilitate interpreting at certain points while still projecting more talk to come, how do they create space for interpreting?

1.3 Video-mediated interaction

While video-technologies provide participants-in-interaction with visual and auditory access to each other, the video-mediated environment has proven to be both enabling and constraining to the accomplishment of interactional practices (Arminen et al., 2016). Some features of the video-mediated environment become especially relevant in the organisation of turn-taking. “Time lag”, a delay in the transmission of signals from the one site to the other has been found to disrupt the turn-taking system (Ruhleder and Jordan, 2001). The asymmetric nature of the visual interactional space due to participants’ lack of full bodied co-presence (Arminen et al., 2016; Luff et al., 2003) and the ways participants organize their virtual and physical space in video-interpreted hospital encounters (Author 1, 2020, in press) may limit participants’ access to embodied resources in the organization of for instance turn-taking. This article analyzes the organization of turn-taking within medical professionals’ extended accounts, and thus the permeability of multi-unit turns. We investigate how the collaborative achievement of a temporary suspension point may be affected by affordances of the video-mediated environment, such as delay, asymmetric and limited visual access, and reduced mutual audibility, especially of simultaneous talk (and other vocal sounds).
2 Data and method

This study builds on video-recordings of video-mediated interpreting in hospital interaction. The data consist of recordings of 11 meetings with the interpreter participating as a remote participant and three with the interpreter onsite. The recordings are from visits in outpatient clinics and meetings with admitted patients. Although the meetings vary in length, number of participants, aims and topics, as well as phases and procedures, in all the meetings the medical professionals produce longer turns, and the temporary suspension of turns is common in both on-site and video-interpreted meetings. However, as video-technology provides a complex material setting and interactional space for the organization of turn-taking, the analysis builds on extracts from video-mediated interpreting.

In the video-mediated meetings, the interpreters are situated at a different location than the other participants and carry out interpreting via video-technology. The interpreting studios are equipped with videoconference units about the size of a personal computer. While some of the wards in the study have videoconference units similar to the ones in the interpreter’s studio, some have full videoconference systems. Three of the video recordings are made from the interpreter’s studio, one is made from the ward and seven are made from both the ward and the interpreter’s studio.

The article tends to the sequential organization of interaction and is carried out within the theoretical framework of multimodal conversation analysis (e.g. Mondada, 2014; Deppermann, 2013; Hazel et al., 2014). The transcripts are based on Mondada’s (2001) convention for multimodal transcript. The first part of the analysis builds on transcripts based on video-recordings from the interpreter’s studio while the final section juxtaposes transcripts from the interpreter’s studio and the ward in order to explore the differences in visual and auditory access to the various participants.

3 Analysis

In the following, we analyse extracts where interpreting occurs within a multi-unit turn by the medical professional. We start by speaker-initiated TSPs, i.e. cases where the medical professional actively accommodates interpreting by suspending the turn in progress and inviting the interpreter to take the next turn. We go on to interpreter-initiated TSPs, i.e. cases where the interpreter seeks...
to create a suspension point by producing pre-beginning signals at potential points of conditional entry into the multi-unit turn in progress. Finally, we analyse extracts exhibiting how video-mediation may cause challenges for the negotiation of a slot for interpreting.

3.1 Speaker-initiated TSPs

The following extract (1) shows a sequence where the medical professional constructs a longer turn in installments, thus temporarily suspending the turn for interpreting. The extract is from a meeting with several participants present. There are four medical professionals present at the ward in addition to the patient (PAT) and the next-of-kin, while the interpreter (INT) participates via video-technology. The extract is from the middle of an extended turn in which an occupational therapist (OT) presents future goals for the patient.

*Figure 1.1: From the interpreter’s point of view, line 3 in transcript*
Throughout the extract, the occupational therapist orients her gaze and bodily orientation toward the patient and glances toward the screen representing the interpreter every now and then during the interpreter’s speech. After a self-repair, the occupational therapist’s first utterance in this extract (line 1-2) takes the form of a complete syntactic unit, a main clause. Prosodically, it is produced with continuing intonation, indicating that the turn has not reached completion. However, the speaker does not continue speaking. The interpreter opens her mouth, and after a pause of 0.4 seconds, the interpreter starts interpreting (line 3). Thus, both the occupational therapist and the interpreter orient to this point in the interaction as relevant to interpreting.

After the interpreter has completed the rendition, the occupational therapist resumes her turn (line 8). This TCU is even shorter than the previous one and constitutes a syntactic increment to it,
a prepositional phrase specifying the verb phrase. The utterance again ends in rising intonation, indicating that the turn has not yet come to completion. However, rather than continuing to speak, the occupational therapist suspends her turn and the interpreter takes the floor after a 0.5 second pause.

After the interpreter has interpreted, the occupational therapist resumes her turn and produces yet another turn component (line 12). This is also presented as a continuation of the sentence introduced in line 1-2 and continued in line 8, as it constitutes a new syntactically dependent complement to the main verb ‘learn’. The utterance ends with rising intonation, projecting continuation, and a pause, once again making it possible for the interpreter to produce a turn.

The three TCUs produced by the medical professional all end at a point of potential syntactic completion, but the final rising intonation projects continuation of the turn. Pragmatically, the activity of presenting plans for the therapy also projects continuation. Thus, after each TCU it is clear that the speaker is not finished with her turn, yet she pauses and leaves a silence. By decomposing her multi-unit turn into smaller chunks and presenting them in installments, the medical professional actively accommodates interpreting. The interpreter also orients to the pause as an opportunity space for her to intervene and start interpreting. Although the medical professional does not turn her gaze away from the patient, she does not treat these turns as interruptive or misplaced. On the contrary, she seems to adapt to this rhythm of alternating by continuing to produce short installments and leaving pauses after each of them. The result is an alternation of speakership with rather short gaps and no overlaps, leading to a smooth coordination of speaking turns. The parties thus in collaboration create and recognize these places in the course of the multi-unit turn as temporary suspension points.

The installments found in the data are seldom as short as in the extract above. Usually, the medical professional speaks for a longer time before letting the interpreter intervene. How long each installment should be is a matter of online negotiation between the parties. In extract (2), only the doctor (DR) and the patient (PAT) are present at the ward. The doctor produces several TCUs realizing different pragmatic actions before turning to the screen and thereby signalling the relevance of letting the interpreter (INT) take the floor.
Extract (2)

1 DR: *og vi& ser samme som vi så under graviditeten din,
   and we see the same as we saw during your pregnancy
dr:  *gaze to patient-->
2   at du har veldig god effekt,
    that you have very good effect
3   og v[irus]*mengden åfall* (.h <dramatisk>),
    and the virus count is falling (.h dramatically
4 INT: [.hh ]
dr:  *gestures down  *
pat: &nods &
fig:  #fig2.1
5 DR: du hadde jo: (0.3) mange hundre (. ) millioner (0.5) e:::
     you had (particle) many hundred (. ) millions (0.5) e:::
6   >me virus<kopier.
    of virus copies
7 (0.5)
8 e:: og så faller det til under to tusen.
   and then it falls to under two thousand
9   >og det er< det vi ønsker,
    and that is what we want
10  under to tusen da som er .h som er e: bra.
    under two thousand then which is .h which is e: good
11  >så vi er< veldig fornøyd,
    so we are very pleased
12  & (0.3)*((0.3)&(0.7)
dr:  --> *turns head to int/screen-->
pat: &nods &turns head to int/screen-->
13 INT: ((begins to interpret))
Both in line 2 and 3, syntax and pragmatic aspects converge to mark the contribution as potentially complete, while the intonation is rising, thus projecting continuation. In the previous extract (1), we saw that such points were used to create a temporary suspension point. In this case, however, the DR does not suspend the turn, but continues by initiating new syntactic constructions and new pragmatic actions, in line 3 adding new information and in 5-7 specifying the decrease in the virus count. In line 7, the specification is potentially complete, and both syntax and intonation converge in marking the TCU as potentially complete. However, the MP expands the turn by adding a conjoined sentence with rush-through prosody (increased tempo in line 9). This new component links on syntactically to the previous TCU, but pragmatically, it introduces a new type of action, a report about the standards used for evaluating the results of virus counts. After this, he extends the turn even further by adding two new actions, an object-side assessment in line 10 and a subject-side assessment in line 11 (Edwards & Potter 2017). Only at this point does he display an orientation to the relevance of interpreting by both leaving a silence and by turning his head and gaze from the patient to the screen (line 12), thus selecting the interpreter as the next speaker.

This extract shows that points of potential completeness are not necessarily oriented to (by either parties) as temporary suspension points. As a result, speakers may construct rather complex, multi-unit and multi-action components before they leave space for the interpreter to intervene. Although the interpreter does produce a pre-beginning signal (line 4), an audible in-breath, this audible in-breath is not audible at the ward, and when the medical professional continues speaking, the interpreter does not attempt to claim the floor again. In this way, the size of the unit to be produced before the interpreter intervenes is a result of co-construction and collaboration, to which both parties actively contribute by either turn holding practices or turn initiating practices.

3.2 Interpreter-initiated TSPs

While extracts (1) and (2) show how medical professionals design turns in installments for interpreting, the following section shows how the interpreter can initiate the temporary suspension of a turn by producing pre-beginning signals. Extract (3) is from a meeting with the medical professional (MP), patient (PAT) and next-of-kin present at the ward in addition to the researcher. After the researcher (RES) has informed about the research project and asked for consent, the medical professional jokingly states that the participants from now on should pretend that the researcher is not present.
Extract (3)

1 MP: nå skal vi—
   now we shall
   mp: --> gaze to PAT and NOK--->>
2 da skal vi prøve å oversee:: ((gestures to researcher))
   then we shall try to ignore
3 RES: ja hehehe
   yes hehehe
4 MP: henne, for resten av samtalen,#
   her    for the rest of the conversation
   fig:   #fig3.1
5 INT: .hh () wobec tego próbujmy zapomnieć
   .hh () then we    try to forget
6   e:: o niej [tutaj siedzącej]
   e:: about her sitting here
   [((laughter)) ]
7 INT: ((continues to interpret))

Figure 3.1: Interpreter’s point of view, line 4

The medical professional is seated facing the patient and next-of-kin. She produces a syntactically complete TCU ending with rising intonation. Pragmatically, she marks the transition from one activity (related to the recording) to another (the meeting proper), thus also projecting more to come. The medical professional continues to gaze toward the patient and next of kin at the TCU boundary (see Figure 3.1). At that point, without waiting to see whether or not the speaker will pause, the interpreter produces an audible in-breath, signalling incipient speakership. The medical professional continues to gaze toward the patient and next-of-kin but withholds further talk during the in-breath
and the following micro-pause, thus providing space for the interpreter to start interpreting. In this case, then, it is the interpreter and not the speaker who signals the relevance of suspending the turn-in-progress. However, the medical professional orients to the interpreter’s pre-beginning signal as a legitimate bid for the floor by withholding further talk. Consequently, they collaboratively create a temporary suspension point by split-second coordination of turn-entry devices and withholding talk, leading to a smooth speaker transition.

In some cases, however, the interpreters’ attempts to create a temporary suspension point are not equally successful. The medical professionals do not always react to pre-beginning signals produced by the interpreter, and the result is a mismatch between the parties in calibrating the size of the turn-in-progress and thus the temporary suspension point. In the next extract, we can observe how the interpreter is not given space to intervene despite several attempts to initiate a turn. The extract is from a meeting with several participants present. The medical professional is presenting plans for the future.

**Extract (4)**

```plaintext
1   MP:   hh+ ha:n e: lager e: (0.2) hjelpemidler e: for e:-
     hh he e: makes e: (0.2) technical aids e for e:-
   mp:   ++turns to center-->
2   for hendene dine,
     for your hands
3   (.)
4   INT:  .h ((opens mouth))
5   (.)
6   MP:   >så du< kan e:: m (.e) bruke en kjøkken::kni::v,
     so you can e:: m (.e) use a kitchen knife
    (0.3)
8   INT:  .h ((opens mouth))
9   (.)
10  MP:    en elektr::isk tannbørste, og en barmaskin.
     an electric tooth brush and a shaving machine
11   (0.3)
12  INT:  .h (.# [on-]
13  MP:    [da ] få:r du utstyret sammen
     then you get the equipment together
14   int:   --> ﹀gaze up, nods﹂gaze down again-->
15   fig:   ﹀fig4.1
16  MP:    med et sånt (0.6) spesialtilpasset grep.
     with such a (0.6) specially adapted grip
17  INT:  .h (.#) +dakle on ﹀radi na izradi pomocnih sredstva
     .h (.#) so he is working on making aids
18  int:   --> ﹀gaze up-->
19  mp:    -->+turns to screen/cam-->
20  INT:  ((continues interpreting))
```
Both in lines 4 and 8, the interpreter treats the slight pause in the MP’s talk as a potential slot for interpreting. She produces a audible in-breath and opens her mouth. However, the MP does not react to these pre-beginning signals and continues her turn-in-progress before the interpreter has produced any vocal sounds. At the end of line 10, there is a convergence of syntactic, prosodic, and pragmatic aspects of turn completion. Again, the interpreter treats this as an opportunity to start interpreting, and this time she starts producing the first vocal sounds of an utterance (line 12). However, the MP overlaps with an expansion of the turn in the form of a new syntactic, prosodic, and pragmatic unit (lines 13-14). Only when this is brought to completion does she let the interpreter take the floor to translate. Thus, the interpreter orients to every TCU ending as a temporary suspension point, while the MP repeatedly blocks her entry into the turn space by continuing to speak. In this way, the calibration of a unit relevant for interpreting is the result of an ongoing process of negotiation, in which turn holding and turn yielding practices are used to propose, accept and resist the establishment of a temporary suspension point. This recording is made from the interpreter’s studio, and from this perspective, it is not possible to tell whether her pre-beginning signals are in fact perceivable from the ward. This raises the question to which degree the video technology influences the coordination of temporary suspension points. This is the topic of the following section.
3.3 Negotiating TSPs in the video-mediated interactional space

This section addresses trouble calculating the length of an installment in the video-mediated environment. Here we will explore how delay in the transmission of the video signal has consequences for the split-second coordination of turn-taking. Furthermore, we address the challenges posed by limited audibility and visibility. In order to do this, we will need to compare the recordings made at the ward with the recordings made at the interpreter’s remote studio.

We start by considering the effects of delay. Extract (5) shows how delay may affect the interpreter’s possibilities to create a temporary suspension point by producing a pre-beginning signal. In the following extract, the interpreter’s pre-beginning signal is transcribed twice (in line 5 and 7) marked with A and B. A is when, from the interpreter’s perspective, the interpreter produces the audible in-breath. B demonstrates at what point in the interaction this audible in-breath becomes audible at the ward. In this consultation, a medical professional and a patient are seated in front of a videoconference unit. The medical professional is in the course of summing up the progress in the treatment and the prognosis of the illness.
Figure 5.1: Medical professional continues speaking
The medical professional produces a syntactically complete sentential TCU with falling intonation, both features indicating potential completion. Pragmatically, however, the report of a potential relapse projects further talk about how such a problem would be addressed. In addition, and more concretely, the medical professional lifts her hands in the following micro-pause (line 5), preparing them for an upcoming gesture (Kendon 2004), thus projecting immediate continuation. At this point, the interpreter produces an audible in-breath (line 6), a pre-beginning signal proposing a temporary suspension point. This in-breath is transcribed as it occurs in the recording made in the interpreter’s studio. However, due to delay caused by the video-technology, the in-breath is not yet perceivable at the ward. Line 7 renders the same in-breath from the recording at the ward. As can be seen, it only becomes perceivable to the medical professional as she produces the first vocal sounds of a new TCU. At that very moment, she immediately cuts herself off and turns her head to the camera, thereby allocating the next turn to the interpreter. She also moves her right hand to her chin, and thus displays alignment as a recipient through embodied action. In this case, then, the delay in the transmission affects the precision-timing of the pre-beginning signal and leads to a hitch in the coordination of the temporary suspension point.

Extract (6) is from a similar setting and involves even greater delay, leading the doctor to abort a TCU well underway. Similar to the previous extract, line A refers to the point in the interaction the interpreter produces the audible in-breath while line B demonstrates at what point this audible in-breath becomes audible at the ward. The extract is from a sequence where the doctor is summing up the progress of the treatment so far. Just prior to this, the interpreter has interpreted the medical professional’s previous utterance.
Figure 6.1: Image from the interpreter’s studio to the left. Only the patient is visible on the interpreter’s screen. Image from the ward to the right. Doctor holds counting gesture which is not visible to the interpreter.
Figure 6.2: Doctor is speaking

Figure 6.3: Doctor turns to screen and gestures to screen
The doctor produces a TCU with rising intonation (line 1). As she begins telling the patient what the results of the treatment is, she clutches her fist with an extended thumb, in a counting gesture, indicating that this might be the first part of a list of things. However, the doctor is seated outside the camera frame and her gestures are not visible to the interpreter. The interpreter produces an audible in-breath shortly after the completion of the doctor’s first TCU (line 3). Here again, due to delay, the interpreter’s pre-beginning signal is not perceivable to the participants at the ward until the doctor has produced the first words of a new TCU (line 4). Shortly thereafter, the medical professional cuts off the utterance in progress and explicitly allocates the turn to the interpreter, both verbally and by turning her head toward the screen and gesturing towards her (line 7). Neither the gesture nor the gaze is visually available to the interpreter, but the verbal turn allocation secures that the interpreter takes the next turn and starts interpreting. The delay here thus again leads to problems of coordination and progressivity, leading to an aborted TCU and an explicit metacommunicative turn allocation.

Problems of visibility have already been observed in the last example, but here we will see how that and audibility problems may affect the coordination of a temporary suspension point. Extract (7) and (8) are from the beginning of a meeting with six participants at the ward, one patient and five medical professionals. Example (7) demonstrates how the participants have trouble coordinating the interpreter’s contributions due to limited mutual audibility. We start by looking at the recording made at the ward (7A) and then analyse the same extract as it appears to the interpreter in the remote studio (7B).

Extract (7A): From the ward’s perspective:

```
1   DR:   hei benjamin*=',
         hi benjamin
2        jeg skal oppsummere det som e <skjedde>,
         i   will summarize     that which e happened
     fig:       #fig7A.1
     pat:   >> gaze to dr -- turn to screen/int-->
     #fig7A.1
3         #fig7A.1
     pat:   #turns back to dr -->
     #fig7A.1
4   DR:   .hh (. ) e*:: du hadde da en e forgiftning med koka\i in
         .hh (. ) e::  you had prt an e intoxication with cocaine
     int:   *sits back -->
```
Figure 7A.1: Doctor orients to patient

The doctor greets the patient, addresses him by name and is oriented toward him through gaze and bodily orientation throughout the extract. The preface in line 2 is a potentially complete TCU in that it constitutes a complete syntactic clause and an independent pragmatic action. It is produced with rising intonation, but the doctor leaves a 0.5 second silence after it (line 3). Consequently, this moment in the interaction appears like a moment where the interpreter could have self-selected but did not. However, this transcription is based on the recording at the ward and captures only how the situation appeared to the participants present there. If we see the same extract from the recording in the interpreter’s remote studio, we get a different picture:

Extract (7B): From the interpreter’s perspective:

2  DR: jeg skal oppsummer e<skjedde>,*
    i will summarize that which e happened
int: -------> *sits back
pat: #gaze to screen-->

3  (0.5)n
pat: -->n gaze to part at ward, away from SCR-->

4  DR: .hh

5  (.)

6  INT: [emm]

7  DR: # du hadde da en e forgiftning med* koka\in
     e:: you had prt an e intoxication with cocaine
int: ---> *gaze notes-->
fig: #fig7B.1
From this perspective it is clear that the interpreter treats the silence following the doctor’s utterance as a potential opportunity to start interpreting. She begins to produce a vocal sound, a hesitation marker functioning as a pre-beginning signal (line 6). The interpreter’s attempt to take the floor is not audible to participants at the ward, as the doctor starts producing a prolonged hesitation marker (line 7), signalling continuation of the turn and blocking access to the floor by other participants. The interpreter orients to this by aborting her turn beginning and thereby loses the opportunity to begin interpreting at this point. The problem in this case is restricted mutual audibility caused by the video technology. It may be that the interpreter’s voice has been muted by the noise cancellation function in the technology, or simply that the sound rendered over the loudspeakers is so low in comparison with the doctor’s onsite voice that it is drowned by it.

Another limitation of video transmission concerns mutual visibility. The camera frame only captures a limited part of the interpreter, and this may have serious consequences for the visibility of the interpreter’s gestures to the participants at the ward. Also, the seating arrangements and the placement of camera at the ward may give the interpreter limited visual access to certain participants. We will see an example of this in the next excerpt (8), which follows directly from the end of excerpt (7). The participants at the ward are seated around a table facing each other and thus must turn their head sideways to see the interpreter. The doctor (DR) speaking in the extract is
seated just within the camera frame, and a colleague of hers (MP) is seated to her right, just outside
the camera frame. We start by showing the transcript from the recording at the ward.

Extract (8A): Ward’s perspective

4 DR: .hh (. ) e*: du hadde da en e forgiftning med koka in
       .hh (. ) e: you had prt an e intoxication with cocaine
int: *sits back -->
5 DR: *mattende februar,*
       eighteenth of february
pat: *turns to screen/int and back again*
int: *looks down to notes-->
6 (0.2)
7 .hh og grunnet det hh gjennomgikk du e
       .hh and due to this hh you went through e
8 hjerte o::*g ee:: respirasjonsstans,
       heart and ee: respiratory failure
int: *looks up and down again-->
9 _ (0.5)_
pat. *nods*
10 og var innlagt akutt på </Sykehus/>,
       and was hospitalized urgently at /Hospital/
11 (0.2)
12 .hh e:: så da* det *#var* litt *komplisert forløp,*%
       .hh e:: so then there was a little complicated process
13 dr
14 pat.
15 int: *gaze up, nod * gaze down-->
       *fingertips*
16 fig: #fig8A.1
17 (1.4)
18 DR: *og du *[had]&de da*-
       and you had prt
19 INT: [.hhh]
20 dr: %turns back to pat-->
21 pat: %turns back to table -->
22 int: *gaze up-->
23 (0.4)&(0.4)&#(0.7)
24 mp &stop gesture + gaze to scr & turns to dr-->
25 pat: *turns to int/screen -->
26 fig: #fig8A.2
27 INT: *kan tolken eh* bare& avbryte litt?
    may the interpreter uhm just interrupt a little?
28 int: *finger appears on screen*
29 MP: --> &gaze to screen-->
The doctor produces an extended report of the events in a series of TCUs ending in rising intonation, projecting continuation (line 4-12). During this period, there do not seem to be any signs of the interpreter trying to take the floor. In the middle of the doctor’s utterance in line 12, the participants at the ward can see the interpreter looking up from her notes to the screen, and her fingertips appear briefly at the low end of the screen (see Figure 8A.1). However, the fingertips disappear quickly,
and the interpreter looks down again at her notes and stays in that position during the 1.4-second silence in line 13. During this silence, the doctor, and the patient both gaze at the screen, displaying an expectation that the interpreter will take the floor. However, as this does not happen, the doctor turns back to the patient and resumes the report by initiating a new TCU (line 14). Only in the middle of this utterance does an in-breath from the interpreter (line 15) become audible to the participants at the ward. The medical professional seated outside the camera frame picks up on this and produces a stop gesture, before she turns her head and gaze to the screen and points to the screen, hence allocating the turn to the interpreter (see image 8b.2). The doctor cut off her utterance quickly after the audible in-breath. So, from the ward’s perspective, this looks like a temporary suspension point created by the doctor and the medical professional, in which the doctor pauses extendedly and turns to the interpreter. It also looks like the interpreter does not take the turn allocated to her, but instead comes in only when the doctor has resumed the floor, thus interrupting her.

If we look at the focal lines (12-17) from the interpreter’s perspective, we can see that the situation looks rather different.

**Extract (8B): Interpreter’s perspective**

12   DR: *.hh e:: føså# da det var flitt komplisert forløp,‰*
       .hh e:: so then it was a little complicated process
       dr: --> %to scr
       pat: --> %to scr
       int: * looks up to screen --> *
       int: *stop gest --> £
       fig: #fig8b.1

13  (0.5)
14  INT: *.hh e
15  DR: *%og ëdu håndde da-
       and you had prt
       dr: %turns back to pat-->
       pat: %turns back to local participation framework-->
       int: *gaze up to screen -->
       int: *stop gest£

16  (0.7)*(0.3)
17  INT: *kan #tolken *eh bare* avbryte litt?
       may the interpreter uhm just interrupt a little?
       int: -->*releases*gaze down to notes-->
       fig: #fig8B.2
From this perspective, we can see that the interpreter looks up at the screen already in the beginning of line 12, while the doctor is drawing her breath and producing a hesitation marker. After this, she lifts her hand in a stop gesture, prompting the doctor to refrain from continuing. This is what turned up on the screen in the ward as fingertips in the middle of the utterance due to delay and restricted visual access to the interpreter’s hands. Furthermore, here we see that the silence in the interpreter’s studio is only 0.5 seconds (line 13) before she draws her breath and produces a hesitation marker, thereby projecting turn initiation (line 14). Upon hearing that the doctor continues speaking after
this, she makes a new stop gesture with her hand (line 15), which is not visible at the ward. After this, she hears that the doctor cuts herself off, however, she cannot not see that the medical professional to the doctor’s right, who is seated outside of the camera frame, has turned toward the screen and allocated the turn to the interpreter by means of a gesture. During the silence that follows, the interpreter raises her hand as to take the turn before she begins speaking. Interestingly, she does not just start interpreting, but produces a metacommunicative preface, requesting permission to intervene (line 17). By doing so, she orients to and accounts for the coordination problems that have occurred by taking the blame for ‘interrupting’ the doctor (Robinson 2006).

This excerpt thus shows how both delay and limited visibility make the interpreter’s attempts to take the turn unsuccessful. Her stop gestures are partly or wholly invisible to the participants at the ward, and the doctor’s go-ahead gesture is invisible to her. Furthermore, both her stop gestures and her in-breath occur so delayed in the ward that they come across as misplaced, occurring in the middle of the doctor’s utterances rather than at TCU boundaries. The consequence is a lack of smooth coordination between the parties, leading the interpreter to resort to explicitly requesting permission from the doctor to speak.

4 Discussion

The article has demonstrated how the temporary suspension of a medical professional’s extended contribution for the purpose of interpreting is achieved through joint effort by interpreters and medical professionals in collaboration. The temporary suspension of a multi-unit turn can be occasioned by the medical professionals by designing their turns in shorter installments. These may vary in length and may span from the syntactically, grammatically and pragmatically incomplete to something seemingly complete. By designing a turn in installments the medical professional contributes to creating temporary suspension points, temporarily halting the progress of the turn to allow interpreting. In addition to pausing their speech, they use a range of resources allowing and even inviting the interpreter to speak, such as gazing toward the interpreter, gesturing to the interpreter and explicitly addressing the interpreter. The interpreters can contribute to creating suspension points for interpreting by producing pre-beginning signals, such as audible in-breaths, using gestures, and explicitly asking for the floor. The interpreters’ signals display an orientation to certain points in the medical professional’s longer contribution as relevant for interpreting. The medical professionals can pre-empt these points by continuing past them and not leaving a silence. Similarly, the audible in-breath can serve to halt the medical professional’s ongoing TCU.
While the construction and design of installments may be similar in situations where the interpreter is co-present and in mediated settings, video mediation poses specific challenges to the negotiation of installment lengths. Delay can cause problems for the participants in negotiating the timing of a temporary suspension point. Furthermore, the video mediated environment and the participation frameworks created for video-mediated interpreting in these settings restrict the resources available for signalling attempts at interpreting. Different technological settings and placement of participants and technology may give the participants different visible and auditory access to each other.

Previous research has shown how decomposing a multi-unit turn into smaller chunks and delivering them in installments may be used to elicit listener responses and secure mutual understanding step-by-step (Author 2 2018). This study expands on the use of this practice to include other activity types and other forms of interpolated contributions by interlocutors. In addition, it shows how the length of each installment is a matter of online negotiation between the parties.

While studies on interpreting may focus on the interpreter’s actions, leaving much responsibility on the interpreter for the coordination of the interaction, this study contributes to an understanding of interpreting as interactionally achieved through collaboration. The study shows how participants other than the interpreter are involved in the achievement of the interpreter’s turns.

The findings reported here can be practically useful for practitioners in the field, both interpreters and medical professionals. Professionals can be trained to adapt their speech to interpreting. Interpreters on their side can be and are in many cases trained to signal pre-beginnings. The video-mediated environment provides a complex interactional space challenging the fine-tuned moment-by-moment negotiation of turn space. Awareness of the complexities of negotiating turn space in a mediated environment would be beneficial to both interpreters and medical professionals. The medical professionals may need to leave longer pauses after their turns and the interpreter’s pre-beginning signals may need to be more explicit in order to be perceived by the participants at the ward.

**Acknowledgements**

This work was partly supported by the Research Council of Norway through its Centres of Excellence funding scheme, project number 223265.
References


Author I (2020) xxxxxxx Manuscript submitted for publication.


Li X (2013a) Language and the body in the construction of units in Mandarin face-to-face interaction. In: Reed BS & Raymond G (eds.) Units of Talk Units of Action. Amsterdam: John Benjamins, 343-375


ARTICLE 2: Recruiting repair: Interpreters’ displays of trouble in video-mediated environments

Authors: Jessica P B Hansen

Prepared for submission
Recruiting repair: Interpreters’ displays of trouble in video-mediated environments

Abstract

This article explores interpreters’ embodied displays of trouble in hospital encounters where interpreting is carried out through video-technology. Having other affordances than verbal repair initiators, embodied displays of trouble can serve as a device to initiate repair within the specific interactional and material setting. With an opaque format, the embodied display of trouble resembles recruitments in how it depends on others’ anticipation of what the problem might be. While the embodied display of trouble may serve as an easy and readily available solution within a given setting, it runs the risk of going unnoticed by co-participants. The video-mediated environment proves to be a complex interactional space for recruiting repair. This article explores interpreters’ embodied displays of trouble, the temporal organization of these and the interactional trajectories following.

Key words: other-initiated repair, recruitments, video-mediated interaction, interpreting

1 Introduction

This article explores interpreters’ displays of trouble in a video-mediated environment. The study is based on video-recordings of video-mediated interpreting in Norwegian hospitals. In these meetings, the interpreter is located at a different site than the other participants and participates in the interaction through use of video-technology. Participants in video-mediated interaction use features of the technology to create an interactional space for the undergone activity, both by orienting to the technological artefacts representing their co-participants (e.g. Due, forthcoming) and by using the affordances of the technology to gain access to audio and video from each other’s sites (e.g Due, forthcoming; Luff et al., 2003; Mondada, 2007). The video-mediated environment provides a complex interactional space for organizing interpreting, affording participants with the possibility to create and use a visual ecology in addition to audio.

Drawing on insights from studies of repair and studies of recruitment, this article explores how interpreters’ multimodal displays of trouble are sequentially organized and how they may be used as a less intrusive initial device to recruit assistance promoting the activity of interpreting within
this specific material and interactional setting. How are interpreters’ embodied displays of trouble used in video-mediated environments?

1.1 Repair or recruitment

The organization of ‘repair’ in interaction tends to recurrent problems in speaking, hearing and understanding (Schegloff, Jefferson, & Sacks, 1977, p. 361). Embodied repair-initiators have been found to engender repair unaccompanied by verbal utterances. In ESL tutoring sessions, two different gestures, a sharp head turn or tilt to the side and a head poke accompanied by the movement of the upper body forward, have been found to initiate repair addressing trouble understanding (Seo & Koshik, 2010). A teacher’s use of cupping of the hand behind the ear in the foreign language classroom engendered repair targeting problems hearing (Mortensen, 2016). In interactional meetings and interactions at a customs post, participants’ lifted eyebrow was treated as trouble hearing while a freeze display suspending movements was treated as trouble understanding (Oloff, 2018). Oloff (2018, p. 41) argued that the repairs following these embodied displays were not simply reactions to a co-participant’s lack of response, “they react to specific embodied displays with which the co-participant other-initiates repair in a possible response slot”. Similarly, Seo and Koshik (2010, p. 2221) argue that the gestures “are not merely embodied displays of puzzlement, but are systematically produced actions that engender particular types of responses”.

Recruitments are commonly studied within the context of requests (e.g. Drew & Couper-Kuhlen, 2014; Enfield, 2014; Kendrick & Drew, 2016), and have been found to be embodied ways to elicit someone’s help with something often immediate and local without having to ask. Recruitments at their most explicit have been found to make a physical need, problem or wish overt and publicly available, thereby providing co-participants with an opportunity to assist in meeting the need to resolve a problem (Drew & Couper-Kuhlen, 2014, p. 28). Recruitments have varied degrees of transparency; “The more transparent the display of a need is, the more it assumes an on-record character that is accountable” (Drew & Couper-Kuhlen, 2014, p. 28).

Drawing up a continuum between offers and requests, Kendrick and Drew demonstrate how less transparent displays of need, such as recruitments, may need a higher degree of involvement from another party in order to achieve a solution to the problem (Kendrick & Drew, 2016). Recruitments have been found to target concrete objects that all the participants have equal access to (Drew & Couper-Kuhlen, 2014a, pp. 28-29). Within this understanding of recruitments, linguistic information or permission to do something cannot be recruited, “although they can be requested”
In this article, I argue that the opaque nature of interpreters’ embodied displays of trouble resemble that of recruitments, and that the interpreters use embodied displays of trouble as a resource to recruit repair.

1.2 The orderliness of interpreting

Interpreters’ turns in interaction are often found to respond to other participants’ turns either by providing a rendition of other participants’ turns or by producing other actions, such as asking for clarification (Gavioli & Baraldi, 2011, p. 211). Interpreters’ utterances make parties’ utterances among other intelligible and actionable to co-participants (Paananen & Majlesi, 2018). The interpreter needs to get access to the floor every now and then in order to interpret. As such, other participants’ longer contributions may need to be produced in shorter installments (Davitti, 2019; Hansen & Svennevig, forth.; Licoppe, Verdier, & Veyrier, 2018; Licoppe & Veyrier, 2020). The participants negotiate the length of the installments, for instance through interpreter’s pre-beginning signals such as audible in-breaths, and create temporary suspension points (TSP), a space within the longer turn for the interpreter to interpret (Hansen & Svennevig, forth.) Interpreting enables multilingual interaction, while it is at the same time interactionally and collaboratively achieved by participants in and through the interaction.

In order to produce an interpreted utterance, the interpreter must hear or perceive what was said and understand (to some extent) this, and based on this be able to interpret what was said into the other language. Interpreter’s repair initiators are produced within a locally constituted system for turn-taking specific for the achievement of consecutive interpreting and an embodied participation framework specific for interpreter mediated encounters. Studies of repair in social interaction find that different repair initiators, verbal and embodied ones, can target specific problems in the interaction. While varieties of repair occur in interpreted talk and are mentioned in several studies (e.g. Friedland & Penn, 2003; Li, 2015; Majlesi & Plejert, 2018; Plejert, Antelius, Yazdanpanah, & Nielsen, 2015; Wadensjö, 1998), few studies focus specifically on the nature of interpreter-initiated repair and the interactional trajectory that follows. This article explores interpreters’ displays of trouble in the video-mediated environment and the displays’ sequential positions and interactional trajectories. The study investigates if studies of recruitment can provide a framework for understanding the interpreters’ embodied displays of trouble.
2 Methods and data

The study builds on video recordings of interpreting in hospital encounters. The dataset consists of video-recordings of eleven hospital encounters with the interpreter participating through videotechnology. The interpreting is carried out in seven different languages: Albanian, Arabic, Bosnian/Croatian/Serbian, Mandarin, Polish, Thai and Vietnamese. The video-recordings are made in different wards in Norwegian hospitals. Some of the meetings are with admitted patients while some are from outpatient clinics. All the interpreters included in the study have formal qualifications such as interpreting education. Three of the meetings are video-recorded from the interpreter’s point of view, one is recorded from the ward, and eight are recorded from both the ward and the interpreter’s point of view. The analysis draws on insights from recordings made on both sites.

This analysis builds on the theoretical framework of multimodal conversation analysis (e.g., Deppermann, 2013; Hazel, Mortensen, & Rasmussen, 2014; Mondada, 2014) and the transcriptions are made using Mondada’s (2001) system for multimodal transcription. In conversation analysis, multimodality includes the relevant resources that are mobilized by participants to build and interpret the public intelligibility and accountability of their situated action (Mondada, 2018). Such resources may be for instance grammar, lexicon, prosody, gesture, gaze, body postures and movement (Mondada, 2018, p. 86). Conversation analysis’ analytic aim is to study social interaction from an emic perspective, to take the participants’ perspective in analysis. In the study of the multimodal organization of interaction, this gives raise to an analytic question: Is a gesture or other bodily conduct used systematically and recognizably and therefore accountably as a resource in social interaction (Hazel et al., 2014; Mortensen, 2016)? This analysis is concerned with embodied displays of trouble and the following interactional trajectories. Although the embodied displays of trouble do not always receive displays of uptake, I argue that these are systematically, and accordingly accountably produced as a resource in the interaction.

---

4 The study is granted approval by the Norwegian Centre for Research Data and the involved hospitals and wards. All participants have given informed consent. The interpreters included in the study have formal qualifications.
3 Analysis

The analysis presents extracts where interpreters produce embodied displays of trouble in video-mediated environments. The first example demonstrates how leaning forward engenders repair. The following examples demonstrate how the video-mediated environment and orders of interpreted interaction may challenge the embodied display of trouble as a resource to recruit repair. While the first three examples demonstrate issues where the embodied display targets problems that can be solved with repair, the final example demonstrates how embodied displays of trouble are versatile and can be used to address other issues, such as auditory disturbances.

3.1 Leaning forward engenders repair

Extract (1) demonstrates how leaning forward is an interactional practice that can engender repair. Patrycja (PAT) who speaks Polish is meeting with several medical professionals who all speak Norwegian. Patrycja and the medical professionals are in a meeting room equipped with a videoconference system. The interpreter (INT) participates in the meeting from a different location using video-technology and sees the participants at the ward on her screen. At the ward, the camera and the screen displaying the interpreter are placed at the end of the table, and Patrycja is seated at the other end of the table, facing the camera and screen. Figure 1.1 below illustrates what the setting looks like from the interpreter’s perspective. The extract is from the opening of the meeting. One of the medical professionals (MP) has just welcomed Patrycja to the meeting and this has been interpreted into Polish. When asked if she knows all the participants, Patrycja refers to the speech therapist (ST) as unknown.
Extract 1

1 MP: 
   som vi sa i sta-, 
   as we said earlier
2  
   du kjenner alle rundt bordet her nå. 
   you know everyone around the table here now
3  
   ikke sant? 
   isn’t that so
4 INT: 
   .h tak jak (.) mówiliśmy przed chwilą 
   .h so like (.) we said earlier
5  
   wszystkich wokół tego stołu pani zna prawda? 
   you (MRS, polite fem) know everyone around this table right?
6  (1.8)
7 PAT: 
   nie może pani #znać?# 
   there is someone here who MRS does not know
   #turns left, touches ST#
8  ST: 
   ja 
   yes
9 PAT: 
   hm [he he he] he he he 
10 ST: 
   [hm hm hm]
11  
   → (0.3)*#(0.6)  
   → int: 
       *head forward--->
       #fig1.2
12 PAT: 
   &(#jed:na) jest jedna osoba której #nie ;znam
   there is one person who I don’t know
   pat: &gaze up toward screen/camera
13  
   ale myśle*# że zaraz z tym nadrobimy. 
   but I think that we will catch up with this soon
   → int: 
       *retracts position, gaze to scr-->
       #fig1.3
14 INT: 
   .h (. ) e: det er en person som jeg ikke kjenner, 
   .h (. ) e: there is one person who I do not know
Figure 1.2 (left) & 1.3 (right): Interpreter pokes head forward and releases.

Orienting to the possible relevance of the other-introduction of parties in the meeting (Pillet-Shore, 2011), the medical professional suggests that Patrycja knows all the people present and completes the utterance with a turn final tag question (line 1). The Norwegian utterance is directed to Patrycja through use of the second person singular pronoun, “du” (you). While the question format would usually allocate the turn to the addressee or at least suggest that a response is relevant, the speaker speaks Norwegian and the addressee speaks Polish. The interpreter will therefore interpret the utterance into Polish before Patrycja responds.

The interpreted utterance (line 4-5) is designed in Polish with a structure similar to the medical professional’s utterance with a turn final tag question. The interpreter uses the female polite form of address in Polish, “pani” (Mrs). Patrycja responds in Polish, stating that there is someone in the setting that she does not know (line 7). Jokingly she mirrors the interpreter’s polite phrasing, referring to herself with the polite form “pani” (line 7). Although the Norwegian-speaking participants most probably do not understand what Patrycja said, Patrycja identifies the speech therapist to her left as relevant to the question that was asked through bodily orientation, gaze and touch. The speech therapist confirms immediately without waiting for the utterance to be interpreted, as such displaying some understanding of the content possibly based on the combination of the content of the question and Patrycja’s embodied actions. Patrycja’s response to the medical professional’s initial question both addresses the topic raised, plays on the interpreter’s formal phrasing in Polish and identifies through gaze and gesture the speech therapist as unknown.
As such, different participants have different possible levels of access to Patrycja’s multimodal utterance at different times.

Patrycja and the speech therapist laugh (line 9-10). The speech therapist has already responded, but does not go on to introduce herself, and the medical professional who opened the meeting does not reclaim the floor. The interpreter does not begin to interpret. Rather, she leans forward toward the desktop unit displaying the participants at the hospital (line 11). By doing so, the interpreter displays that there is a problem. Her change of posture does not make transparent to other participants what the problem is, beyond that she for some reason cannot interpret. As a repair initiator, the interpreter’s embodied display of trouble resembles the format of an open class repair initiator leaving it up to the producer of the trouble source utterance to identify the problem (Drew, 1997). Although the open-class repair initiator may often be treated as a problem with hearing (Drew, 1997; Svennevig, 2008), in this case, Patrycja does not merely repeat her previous utterance. By elaborating on her initial utterance, Patrycja treats this as a problem with the comprehensibility of her utterance. She repeats that there is someone present that she does not know by using a more common phrasing and does not repeat or explain the wordplay (line 12-13). She suggests that the status of their acquaintance will soon change, perhaps downplaying the urgency of this matter. By having to repeat and explicate what was initially a short and joking remark, the content of Patrycja’s comment becomes more emphasized and may perhaps be perceived as more serious and even urgent. Patrycja directs the utterance to the interpreter by looking up toward the screen and camera while speaking. Still, the interpreter renders Patrycja’s utterance into Norwegian (line 14-15), making relevant an introduction, and the medical professional introduces herself as the speech therapist (line 16).

The interpreter releases the embodied display of trouble and sits back again as Patrycja reformulates her utterance (fig 1.3) (line 13). While the interpreter’s initial change of posture indicates that there is a problem, it leaves open to the participants at the ward to propose a solution to the problem. In this way, the interpreter’s embodied display of trouble resembles the format of a recruitment. The interpreter’s action enables other participants to recognize or anticipate that there is a problem, but does not make clear what the resolution may be (Kendrick & Drew, 2016, p. 11). The sequential positioning of the interpreter’s change of posture indicates that the problem is interactional. The activity of interpreting now depends on, if not the assistance of other participants, at least the actions of others in the interaction. The problem is solved when the previous speaker reformulates the previous utterance and, in this way, treating the problem as a problem of understanding. The interpreter’s embodied display of trouble resembles a recruitment in format while it engenders repair.
3.2 Embodied displays of trouble within the video-mediated environment

In extract 1, Patrycja was facing toward the screen representing the interpreter. In extract 2, Pawel, the speaker of the trouble source turn is seated near the videoconference system. This requires him to turn to the left in order to see the interpreter. Extract 2 demonstrates how the embodied display of trouble depends on other participants’ gaze and visual attention and furthermore how the embodied display of trouble can be expanded with a verbal utterance.

Pawel is responding to a question from one of the medical professionals concerning how he is planning to keep active when he returns home from the hospital. He mentions a range of activities he will carry out at home before he mentions one he will not be doing at home.

![Fig 2.1](image1) ![Fig 2.2](image2) ![Fig 2.3](image3)

**Extract 2**

1  PAW:  (wiadomo) że na# stolarni nie będę
   it’s clear that I will not work at the
   paw:     >>gaze to interpreter-->
   fig:    #fig2.1
2  robil   ale&
   carpenter’s workshop but
   paw:     &to participants at the ward-->
   fig:    #fig2.2
3  (0.7)*(0.4)
4  int:  *leans forward-->
4  INT: .h prz#e&praszam, pop- pro-
   sorry
   paw:     &turns to screen/camera-->
   fig:    #fig2.2
5  PAW:  na stolarni nie będę robil ale
   I will not work at the carpenter’s workshop but
   hhh *hhh bo* (spalanie)robiłem na stolarni
   hhh  hhh because we did work in the workshop
5  int:  *smiles*bites lip-->
   (0.4)
Pawel is visually oriented to the camera and the screen displaying the interpreter while talking before he turns away from the technology and back to the participants at the ward as his utterance comes to completion (line 2). During the silence following Pawel’s turn, rather than interpreting the utterance, the interpreter leans forward toward the videoconference unit (fig2.2) (line 3). Pawel is facing in a different direction and cannot perceive the interpreter’s change of posture. Resources that might be effective in face-to-face interaction, for instance to get the attention of a co-participant, are not always effective in the video-mediated environment (Heath & Luff, 1993; Hutchby, 2001). Upper body movement is more prominent when carried out by participants who are physically co-present than when depicted on a screen in participants’ peripheral vision. Within this specific socio-material setting, the interpreter has only limited possibilities to coordinate the display of trouble with the co-participants’ line of sight (Kendrick & Drew, 2016, p. 15). Possibly recognizing that Pawel will not be able to see her gesture when he is turned away from the screen, the interpreter begins to produce a verbal repair initiator before her movement comes to a halt (line 4). She produces a short verbal open-class repair-initiator in Polish “przepraszam” (I’m sorry) while holding the position leaning forward. The open class repair initiator does not identify the specific problem in the interaction but leaves it up to the speaker of the trouble source turn to identify the problem. Pawel turns back to the interpreter and repeats what he initially said accounting for the supposition of his statement (line 5-6). The interpreter responds to his utterance producing a change-of-state token, “åja” (oh), in Norwegian before she releases her position, indicating that the problem has been resolved, and interprets the utterance (line 7).

Lack of interpreting in itself does not necessarily engender repair. The participants did not display an orientation toward a possible problem until the interpreter began producing the verbal repair initiator. By supporting or expanding the display of trouble with a verbal repair initiator while still moving forward, the interpreter utilizes an assumedly more effective resource within the specific setting. Similarly, extract 3 demonstrates how the interpreter’s embodied display of trouble does engender repair. However, the participants have trouble timing their actions due to delay. The

---

5 Although “åja” is a third position receipt to the repair sequence, it is not produced in the language of the repair sequence but in the language of the interpreter’s next utterance, the interpreted rendition in Norwegian. This may be a demonstration of the complexity of the linguistic setting. This also occurs in extract 3.
interpreter produces several verbal repair initiators as she and the speaker of the trouble source turn have problems synchronizing the repair sequence. Extracts from the interpreter’s studio and the ward are juxtaposed in this example to demonstrate how the sequence emerges differently at the two sites due to delay in the transmission of signals between sites.

Alesandro (ALE) who speaks Albanian is meeting with six medical professionals. Just prior to the sequence in this example, one of the medical professionals asks her colleague (MP) if she has a plan B if the patient is not accepted for the treatment for which they have applied. The interpreter has interpreted this from Norwegian to Albanian, and after a long silence, both Alesandro and the medical professional begin to speak at the same time. The image to the left illustrates the setting from the wards perspective. The interpreter is displayed on the screen at the end of the table, and Alesandro is seated to the left of the screen. The image to the right demonstrates how the participants are displayed on the screen in the interpreter’s location. The first extract is transcribed from the interpreter’s perspective.

Figure 3A: Ward’s perspective to left: patient seated to the left of the screen. Interpreter’s perspective to the right: patient displayed partially covered by interpreter’s selfview.

Extract 3A: Embodied displays and delay in the media

1   (2.6)
2   MP:  ts *.hh #(0.3)  [e:  ]
3   ALE:  *[fitn]ess
ale:  *gaze to int*turns back to participants at ward-->
fig:  #fig5.1
4   *(0.2)≈(0.8)  ≈(0.3)
ale:  *turns to interpreter--->
int:  ≈head forw.#holds pos.--->
     #fig5.2
5   INT:  unn[skyld]
          sor[ry ] (NOR)
6   ALE:  *[fi  ]ness
int:  ->holds head position-->
Alesandro gazes toward the camera and the screen displaying the interpreter (line 3), before he turns away from the screen and back to the table where the medical professionals are seated and suggests “fitness” as an alternative to the treatment (line 3). The medical professional produces a pre-beginning vocal sound at the same time (line 2), which overlaps with Alesandro’s utterance. She cuts herself off and lets him complete. By turning to the table while producing his utterance, Alesandro has directed the utterance to his co-participants at the ward. Right after having made this suggestion, he turns back to the camera and screen (line 4) indicating that the interpreter will be the next speaker. Overlapping talk can be difficult to perceive through video-technology. Both utterances were short, making it difficult to make out even parts of their utterances. Rather than interpreting, the interpreter leans forward displaying that there is a problem. Alesandro is turned to the screen, so this is visually available to him. After having held the position for 0.3 seconds, the interpreter produces a verbal repair initiator in Norwegian, “unnskyld” (sorry), while still holding the position (line 5). The interpreter is speaking Norwegian and gazing to the left side of the screen where only medical professionals are displayed, which indicates that she has not yet identified Alesandro as the speaker of the trouble source turn. Overlapping with the final syllable of the interpreter’s verbal repair initiator, Alesandro reiterates his suggestion “fitness” (line 6). While these two utterances are partially overlapping from the interpreter’s perspective, extract 3B from the ward reveals how the utterances are ordered differently at the ward.
Extract 3B: Delay and the re-ordering of turns

1 (2.0)
2 MP: ts *.h (0.3) [e: ]
3 ALE: * [fitn]*ess
   ale: *gaze to int *turns to parts at ward-->
5 * (1.0)##(0.6) ##(0.1)
   ale: *turns to interpreter---->
   int: #head forw.#holds pos.---->
   #fig3b.1#fig3b.2
6 ALE: fitness=
7 INT: =unnskyld
   =sorry (in Norwegian)
   int: >-holds head position--->

Fig 3b.1

Fig 3b.2

Alesandro is gazing at the screen representing the interpreter in the ward, and her embodied display of trouble is visible at the ward after a second of silence (line 5). Alesandro does not treat the silence alone or the movement itself as a display of trouble, but repeats his utterance only after the interpreter leans forward and holds the position. Alesandro repeats his utterance (line 6) and as such treats the interpreter’s posture change as a repair initiator. By repeating what he said the first time, Alesandro treats this as a problem hearing. The interpreters’ displays of trouble leave it up to the speaker of the trouble source turn to identify what the problem with the previous utterance might be as such, the displays of trouble are not indicative of the nature of the problem.

Latching onto Alesandro’s utterance, the interpreter now produces an open-class verbal repair initiator in Norwegian (line 7), “unnskyld” (sorry). From the interpreter’s point of view, the interpreter began producing the Norwegian repair initiator before Alesandro produced the repair and the two utterances overlapped. From the ward’s point of view, Alesandro repeated the word before the interpreter produced the verbal repair initiator. Due to delay in the transfer of signals, the two utterances have a different order at the two sites. The problem has not yet been resolved.
Returning to the interpreter’s perspective, the extract below demonstrates how the repair sequence continues to emerge as the interpreter identifies Alesandro as the speaker of the trouble source turn.

Extract 3C: Identifying the speaker in the mediated environment

<table>
<thead>
<tr>
<th>Line</th>
<th>Interaction</th>
<th>Language</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0.4 (0.1)</td>
<td></td>
<td>turns head to the right, holds--&gt;</td>
</tr>
<tr>
<td>8</td>
<td>INT: çfarë the?</td>
<td></td>
<td>what did you say?</td>
</tr>
<tr>
<td>8</td>
<td>ALE: palestër.</td>
<td>gym (ALB)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After yet another silence, while still holding the position, the interpreter turns her head, now seemingly gazing toward participants displayed on the right side of her screen (line 7). This is the area where Alesandro is seated. Still holding the position leaning forward, she now produces an open-class repair initiator in Albanian (line 8). Through change of visual orientation and language choice for the production of the repair initiator, the interpreter displays that she has identified Alesandro as the speaker of the trouble source turn. From the interpreter’s perspective there is a 0.4 second silence following her utterance (line 8) before Alesandro’s responds, now choosing a different word, the more typical Albanian word “palestër” (line 9). Returning to the wards perspective, the following extract shows how the problem is resolved and the sequence is closed.

Extract 3D: Resolving the problem

<table>
<thead>
<tr>
<th>Line</th>
<th>Interaction</th>
<th>Language</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>INT: palestër %*#fitness</td>
<td>gym (Alb)</td>
<td>fitness #</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>int:</td>
<td>retracts head, gaze straight forward---&gt;</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>ale:</td>
<td>*leans in to table, turns to the table&gt;&gt;</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>fig:</td>
<td>#fig3d.1#fig3d.2</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>MP: [ja]</td>
<td>yes</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>INT: [tr]enigsstudio</td>
<td>gym ((in Norwegian))</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td>(0.4)</td>
</tr>
</tbody>
</table>
Although the silence that follows Alesandro’s utterances is 1.1 seconds long from the ward’s perspective (line 1), Alesandro does not treat this as an indication that there is still a problem. Following the silence, the interpreter first repeats the word Alesandro used last, “palëster”, and retracts the position as she repeats the word he used first, “fitness” (line 2). Simultaneously, Alesandro turns away from the screen and back to the medical professionals seated around the table (line 2) displaying that he recognizes that the problem has been resolved. By repeating the word Alesandro first used, the interpreter shows that she has not only perceived what he said now, but made sense of what he was saying earlier as well. After yet some silence, the interpreter continues “ja” (yes)\(^6\) in Norwegian (line 4) although the repair sequence was between her and the Albanian speaker and continues to produce a Norwegian word for gym, “treningsstudio” (line 7).

In these settings, participants are speaking several languages. In order to address verbally problems that arise, the interpreter has to choose a language for the production of the utterance. In order to do so, the interpreter must be able to identify the speaker and the language they speak. Extract (3) has demonstrated how, when overlapping talk occurs in this complex material and linguistic setting, identifying the speaker and accordingly the relevant language might be difficult in the mediated environment. The extracts above have demonstrated how embodied displays of trouble can engender repair within a complex participation framework and that embodied resources allow the interpreter to display that there is a problem quickly without producing a verbal turn and consequently without choosing a language. However, in both extracts (2) and (3) the interpreters expand the displays of trouble verbally, in extract (2) in order to secure the attention of the speaker of the trouble source turn and in (3) to enhance the display of trouble after assumed lack of uptake. In both cases, the expansion is produced as a verbal repair initiator drawing not only on the symbolic affordances of verbal talk as such but as auditory cues as a way to attract attention. As the verbal

---

\(^6\) Similar to in extract 2, the interpreter produced a third position response in Norwegian after the repair sequence was resolved although the repair sequence was in fact conducted in the other language.
repair initiator seems to have affordances that work well in this environment, why do not interpreters begin with the verbal repair initiator? Considering the complexity of the setting, producing an embodied display of trouble may simply be easier. The modality of the action allows it to be produced simultaneously to speech and as such does not require the same attention to the unclear boundaries of turn taking in interpreted talk and hence the risk for overlapping talk. Furthermore, the embodied display of trouble does not require the interpreter to identify the speaker of a trouble source turn and choose a language. Finally, the embodied display of trouble does not require a response and can be carried out simultaneously to ongoing talk. As such, the interpreters may be trying the easiest solution first (Svennevig, 2008).

3.3 Same embodied format for a different problem

The three prior examples have demonstrated how interpreters’ embodied displays of trouble are positioned at temporary suspension points or transition relevance spaces in the interaction and are as such organized relative to the ongoing talk. Extract (4) demonstrates how an embodied display of trouble can be produced simultaneous to ongoing talk, initiating a simultaneous line of conduct, to indicate early that is a problem.

At the beginning of a meeting with several participants, a medical professional (MP) produces an utterance while visually oriented away from the screen displaying the interpreter. While she speaks, another participant moves a notepad across the table and shortly after another does the same. The shuffling causes noise while the medical professional is speaking, making it difficult to hear what she is saying through the technology. The interpreter leans forward shortly after the noise occurs.
Extract 4

1 → MP: %mm. (0.2) e: ja,*# (0.5) Piotr er* jo en ffin *mann #få
        mm e: yes Piotr is a good man to
    mp: %turns to Paulina--->
        → int: *gaze notepad-*gaze up-*leans forward-->
    mp1: *moves notepad
    mp2: *moves notepad
    #fig4.1
    #fig4.2
7 2 MP: samarbeide me:,
        collaborate with
        (0.2)
7 3 INT: .h
7 5 (0.2)
6 MP: posi*tiv# type.
        a positive guy.
    → int: -*turns right ear to screen-->
        #fig4.3
7 9 (0.6)
8 INT: e:*m:
    mp: *turns to screen/camera-->
7 9 (0.5)
10 → INT: e- k*unne du gjentatt?
        e- could you repeat?
    → int: -*releases head position, gaze still forward-->
11 → tolken fikk ikke med seg hva som ble sagt.
        the interpreter did not get what was said.
12 (0.6)

Objects being moved across the table causes auditory disturbances, which is picked up by the microphone. The disturbance makes it difficult to distinguish what the current speaker is saying (line 1). Shortly after the disturbances begin, the interpreter leans toward the screen and camera (figure 2.2) and holds this position during the progress of the turn (line 1-2). The interpreter begins to change the position of her upper body as the medical professional is still speaking. This is a clear distinction from the previous examples, where the interpreter’s actions were occasioned by the previous utterance and produced in sequential order relevant to the problem source turn. In this case, the interpreter’s embodied action is structured as an action simultaneous to the talk. The interpreter treats the short silence at the possible completion of the medical professional’s turn as a possible temporary suspension point, a point in the interaction relevant for interpreting (Hansen &
and produces an audible in-breath (line 4), a pre-beginning signal. However, the medical professional does not suspend her turn but continues to speak (line 6). While the medical professional is still speaking, the interpreter increases the intensity of her embodied display of trouble by leaning even further in towards the screen and turns her ear to the screen (line 6, figure 4.3). While the first display of trouble, leaning forward, is perhaps opaque as to indicating what the problem is, turning an ear to the screen is more transparent and suggests that the problem is related to hearing. The medical professional who produced the utterance is seated to the side of the table, facing away from the screen, and directs the utterance toward the participants at the ward. The interpreter on the screen is not within the medical professional’s sightline, and movement displayed on a screen is less likely to be perceived in the peripheral vision. The medical professional does not indicate that she has perceived the interpreter’s embodied display of trouble. The interpreter continues to hold the position during the silence following the medical professional’s turn (line 7). However, this display still does not receive any uptake. While the initial problem was caused as participants moved objects near the microphone and could be solved while the turn was still emerging if the participants ceased to do so, a new problem has now emerged; the interpreter is prevented from interpreting due to not having heard what was said. The interpreter produces a vocal sound, which gets the attention of the previous speaker who now turns to the screen (line 8). After a further silence (line 9), although the problem is not resolved, the interpreter releases the embodied display of trouble and sits back as she begins to produce a verbal repair initiator (line 10). The interpreter does not specify whether the trouble is hearing or understanding, but accounts for the request using the ambiguous phrase “få med seg” (to get) what was said.

The interpreter’s embodied display of trouble in this extract differs from the previous examples when it comes to temporal organization and the following order of events. Most repair initiations by parties other than the speaker of the trouble source turn occur in the position of the next turn. Trouble source turns are rarely interrupted by other-initiation of repair (Schegloff et al., 1977). This example demonstrates how the interpreter’s embodied display of trouble targets a different problem, auditory disturbance, and is produced during the ongoing turn. Drawing on the visual ecology afforded by the media the interpreter’s embodied action was designed and positioned interactionally so it could recruit assistance with the ongoing problem – the noise made by other participants – without interrupting the ongoing turn. The embodied display has other possibilities, affordances, and constraints than verbal utterances. Affording simultaneity of actions, it is less intrusive in the interaction than a verbal utterance in the same position, as it can be produced during the ongoing verbal activity without interrupting. As such, the interpreter treats visual resources as less intrusive, and chooses the less invasive modality first (Frøili, 2001; Svennevig & Marstrand,
Furthermore, the embodied display of trouble could invite participants to solve the problem immediately either by halting the actions causing disturbance or by the speaker producing self-initiated self-repair thus promoting the progression of the interaction. However, this would require that the interpreter’s actions were perceived by the participants at the ward. Seeing that the participants are gazing to each other – and away from the screen – the interpreter’s actions are not perceivable to the current speaker. In the prior examples, when the interpreters expanded the display of trouble with a verbal repair initiator, they held the position until the problem was resolved. In this case, the interpreter released the position as she began to produce the verbal repair initiator, indicating that the two actions, one embodied and one verbal, are two separate actions. The first is an attempt to recruit help to halt the disturbance. The second is a repair initiator. If the interpreter’s action had been perceived and the disturbances halted at an earlier point, there might not be need for repair – the repair sequence could have been prevented altogether.

4 Discussion and conclusion

In this study, I have explored interpreters’ embodied displays of trouble in complex interactive settings. The analysis has demonstrated how interpreters use embodied displays of trouble to address different types of problems. Three of the extracts in the analysis demonstrated how the interpreters’ gestures addressed a problem that could be solved through repair. The embodied display of trouble was produced at a point in the interaction relevant for interpreting and was as such organized relative to the temporality of the ongoing talk. The final example demonstrated how the same embodied display of trouble was produced while the current turn was still ongoing, a point in the interaction not relevant for interpreting, and as such, invited action in a line of conduct simultaneous to the ongoing one. The embodied display was used here to address a problem other than repair. However, the display of trouble did not receive uptake. Since the problem was not solved, the embodied display of trouble was abandoned and followed by a repair initiator. In all cases, the problems could only be solved with assistance from participants at the ward.

Since embodied resources afford simultaneity without interrupting ongoing talk, embodied resources run less of a risk of resulting in overlapping talk. With the unclear turn boundaries of consecutively interpreted talk, embodied displays of trouble could be a useful resource to solve problems. However, the interactional settings in which embodied display of troubles are produced are quite complex. The participants use technology to create an interactional space for their conduct. While the technology affords the participants with visual and auditory access to each other, visual
access depends on the camera frames participants choose and how they use the screen displaying each other in the organization of interaction. The participants at the ward are seated around a table, and direct their utterances to each other, gazing toward the interpreter now and then. Securing the sightline of participants at the ward is not always possible for the interpreter, and the embodied display of trouble does not attract attention from a participant gazing in another direction. Embodied actions have different affordances when being displayed on a screen compared to what they do when perceived by a co-participant at the same location and do not seem to get attention in the same way in the video-mediated environment (Heath & Luff, 1993). Transmission delay is a feature of the video-mediated environment that may change the temporal unfolding of actions at each site.

The embodied display of trouble serves as a versatile device that engenders repair addressing trouble hearing and trouble understanding. The same bodily movement in a different sequential position can be used to address ongoing problems with disturbances at the ward. Furthermore, the embodied display of trouble does not require the interpreter to identify or even choose a language in which to produce a verbal repair initiator. These examples have shown how initiating repair can be quite complex in these interactional, linguistic and material settings and that embodied displays of trouble may serve as a recruitment to elicit the help of others in performing repair. Participants treat the embodied displays of trouble as less intrusive than verbal requests and accounts. However, the embodied display of trouble, being more opaque than a verbal request or repair initiator might easily be overseen. In several of the examples, the embodied display of trouble were expanded with verbal support in order to elicit the needed assistance. While they may go unnoticed by co-participants, they are systematically produced and interactionally organized, and the following trajectories of the interaction suggest that they are not merely displays of puzzlement.

The concept of ‘recruitment’ is developed to encompass the linguistic and embodied ways in which assistance may be sought – requested or solicited – or in which someone comes to perceive another’s need and offer or volunteer assistance (Kendrick & Drew, 2016, p. 1). Kendrick and Drew suggest that recruitments for assistance are distributed along a continuum from the most explicitly articulated requests to the less explicit articulations, such as embodied displays of trouble and even projectable trouble where the problem is projected by the other participants (Kendrick & Drew, 2016). Recruitments rely to a larger degree on others anticipating that there is a problem and what it might be. Similarly, the examples above, although most of them demonstrated how the interpreter’s embodied display of trouble targets interactional problems, demonstrate how the embodied display of trouble leaves it up to other participants to identify the problem and assist in solving the problem. Repair initiators can be produced distributing different levels of responsibility for identifying a problem on the producer of the repair initiator and the producer of the trouble
source turn. Embodied actions have been found to initiate repair and different embodied actions have been found to elicit different types of repair within different settings. Based on the analysis above, it might be relevant to consider continuums similar to those of recruitments in the study of the organization of repair.

5 References


ARTICLE 3: Invisible participants in a visual ecology: Visual space as a resource for organising video-mediated interpreting in hospital encounters

Authors: Jessica P B Hansen

*Social Interaction. Video-based Studies of Human Sociality, in press*
Invisible participants in a visual ecology: Visual space as a resource for organising video-mediated interpreting in hospital encounters

Abstract

This paper explores multilingual hospital encounters in which medical professionals and patients do not speak the same language, and where interpreting is facilitated through the use of video technology. The participants use video technology to create an interactional space for interpreting. While video technology affords the participants visual access to each other, and the participants may use embodied actions in interaction, participants in interaction do not necessarily organise their interactional space in ways that secure congruent views of each other. While the participants’ incongruent views of each other may cause problems in the organisation of interaction, the participants rarely discuss the visual setting. This article explores how the participants orient to the visual materiality of the setting and how they use the visual ecology they create, in and through the interaction, to best achieve the multilingual activity of interpreting in hospital encounters.

Keywords: visual ecologies, interactional space, video-mediated interaction, multimodality, interpreting, hospital interaction

1 Introduction

Norway is a geographically vast country with a widely dispersed population. Video-technology is believed to increase access to qualified interpreters across the country, and accordingly facilitates equal access to public services, such as healthcare, to an increasingly diverse population (e.g. Det Kongelige Kunnskapsdepartement, 2019; NOU, 2014:8, 2014). Trials of video technology for the purpose of interpreting have been conducted in Norway since the late 1990s (Skaaden, 2001). However, estimates indicate that only 0.5% of interpreting assignments in the public sector are carried out using video technology (Det Kongelige Kunnskapsdepartement, 2019). Video-mediated interpreting is widely considered to be a superior option to using the telephone, since this technology
affords participants visual access to each other. This has encouraged governmental bodies to argue for an increased use of video-mediated interpreting in Norwegian public services (Det Kongelige Kunnskapsdepartement, 2019; NOU, 2014:8, 2014).

While documents that propose an increased use of video-mediated interpreting emphasise the visual affordance of the media (e.g. NOU 2014:8, 2014), little is known about how participants in interaction use the visual affordance of the media when accomplishing interpreting. Interpreting studies have compared the quality of video-mediated interpreting to onsite interpreting, and in some cases to telephone interpreting, based on simulations and role-plays (e.g. Balogh & Hertog, 2012; Braun & Taylor, 2012; De Boe, 2020), however, “the potential benefits of an audiovisual channel compared to an audio-only channel” have yet to be explored (De Boe, 2020, p. 57).

Few studies have explored video-mediated interpreting in medical encounters using naturalistic data. Based on video recordings of video-mediated interpreting in hospital encounters, this study explores how participants use and orient to the visual affordance in the organisation of interpreting in hospital encounters. The study tends to the organisation of interpreting in a video-mediated environment as something “in its own right” (Dourish, Adler, Bellotti, & Henderson, 1996), and aims to explore the utilisation of the media’s visual affordance in these meetings from a member’s perspective (Arminen, Licoppe, & Spagnolli, 2016).

1.1 Interpreting as interaction

Interpreting facilitates interaction in encounters where participants do not speak the same language. Interpreters’ turns in interaction often respond to other participants’ turns, either by providing renditions of their turns, or through other actions, such as asking for clarification (Gavioli & Baraldi, 2011, p. 211). Interpreter mediation may place certain constraints upon – and also open up possibilities regarding – the organisation of interaction, for instance concerning turn-taking (e.g. Davitti, 2019; Gavioli & Baraldi, 2011; Hansen & Svennevig, forth.; Licoppe, Verdier, & Veyrier, 2018). In their work, interpreters have been found to “continuously monitor and analyse the unfolding interaction and make moment-by-moment decisions about their actions” (Bolden, 2018, p. 135). There is a growing field of studies that acknowledge and explore interpreting as interaction and, consequently, the interpreter as a participant in interaction (e.g. Angermeyer, 2007; Bolden, 2000; Davitti, 2019; Gavioli & Baraldi, 2011; Li, 2015; Paananen & Majlesi, 2018; Wadensjö, 1998). Interpreting enables multilingual encounters, while at the same time it is interactionally and collaboratively achieved by participants in and through the interaction.
Interpreters’ and medical professionals’ behaviour is informed by the norms and conventions of professional practice, whereas interpreting is organised in situ. For instance, interpreters are expected not to contribute to the substance of the conversation (Wadensjö, 1998, p. 67). A similar position can be found in interpreter training, where interpreters might be explicitly encouraged to avoid “getting in the way” of the interaction, for instance by engaging primarily in interaction with one of the parties present (Skaaden, 2013, pp. 151–155). For medical professionals, guidelines for working with interpreters tend to promote a simplistic view of interpreting, based on an understanding of the interpreter as a conduit of information, and are inclined to relegate the interpreter to the role of a side-participant or an overhearer in the interaction (Li et al., 2017). Participants in interpreted hospital encounters deviate from the practices suggested in such guidelines in order to solve interactional problems and accomplish the interaction (Hansen, 2018).

Participants’ understanding of the interpreter’s role is reflected by their actions in interaction. For instance, in a French courtroom setting in which the defendant was sitting remotely, the participants’ view of the interpreter’s role was demonstrated by how the camera operator actively chose to frame the interpreter, and how this framing was linked to talk-in-interaction (Licoppe & Veyrier, 2017). The interpreter “is ‘enough of a speaker’ that s/he should be made visible, but not ‘enough of a speaker’ that the other parties for whom he/she is interpreting may be visually ignored” (Licoppe & Veyrier, 2017).

1.2 A visual ecology for interpreting

Schutz’s (1953) concept of reciprocity of perspectives suggests that in common-sense thinking, “the world taken for granted by me is also taken for granted by you”. This presupposed reciprocity of standpoints is made possible by what Schutz calls the idealization of the interchangeability of standpoints. He suggests that people take it for granted that if were thiey to switch places, they would see the world in the same typicality as the other (Schutz, 1953, p. 53). In video-mediated interaction, participants have been found to presuppose that they have visual access to each other (Luff et al., 2003) and have difficulties ensuring that they have congruent views of each other in the interaction (Arminen et al., 2016; Luff et al., 2003). While video technology emulates co-presence at a distance, video-mediated environments have proven to constitute a complex interactional space (Arminen et al., 2016).

However, the interactional space is not predetermined by technology – rather, the technology serves as a resource available to the participants, with which they can shape a space relevant to their
ongoing work (Mondada, 2007, pp. 51–52). Studies of video-mediated interpreting in authentic settings have addressed issues of spatial arrangement and the use of camera actions in courtroom hearings (Licoppe & Verdier, 2013; Licoppe & Veyrier, 2017). Participants not only create a mediated space they find appropriate for the activity unfolding, but also display an understanding of the activity itself and frame it accordingly (e.g. Licoppe & Veyrier, 2017). This study focuses on interpreted interaction in which the interpreter – the person who knows both languages and who participates in the interaction to facilitate the multilingual encounter – participates from a remote location. This article explores ways in which the participants in medical encounters use and orient to the visual ecology provided by the media when accomplishing interpreting. How do participants in interaction use and orient to the visual affordance of the media in the organisation of interpreting in hospital encounters? The study contributes to the understanding of video-mediated interpreting as an interactional achievement within a specific sociomaterial setting, and to an understanding of participants’ use of visual materialities in the organisation of various activities.

2 Data and methods

The data consist of video-recordings of 11 hospital encounters with video-mediated interpreting. Three meetings were recorded from the interpreter’s location, one from the hospital facilities, and seven from both the hospital facilities and the interpreter’s location. The analysis draws on insights from the recordings from both the hospital facilities and the interpreter’s location.

The recordings include meetings in different wards, in the form of both meetings with admitted patients who may have spent several weeks at the hospital, and polyclinical appointments, consist of a brief consultation. In most cases, the medical professionals and patients have met many times before. The medical professionals involved in these meetings have varying degrees of experience using the technology and with video-mediated interpreting. All of the interpreters in the study have formal professional qualifications in interpreting. Some of the interpreters in the study are participating in video-mediated interpreting for the first time, while some have done it many times before. With their diverse experiences and varied prior knowledge of this type of situation, the participants deduce how to accomplish video-interpreting in collaboration in situ.
The interpreter participates from a remote location – an interpreting studio equipped with a desktop system for videoconferencing. The device is about the size of a personal computer and is designed for videoconferencing from an individual workspace. It has a screen, camera, loudspeakers, microphone and a control panel. The camera has a narrow angle, and can only capture a narrow area in front of the camera. The room is also equipped with a separate telephone. In the setting illustrated above, the interpreter has a personal computer in their workspace. The interpreter usually has a pen and a notepad, and can take notes during the session.
A large screen at the end of the meeting room table serves as both a screen for the computer in the room and as a video-conferencing screen. In addition to the camera above the screen, a multidirectional microphone is connected to the system by cable and can be placed on the table. The room is furnished for video-conferencing, and the camera can capture all of the participants around the table, depending on the chosen settings. Adjustments to the video-conference system and technical settings are made using a control panel on the table next to the screen. Some wards are equipped with smaller systems, like the desktop unit in the interpreter’s studio (see illustration C).
The policlínical setting illustrated above is in a multipurpose room used both for specific types of examinations and appointments that require interpreting. A computer for access to journal systems and a desktop video-conferencing unit are available on the desk. The device is designed for video-conferencing from an individual workspace and has a narrower camera angle than the more advanced system illustrated above. Since adjustments cannot be made to the camera angle, adjustments would have to be made to seating arrangements in front of the camera in order to capture all participants at the ward.

The analysis builds on video-ethnography and multimodal conversation analysis tending to the sequential organisation of interaction within this specific linguistic and material setting. The analysis builds on the theoretical framework of multimodal conversation analysis (e.g. Deppermann, 2013; Hazel, Mortensen, & Rasmussen, 2014; Mondada, 2014).

The research project is carried out with the approval of the Norwegian Centre for Research Data, and the hospitals and wards involved. All participants have given their informed consent.
3 Analysis

This analysis demonstrates and discusses how participants-in-interaction use and orient to the media’s visual materiality and the visual interactional space or ecology afforded by the technology. The analysis’ three sections demonstrate different ways in which the visual affordance of the media becomes relevant to participants in the organisation of interpreting and the interpreted encounter. The first section demonstrates how the visual ecology informs interpreters’ work. The second demonstrates how participants use embodied resources in the organisation of interpreting within this interactional space. The final section discusses how participants orient to the visual materiality of the media.

3.1 The visual ecology as a source of information

The ongoing activities at the wards are made available to the interpreter through the use of video technology. For the interpreter to be able to interpret, he/she has to be able to hear what the participants at the ward are saying. While the interpreter’s need for access to sound from the ward may seem obvious, how visual access contributes to the interpreter’s work is perhaps less so. The following section demonstrates how the interpreter’s access to the ward serves as a source of information and informs the interpreter’s actions and linguistic choices.

Extract (1) demonstrates how the visual transmission from the hospital meeting room informs the interpreter’s work. The medical professional (MP) and researcher (RES) are present in the hospital meeting room. The meeting room is equipped with a video-conference system. The camera captures almost the entire meeting room, and the video technology transmits this signal to the interpreter, who is participating from a remote location. Shortly after contact is established between the hospital meeting room and the interpreter’s studio, the medical professional begins to inform the interpreter about the order of business. The patient (PAT) and next-of-kin (NOK) arrive as she is doing so.
Figure 1: Interpreter’s view to the left. Ward’s perspective to the right.

**Extract (1)**

1. (0.6)
   
2. MP: e: ja. dere har sikkert snakket sammen.
   
   e: yes you have probably spoken together

3. MP: *så da bruker me:*
   
   so then uses e:-

4. (0.6)*(0.5)
   
   mp: --> turns to doorway-->

5. int: *nods *

6. (0.8)*(0.6)#
   
   mp: gestures to PAT/NOK-->

   pat: walks into the room toward seats -->

   int: *nods*

   fig: #fig1.1 #fig1.2
**Figure 1.1 and 1.2: Patient and next-of-kin arrive**

![Figure 1.1](image1.png)  ![Figure 1.2](image2.png)

**Figure 1.3: Participants at the ward have a seat**

```
7  INT:  bardzo proszę zapraszam
        please welcome
8   (1.1)  pat:  &takes a seat -->>
9  MP:  så da e bruker e:hm (0.3) melissa litt tid først,
       so then melissa will use a little time first
10   (1.1)  til å informere de,
       to inform them,
11 fig:  #fig 1.3
```
The medical professional is about to give the interpreter further information about the meeting as the patient and next-of-kin arrive at the doorway. The medical professional cuts herself off and turns to the newcomers (line 3). Now facing the patient and next-of-kin, she shows them to their seats. The arrival of the new participants and the medical professional’s change of physical positioning is visible to the interpreter through the video transmission from the ward (Fig 1.1). The interpreter nods as the new arrivals enter the room, displaying an orientation to a change of participation framework. So far, the interaction has been in Norwegian. This changes as the patient and next-of-kin arrive in the meeting room. While the medical professional’s physical orientation and gesture were both available to the patient and next-of-kin, and they act accordingly, the interpreter still produces a rendition of the medical professional’s utterance in Polish (line 7). By interpreting the utterance, the interpreter displays recognition of the new arrivals as Polish-speaking participants, and makes known her presence as the interpreter. The interpreter’s utterance not only serves to make the medical professional’s utterance intelligible to the patient and next-of-kin, but also establishes the interpreter’s position in the interaction.

After having shown the Polish participants to their seats, the medical professional resumes the utterance that she previously cut off (line 9–10). She initially directed this utterance to the interpreter through bodily orientation, and began producing it before the patient and next-of-kin entered the meeting room. She now continues to produce the utterance, seemingly directed to the interpreter, while she moves to her seat (fig 1.3), mentioning the patient and next-of-kin in the third person, “de” (them). The interpreter treats the utterance not as directed to herself, but to the patient and next-of-kin, as indicated by interpreting the utterance into Polish (line 15–16) and through the choice of pronouns in the rendition. While the medical professional’s utterance mentions the patient
and next-of-kin as “de” (them), in the interpreter’s rendition, the patient and next of kin are addressed using the polite form, “państwa” (you, polite, plural, includes male and female gender). The interpreter’s visual access informs both her actions (she begins interpreting as the newcomers arrive) and her linguistic choices (based on the information made available by the visual channel, the interpreter chooses pronouns relevant to the attending participants). She continues, without leaving it up to the participants at the ward to produce a next turn or select the further course of action, by producing a next turn in Norwegian. Treating the medical professional’s utterance as a possible allocation of turns to the researcher, the interpreter negotiates the order of business by requesting permission to inform the participants at the ward about her work before the meeting continues. The medical professional accepts this (line 22). The arrival of the Polish-speaking participants is made available to the interpreter through video transmission from the ward and occasions the interpreter’s conduct as interpreter. Video transmission from the hospital meeting room provides the interpreter with information that informs her choice of actions and linguistic choices in the accomplishment of interpreting.

In some cases, interpreters do not have full access to the participants at the hospital. Extracts (2A) and (2B) demonstrate how having only limited visual access to the participants at the ward shapes the interpreter’s understanding of the situation, and therefore her work. The participants’ differing linguistic access to the situation is relevant to how the interaction proceeds. Extract (2A) is from the beginning of a meeting. The patient (PAT) and next-of-kin (NOK), the patient’s wife, are seated outside of the area captured by the camera, and are therefore not visible to the interpreter. The patient and his wife have met with the doctor (DR) before. The doctor directs his gaze to the patient and formally opens the meeting by asking him how he is doing. The doctor is partially displayed on the interpreter’s screen, but the interpreter (INT) cannot see who is present in the room together with the doctor, or at whom he is gazing.
Figure 2: Interpreter’s point of view to the left. Ward’s point of view to the right.

Extract (2A)

1 (.)
2 DR: 
vordan har du det?
   how are you
3 (.)
4 DR: er- f[øler du- fø]ler du deg bra?
    ar- do you feel- do you feel good
5 PAT: [ja bra ]
    yes good
6 *(!)
   pat: *nods
7 PAT: ja
    yes
8 INT: .h Æ (.) dα, chi ra sao?
    .h uh (.) yes, how is big sister
9 Chi cảm thấy trong người khỏe hay không?
   does big sister feel healthy or not
10 NOK: không, n anh chû hông có chi dâu.#
   no, it is big brother and not big sister
  nok: »leans forward and turns to screen--->
   fig:  #fig 2.1
Figure 2.1: At the ward

![Figure 2.1](image)

Figure 2.1

11 NOK: Chái di theo: Dọg thọi.
Big sister followed. I am only the wife.
(0.7)

12 INT: .h e:: åñei det er han og ikke meg
.h e:: oh no it is him and not meg

13 †jeg er bare med jeg er bare kona.
I am only with I am only the wife
†turns to the doctor and participants at the ward-->

14 NOK: ehehehehehehe

15 DR: †jehe:g vet det
.i::h know

Fig: #fig 2.2

Figure 2.2: In the interpreter’s studio

![Figure 2.2](image)

Figure 2.2
The Norwegian word “du” (you, second person singular pronoun) does not distinguish between genders. While the participants at the hospital meeting room can see at whom the doctor is directing his utterance, the interpreter cannot. Interpreting to Vietnamese, the interpreter has to choose an appropriate pronoun for the addressee. The patient has already responded to the doctor’s utterance in Norwegian (line 5, line 7). However, he is seated to the side of the technology, and parts of his utterance overlap with the doctor’s utterance. This makes his utterance difficult to perceive remotely. Although the patient has responded, and done so in Norwegian, the participants still treat interpreting as relevant. In her rendition, the interpreter uses the pronoun chí (big sister), a polite pronoun common for addressing female participants in interaction. The patient’s wife turns to the screen representing the interpreter (fig 2.1), the screen-interpreter, and responds in Vietnamese to the interpreter’s rendition of the question. In the form of an other-repair (Schegloff, Jefferson, & Sacks, 1977) directed through gaze and bodily orientation to the screen-interpreter, the next-of-kin makes it clear that it is her husband who is the patient, not herself (line 10–11). Although this utterance gives the interpreter information about the participants at the ward, she cannot see to whom the utterance is directed. To the interpreter, the verbal utterance is disconnected from the environment, the participation framework, in which it was produced. As such, for the interpreter, the utterance is fractured from the ecology in which it was produced (Luff et al., 2003). The interpreter proceeds to interpret the utterance into Norwegian (line 12–13). As the interpreter completes the rendition, the next-of-kin begins to laugh. The linguistic problem in Vietnamese is not available to the doctor, who only speaks Norwegian. The interpreter’s utterance is designed as a rendition of what the next-of-kin said. To the doctor, the interpreter’s rendition appears to be directed to him, which in turn occasions his response – he tells the next-of-kin that he knows this and begins to laugh. While the next-of-kin’s utterance was visually directed to the screen-interpreter, she turns back to the interpreter as the interpreter renders the utterance into Norwegian. To the doctor, the content of the next-of-kin’s utterance is only
available after the interpreter has interpreted it. As such, the content of the next-of-kin’s utterance becomes disconnected from the ecology – the next-of-kin’s embodied actions and direction of utterance – in which it was produced.

The participants at the hospital know who the patient is. As both the problem and the following interactional trajectory are occasioned by the interpreter’s lack of visual access, the participants do not topicalise the interpreter’s lack of visual access. The doctor simply treats this as an utterance produced on behalf of the next-of-kin, and they proceed in the interaction. Similarly, in the study of mundane interaction, Rintel (2010) found that couples rarely address the technology that occasions the problem, but instead orient to the problem within the content.

Extract (2B) demonstrates how, even when the presence of participants at the ward has been made clear, the participation framework is locally negotiated based on the verbal interaction. Limited visual access to the participants at the ward results in the interpreter becoming more dependent on auditory cues when choosing relevant address terms. The next extract is from a later point in the same interaction as Extract (2A). The meeting is approaching its conclusion and the doctor has just suggested a time for their next appointment.

Extract (2B)

1 DR: >har [du no]en< spørsmål?
   do you have any questions
2 NOK: [ja-]
3 giống như [hỏi-]
   like now ask-
4 INT: [chị] có Đặc mê c gi không?
   do you (sister) have any questions

Again, the doctor directs his utterance to the patient using gaze, while the interpreter selects a pronoun addressing the wife. The doctor asks if the patient has any questions (line 1). Before the interpreter produces a rendition, the next-of-kin begins to produce a question in Vietnamese (line 3). The interpreter has only access to the verbal utterances, not the visual ecology that frames the activity, and makes her linguistic choices on these grounds. From the interpreter’s point of view, the doctor’s gaze is available, but not the person he is gazing at. By the time she begins to produce an interpreted rendition, the next-of-kin’s beginning utterance becomes a part of the interpreter’s source of information for establishing who is the addressee. If the interpreter were to assume that the question was most likely directed to the patient, and thereby selected the male pronoun in this specific context, she would not only repeat the doctor’s utterance, but actively allocate the turn to
someone other than the current speaker. In her rendition of the doctor’s question, the interpreter uses the female pronoun (line 4).

In video-mediated interaction, the video image acts both as a link between the participants and as a central tool for the participants’ accomplishment of activities (Mondada, 2007, p. 53). The extracts in this section have demonstrated how the interpreter’s visual access to activities at the ward informs her work. The participants at the ward did not make explicit changes in the participation framework or provide relevant information about the setting. When trouble arose, the interactional problems were dealt with without addressing the participants’ asymmetric visual access or making adjustments to the visual materiality of the setting.

3.2 Embodied resources in a visual ecology

By rendering only parts of the participants and their surroundings visible to co-participants, video-mediated interaction is asymmetric (Arminen et al., 2016; Heath & Luff, 1993). The following examples demonstrate how participants attempt to use embodied resources to organise the interpreted interaction. However, the visual ecology does not afford them the visual access to each other that they presuppose they have. Their embodied actions therefore go unnoticed and they resort to auditory resources to organise the interaction.

In extract (3), the participants have trouble coordinating the interpreter’s turns, and the doctor attempts to use gesture to solve this problem. The doctor and the patient are seated in front of a desktop unit in the hospital ward. The interpreter can only see parts of the doctor, and the patient is seated beyond the range of the camera. The participants at the ward only have visual access to the interpreter’s head. The doctor and the patient are discussing medication. After some problems coordinating turn-taking, the doctor uses gaze and gesture to allocate the turn to the interpreter.
Figure 3: The interpreter’s point of view to the left. The doctor and patient at the ward to the right.

Interpreter's perspective  Ward's perspective

Extract (3)

1 DR: e:::m  
   uh:mm  
2  
3 (0.6)  
3 er det greit for deg å fortsette med dette her?  
   is it okay for you to continue with this  
4 (0.3)  
5 PAS: mhm  
6 DR: med medi[sinen],  
   with the medicine  
7 INT: [.h ]  
8 (0.3)∩(0.5)∩(0.4)  
9 dr:  
    turns to screen  
9 dr:  
    stretches out arm to screen  
9 DR: hmn  
10  
11 PAT: hehe  
11 DR: hehe vi må bare bruке deg altså,  
   hehe we just have to use you that is  
12 ((laughter))  
13 INT: .hh(.) ((begins interpreting))
The medical professional is visually oriented to the patient and discussing medication. He asks if it is okay for her to continue “med dette her” (with this) (line 3). After a short silence, the patient produces a minimal response (line 5). The doctor produces a delayed self-repair to specify the referential expression used, “med medisinen” (with the medicine) (line 6), and as such pre-empts misunderstanding. The interpreter’s audible in-breath, a commonly used pre-beginning signal in the data, becomes audible to the participants at the ward during the doctor’s self-repair (line 6–7). The doctor completes his utterance. After some silence, he turns to the screen (line 8). After further ensuing silence, he gestures to the screen with his right arm stretched out and his palm facing up (line 8), seemingly handing something (Streeck, 2009) to the interpreter (fig 3). However, the doctor is only partially visible to the interpreter. The doctor’s arm is not extended within the area captured by the camera, and is therefore not visible to the interpreter (fig 3). The silence following the doctor’s turn and orientation of gaze itself is not enough at this point to elicit interpreting. The doctor and patient begin to laugh, seemingly at the lack of interpreting. The interaction has come to a momentary standstill. The doctor finally retracts his arm and laughingly states that “vi” (we, first person plural), referring to the patient and himself, just have to use “deg” (you, second person singular), using gaze to refer to the screen-interpreter (line 11). The interpreter then begins to interpret.

Although the participants in the interaction have not secured compatible views of each other for the interaction, the doctor uses embodied resources in the organisation of interaction. When the interpreter does not display uptake, i.e. they do not begin to interpret, the doctor proceeds to allocate the turn to the interpreter verbally. As such, although the patient has already responded to his utterance, the doctor treats the lack of interpreting as problematic and his embodied action as inefficient. To solve the problem, he explicitly allocates the turn to the interpreter by referring to the interpreter and the interpreting. Explicit referrals to the interpreter do not occur frequently in the data. The interpreter does not respond to the doctor’s utterance in Norwegian or account for the lack of interpreting or missed cue. Rather, she produces a new pre-beginning signal, an audible in-breath, and begins to interpret.

Extract (4) is from the same meeting, from the opening phase. This extract demonstrates how the interpreter uses gesture to initiate the transition to a next activity. However, the participants’ incongruent views of each other result in the gesture going unnoticed. The interpreter has asked in Norwegian if she can provide information about the interpreter’s role, and proceeds to do so in Thai.
As the interpreter’s description of her work comes to an end, the patient responds to the interpreter’s utterance in Thai (line 4). After a longer silence, the interpreter gestures toward the screen displaying the participants at the ward, seemingly signalling the completion of her utterance (fig
To the participants at the ward, the interpreter’s gestures are not visible, because the interpreter’s hands are below the camera angle (fig 4.1). To the doctor who speaks Norwegian, the vocal production of utterances in Thai and aspects of the turn-taking are available, but the content of the interpreter’s utterance and the patient’s response are not. After some silence, the doctor signals the possible closing of this activity by first asking “okay?” (line 9), before asking if that was okay (line 12). As such, he expresses uncertainty regarding the completion of the previous activity, which was carried out in Thai. The doctor’s lack of access to the interpreter’s gestures has consequences for the unfolding interaction. The doctor solves the problem of his lack of knowledge regarding ongoing activities by means of a verbally produced utterance.

The participants produce embodied actions that presuppose that they have views of each other that are compatible for the efficient uptake of such embodied cues. When the embodied actions do not receive uptake, the participants do not adjust the embodied action according to the visual ecology (as in Licoppe et al., 2017), nor do they address the incongruent views or make adjustments (similar to findings by Rintel, 2010).

3.3 Talking about the visual ecology

The previous extracts have demonstrated how the visual ecology may serve multiple purposes in the organisation of video-interpreted hospital encounters. When participants do not ensure visual access to each other that is appropriate for the activities being undertaken, this may cause problems in the interaction. While the participants solve the interactional problems that occur, they do not topicalise the asymmetric views they have of each other or the visual materiality of the setting. The opening phases of these meetings are possible spaces for making adjustments to the setting and securing an appropriate visual space. However, the professional participants rarely mention this during the opening phase of these meetings.

Extract (5) demonstrates how the doctor (DR) topicalises the interpreter’s (INT) visual access to the ward. However, this does not lead to a collaborative configuring of a visual space for interpreting. As contact is established between the ward and the interpreter’s location, the doctor is only partially visible to the interpreter, and the patient (PAT) is seated beyond the range of the camera. The interpreter’s head and shoulders are visible to the participants at the ward. The interpreter has said some words explaining her work, and the participants are ready to move on to the next topic. The interpreter is holding a pen, indicating that she is ready to interpret.
Figure 5: The interpreter’s visual access to the left. The ward’s perspective to the right.

Extract (5)

1  DR: *ser du meg eller?
   do you see me or
   dr: *leans in to cam frame, body and gaze to SCR/INT -->
2  INT: &ja: jeg ser ¤deg.¤
   ye:s i see you
   int: #looks down and to each side#
   res: &RES appears in image on INT’s scr-->
3  (0.5)
4  DR: hh ja ve- *%hehehe
   hh oka-       hehehe
   dr: *grins-->
   pat: %smiles-->
5  DR: ikke så interesse¤rt,¤
   not that interested
   int: #smiles-->
   int: #claps hands together, bows head-->
6  DR: .HH [he] he [he] .HH
7  INT: [næ] [hm]
8  (0.3)
9  DR: e:*% >du ka-< jeg kan intr¤odusere meg selv da,
   e: >you ca-< i can introduce myself then
   dr: *stops smiling -->>
   pat: %stops smiling -->>
   int: ---> #stops smiling -->>
10 DR: jeg heter Jørn Stenberg,
     my name is Jørn Stenberg
11 DR: og jeg er legen i dag,
     and I am the doctor today
The participants are about to proceed to a next activity. The interpreter, holding her pen, is seemingly ready to interpret. However, rather than opening the meeting, the doctor moves toward the screen and into the camera frame (line 1) and asks the interpreter if she can see him. This question introduces a new topic into the conversation and thus initiates a topic-proffering sequence (Schegloff, 2007, p. 171). The preferred response to a topic-proffering question is an extended response that topicalises the matter introduced. Prior to the point at which the doctor moves closer to the camera frame, the interpreter has seen only half of the doctor. Furthermore, she has no visual access to the patient. Still, the interpreter answers the question minimally, confirming that she can see the doctor, and looks away from the screen (line 2), thereby rejecting visibility as a topic to be addressed in more detail.

The interpreter did not design her response as a dis-preferred response. However, the doctor’s receipt indicates that it was indeed unexpected, unlikely or even inappropriate. With his gaze still oriented to the screen-interpreter, the doctor begins saying “ja ve-” (okay-) and laughs (line 4). “Ja vel” serves as a third position receipt, indicating slight misalignment with the previous turn, the interpreter’s response. Grinning, the doctor continues, while laughing, “ikke så interessert” (not that interested), expanding on his turn, and then continues to laugh (line 5). Still holding the pen in her right hand, the interpreter claps her hands together, coinciding with the completion of the doctor’s utterance. She then bows her head, laughing visibly more than audibly. The doctor’s laughter following his utterance begins with an equivocal initial sound (line 6), an audible, vocalised in-breath. This format, with an utterance followed by laughter, is found to initiate laughing sequences (Jefferson, 1979). Although the doctor’s remark indicates that the interpreter responded in an unexpected manner, the invitation to joint laughter facilitates a joint display of alignment before the meeting proceeds. The extract demonstrates how, in a case where the doctor topicalises the interpreter’s visual access, this does not lead to a collaborative configuration of a visual space for the interaction.

4 Discussion and conclusion

In this article, I have investigated how participants in video-interpreted medical encounters orient to and use visual ecologies in the organisation of interpreting. The analysis has demonstrated how visual ecologies in video-interpreted hospital encounters serve multiple purposes for participants in the interaction. The visual affordance may enable access to information relevant for the accomplishment of the interpreters’ work. The participants presuppose that the media affords an efficient use of embodied actions. However, the participants do not always ensure that their views of each other and each other’s surroundings are congruent with the activities and actions they are
attempting to accomplish. The interaction in these settings is quite complex, as it is not just video-mediated, but also interpreter-mediated. Due to video-mediation, participants’ utterances may be disconnected from the ecology in which the utterances are produced (Luff et al., 2003). Similarly, due to the multilingual nature of the interaction, the linguistic content may become disconnected from the embodied actions that encompass the linguistic content in the original utterance, such as gesture and gaze (see, e.g. Extract (2A)). This can cause complications for participants when making sense of participation frameworks and co-participants’ actions. While the participants’ lack of or incongruent visual access to each other may cause problems in the interaction, they do not attribute the interactional problems to the insufficient visual ecology or make adjustments to the setting. They simply solve the immediate interactional problem and proceed.

Video recordings made by participants in video-mediated interaction, and what they capture and share with remote co-participants, make possible a space of collective action (Mondada, 2007, p. 52). The visual image may be essential for the accomplishment of some professional activities, e.g. during the physical examination phases of telemedicine consultations (Pappas & Seale, 2010) or in surgical settings where laparoscopic surgery, remote expert online advice and remote learning are made possible via the video image (Mondada, 2007). What is a relevant visual ecology for the collaborative accomplishment of interpreting, and how this visual materiality does in fact inform interpreting and the interpreter’s work, might not be entirely clear to the participants in the interaction. While problems hearing might easily be associated with problems in accomplishing interpreting, the participants do not readily connect interactional troubles to insufficient visual access for the accomplishment of ongoing activities.

According to Schutz, the reciprocity of perspectives, in addition to building on the idealisation of the interchangeability of standpoints, builds on the idealisation of the congruency of the system of relevances (Schutz, 1953). While a person’s biographical situation determines at any given moment her purpose at hand, the idealisation of the congruency of the system of relevances suggests that the differences in perspectives originating in people’s unique biographical situations are irrelevant for the purpose at hand (Schutz, 1953). The participants in interaction seem not only to presuppose mutual visibility, but to take for granted that they share an understanding of the purpose at hand. The complexity of the setting, and participants’ various degrees of access to a visual ecology, to linguistic content (and knowledge about the differences of linguistic systems), to background knowledge and even to the purpose at hand at a specific moment in a specific encounter, may in some ways challenge the participants in the accomplishment of that interaction. Engaging in meta-talk about the accomplishment of interpreting in a mediated space could be a possible way to reduce some of the asymmetries in the interaction.
Except for interpreters’ requests to explain their work during the opening phase of the interaction (see extracts (1) and (5)), the professional participants tend not to engage in meta-talk about interpreting or the technology involved. In Extract (5), the medical professional provides an opportunity to talk about the interpreter’s visual access. However, the participants do not engage in the collaborative configuring of a visual space. The way in which the interpreters introduce their work is found in textbooks on interpreting (e.g. Skaaden, 2013) and is practised during interpreting studies. Engaging in meta-talk with a participant in one language poses a risk that the meta-talk may be perceived by speakers of the other language as side-talk, possibly excluding the speakers of the other language (Meeuwesen, Twilt, ten Thije, & Harmsen, 2010). Video-mediated interpreting is still not very common in Norway, and the participants engaging in the interaction have varied experience with accomplishing such meetings. Procedures similar to the introduction of the interpreter’s work have not yet been developed for video-mediated interpreting, although the increased use of technology for the provision of interpreting during the COVID-19 pandemic has occasioned discussions on current practices. Over time, the use of technology has been found to cause an evolution of behaviours, e.g. regarding communicative practices (Dourish et al., 1996). Similar evolutions are likely to occur within the organisation of video-mediated interpreting.

This study has provided insights into the organisation of video-mediated interpreting in hospital encounters and how participants in these settings use and orient to a visual ecology in the organisation of interpreting. As such, the study contributes to the body of knowledge describing various professional activities in mediated environments, and specifically to the understanding of interpreted interaction within a mediated environment. Using multimodal conversation analysis in the investigation of authentic video-interpreted hospital encounters provides insights into the complexity of these interactional settings and the variety of semiotic resources on which the participants draw in their interaction.

Acknowledgements

I am grateful to the reviewers for their comments and suggestions and to Professor Jan Svennevig for his input and advice during work with this article. I am grateful also to Jenny Gudmundsen for useful discussions during work with the article.

This work was partly supported by the Research Council of Norway through its Centres of Excellence funding scheme, project number 223265.
Bibliography


ARTICLE 4: Remediating the mediator: Media ideologies in policies and practices of medical interpreting

Authors: Jessica P B Hansen

Prepared for submission
Remediating the mediator: Media ideologies in policies and practices of medical interpreting

Abstract

Based on the analysis of government documents and interviews with medical professionals and interpreters, this article addresses media ideologies (Gershon, 2010a, 2010b, 2017) drawn upon by stakeholders in debate about video-mediated interpreting in society and in medical encounters. Employing discourse analysis in the analysis of government documents and interviews with practitioners, the article explores how stakeholders compare screen interpreting to onsite interpreting and telephone interpreting. The article finds differences and similarities between remediation in government documents and in practitioners’ talk about their experiences, among other grounded in the ways in which stakeholders deal with the media as being a means of service or a tool for conducting interaction. Remediation is of institutional relevance in both government documents and in interviews where stakeholders express concerns about making interpreting services more efficient and problems in conducting interaction.

Key words: media ideologies, remediation, interaction, video-mediated interpreting,

1 Introduction

Interpreting is generally considered to be a measure to overcome language barriers in language discordant meetings between public services and individuals. Interpreting was made relevant in Norway with increased working migration in the 1960s and 1970s, and the field of interpreting has developed alongside fluctuations of migration flow since then (NOU 2014:8, 2014, p. 25). While the telephone has been used for the provision of interpreting in Norway since the 1970s (NOU 2014:8, 2014, p. 26), early attempts to use video-technology for organizing interpreting services began in the late 1990s (Skaaden, 2001). Recent years’ rapid technological development and increased availability recurrently make relevant discussions on expanded use of video-technology for the provision of interpreting. Although Norway is a country with relatively well developed technological infrastructures and good access to technology across the working population
(Rønning, Sølvberg, & Tønseth, 2005), video-mediated interpreting, or *screen interpreting* (skjermtolking) as it is commonly called in the Norwegian setting, has not yet been widely adopted by Norwegian public services.

Based on the analysis of government documents and interviews with medical professionals and interpreters, this article addresses media ideologies (Gershon, 2010a, 2010b, 2017) drawn upon by stakeholders in debate about video-mediated interpreting in medical encounters, specifically focusing on *remediation*. Employing discourse analysis in the analysis of government documents and interviews with practitioners, this article explores how stakeholders compare screen interpreting to two other channels of communication, onsite interpreting and telephone interpreting. The article finds differences and similarities between remediation in government documents and in practitioners’ talk about their experiences, among other grounded in the ways stakeholders deal with the media as being a means of service or being a tool for conducting interaction. Remediation is of institutional relevance in both government documents and in interviews where stakeholders express concerns about making interpreting services more efficient and about problems in conducting interaction, for instance due to materialities of the media.

The analysis is structured in three sections. The first section outlines a media ecology in which video-mediated interpreting is positioned. The second takes a critical stance to the construction of arguments in government documents. The final section explores how practitioners compare video-mediated interpreting to other channels when talking about their practice.

### 1.1 Media ideologies

Media ideologies involve assumptions that people hold about how the choice of a medium over another impacts the accomplishment of communication (Gershon, 2010a). Media ideologies are a part of broader focus on semiotic ideologies (Gershon, 2010b; Keane, 2003) and pay attention to the semiotic ideologies of voice, body, image and sound (Gershon, 2010b, p. 284). Gershon (2010b) construes media ideologies across several themes, for instance remediation, referentiality, address, materialities and newness. Dimensions of media ideologies taken up in this analysis are remediation and, at a second level, materialities. Remediation addresses how there exists a comparison, tacit or explicit, between a medium and all others available within a media ecology (Gershon, 2017, p. 20). When talking about a recently introduced channel of communication, people put media ideologies and media practices associated with this channel in dialogue with the ideologies and practices of other older channels (Gershon, 2017).
The comparative approach to video-mediated interpreting is not just common in public discourse; even strands of research compare video-mediated interpreting to onsite interpreting and telephone interpreting. For instance, studies have set out to compare the quality of the interpreter’s performance in video-mediated settings compared to various other settings (Braun & Taylor, 2012; Miller-Casino & Rybinska, 2012) and whether the process (technology) influences the product (interpreting) (Balogh & Hertog, 2012). Another dimension, which has been explored using different sources of data, is satisfaction with the different ‘modes of communication’ in medical encounters. For instance, Locatis et al. (2010) in their comparison of video-mediated interpreting, onsite interpreting and telephone interpreting asked whether participants in the study found the communication method distracting. The participants preferred onsite interpreting and video compared to telephone. Similarly, Price, Pérez-Stable, Nickleach, López, and Karliner (2012) investigate interpreters’ perspectives on video-mediated interpreting compared to onsite and telephone interpreting, finding that telephone was the lesser option, onsite the best, and that video-mediated interpreting “represented an improvement” to telephone interpreting (p. 226). By engaging in questions about what benefits one channel provides over another, research may engage in remediation ideologies.

2 Materials and methods

The material is gathered through ethnographic processes and includes field observations, documents, interviews and video-recordings of video-mediated interpreting in hospital encounters (Hansen, forthcoming). The table below gives an overview of the documents in the analysis. The Official Norwegian Report (ONR) “Interpreting in the public sector – a question relating to the right to a due process of law and equal treatment” (NOU 2014:8, 2014) is the core document in the analysis, and the analysis is structured around arguments form the ONR. The ONR is intertextually linked to other texts (Reisigl & Wodak, 2016). Of the many texts intertextually connected to the ONR, I have included only a few: the hearing note to a draft law (HN) and selected documents from the healthcare sector.
The included documents have different areas of authority and fields of action. The ONR and the HN are both instruments for governing. The guidelines for medical professionals address a narrower field and direct recommendations to healthcare professionals and managers in healthcare services. The project reports report from projects with various aims and scopes and are produced by various authors (myself included, which I return to below). The analysis is structured around the ONR, demonstrating how the ONR constructs knowledge through among other the intertextual relations with the other texts.

Interviews with practitioners affiliated with three larger hospitals in the urban parts of Southern Eastern Norway were carried out in addition to interviews with other stakeholders. The urban Oslo area is an area with a higher population density, and a higher density of qualified interpreters available than other parts of the country (NOU 2014:8, 2014). The interpreters included in the study have formal qualifications.
<table>
<thead>
<tr>
<th>Interpreters</th>
<th>Medical professionals</th>
<th>Other stakeholders</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 interviews</td>
<td>11 interviews</td>
<td>4 interviews</td>
</tr>
<tr>
<td>20 participants</td>
<td>14 participants</td>
<td>4 participants</td>
</tr>
</tbody>
</table>

With a complex dataset containing ethnographic data, interviews, texts and observations, the discourse analysis combines historic, although short term, perspectives (Reisigl & Wodak, 2016) and text analysis (Fairclough, 2003, 2016). The concept of remediation is operationalized as an analytical concept in the aim to identify processes of remediation in these different data sources.

While discussions about video-mediated interpreting may occur parallel within the field of signed language interpreting and the field of spoken language interpreting and the same technologies might be up for discussion, for stakeholders within the communities, the stakeholders treat the two fields and thus their media ecologies separately. The field of interpreting I investigate here is defined and delimited among other within the ONR, and is the one of spoken language interpreting.

I am a trained interpreter and have worked as an advisor in the field of interpreting. I was co-author of the Oslo report mentioned above. I started making the interviews for this project before the analytical framework was decided on. Both in my prior work and through the questions I asked in interviews, I have engaged in remediation although unknowing of the concept. My questions in the interviews were constructed to elicit narratives, and as such, I have contributed to constructing the data (Gubrium & Holstein, 1998).

3 A field of interpreting as a media ecology

Analysing media ideologies connected to one media, makes relevant how the specific medium is positioned within a media ecology (Gershon, 2017, p. 21). ‘Media ecology’ as a term is used in a number of ways (Fuller, 2005), and can, among other, be understood as an environment of practice (Madianou & Miller, 2012). As a background for the analysis of remediation of video-mediated interpreting in a Norwegian setting, this following section outlines a media ecology in which video-mediated interpreting is positioned.

Interpreting in Norwegian public sector is considered to be a means to secure a linguistic diverse population equal access to public services and is based on values of the welfare state. Djuve, Lunde, and Sandbæk (2011) suggest that language or degree of Norwegian proficiency “makes the most
visible barrier for immigrants in their encounters with the public” (Djuve et al., 2011, p. 24, my translation). The Official Norwegian Report (ONR) “Interpreting in the public sector – a question relating to the right to a due process of law and equal treatment” 7 was the first to address solely matters of interpreting. Merely by becoming the topic of an ONR, interpreting or topics related to interpreting in Norwegian society were defined as a societal problem or an area in need of measures. To address the dilemma of public authorities spending substantial resources on interpreting services while the resource use was not found proportionate to the quality of the services (NOU 2014:8, 2014, p. 13), the Ministry of Children and Families appointed a committee to investigate and report on the situation. The initial suppositions were that Norwegian public services underused interpreting, qualification requirements were lacking and public services had poor procedures for booking interpreters (NOU 2014:8, 2014; Summary, 2014). Compiling information from dialogues with actors and public services with findings from research, the committee found this to be an accurate picture. Two of several suggested measures to solve the problems of interpreting in the public sector were (i) to develop new legislation regulating use of interpreting in public services, among other requiring public services to use interpreters with formal qualifications and not bilingual individuals without interpreter training, and (ii) to increase the use of screen interpreting.

From a political standpoint, interpreting is considered to be a measure to solve the problem of language discordance and thus to be a measure to secure equal access to public services.

Norway is a country with large distances and a widely-dispersed population. Needless to say, the need for qualified interpreters in more than 100 languages cannot be satisfied locally. This fact notwithstanding, all inhabitants much (sic) be offered equal public services, regardless of where they live. The Committee believes that increased use of remote interpreting (telephone and screen interpretation) is necessary to safeguard the right of due process and to provide equal services to the entire population.

(NOU 2014:8, 2014, p. 17; Summary, 2014, p. 6)8

Remote interpreting is implicitly opposed to onsite interpreting. Basic reasoning by government bodies that propose the increased use of video-technology to provide interpreting, builds on how video-technology, like the telephone, affords interaction at a distance. The official report suggests that screen interpreting should be the “preferred method in suitable situations” and that half of all interpreting assignments in the public sector should be carried out through screen interpreting by

---

7 The ONR explicitly excludes sign language interpreting from the mandate.
8 Where extracts are available in English, I use the English version. However, I refer to the placement in the Norwegian main document and the English version. In cases where I have translated the extract, this is expressed in the citation.
2023 (NOU 2014:8, 2014, p. 16; Summary, 2014, p. 6). While video-mediated interpreting existed prior to the ONR, the ONR explicitly articulates political aims related to the use of technology.

The ONR and following government documents articulate politics of video-mediated interpreting. In her study investigating signed language video-interpreting services, Haualand (2011) points out that video interpreting services are political technologies, both serving to perform politics and as objects of politics. The video interpreting services Haualand (2011) investigated already existed, comprised of all the elements through which the service came into existence: “the telecommunication infrastructure, which connects the involved actors to each other, and the research, invention, and development processes that resulted in the technical equipment involved, and the continual development of new and enhanced functions”. The screen interpreting organization implicitly suggested by the ONR does not exist. However, the ONR occasions projects attempting to try out or explore screen interpreting through different organizations. For instance, Oslo University Hospital (OUH) carried out a project to explore technological, financial and communicative aspects of screen interpreting (Hansen & Løfsnes, 2016). Haukeland University Hospital carried out a project to test screen interpreting as an attempt to increase quality of interpreting services compared to telephone interpreting (Haukeland universitetssjukehus, 2017). Helse Førde engaged in an innovation project with a technology developer to develop a technological organization to increase access to interpreters and to increase the use of screen interpreting compared to telephone interpreting (Helse Førde, 2018). Whereas OUH already used video-remote interpreting in addition to onsite interpreting and telephone interpreting, the two hospitals situated on the west coast of Norway aimed to test video-mediated interpreting in an environment where telephone interpreting was more common. The ONR legitimized projects, and projects were later referred to in the hearing note to the draft law (HN). This in turn reinforced governmental incentives to increase the use of screen interpreting in society.

While the ONR and HN are cross-sectoral documents, the practitioners I have interviewed are affiliated to hospitals. The practitioners in the interviews often suggest that onsite interpreting is an ideal and onsite interpreting is often the baseline for comparison. Screen interpreting is “almost like being there”, some interpreters suggest, while medical professionals might suggest that it is “almost” like having the interpreter present. However, to practitioners in other parts of the country, telephone interpreting might be more common and therefore the baseline of comparison, such as might be the case with the projects in Haukeland and Førde. Some actors in public services may consider telephone interpreting the only way to get access interpreting services. The interpreters I interviewed take assignments for hospitals and for other institutions. They draw on experiences from different institutions in different sectors, such as judicial settings and asylum interviews, in
addition to their experiences from healthcare, when they talk about interpreting. The medical professionals have experiences from working with qualified interpreters, but may also draw on experiences from situations where bilingual persons without interpreter training conducted the interpreting, without necessarily distinguishing between qualified and unqualified labor.

In government documents dealing with interpreting in Norwegian society and in interviews with medical professionals and interpreters, onsite interpreting, screen interpreting and telephone interpreting, comprise three related but clearly distinct channels of communication. Societal definitions of these channels do not simply reflect which technological artefacts are used to conduct the interaction but are tightly linked to service provision. Interpreters may engage in telephone interpreting or video-mediated interpreting in settings that are not defined as screen interpreting or telephone interpreting. For instance, an interpreter might travel to a hospital ward to find that there are several remote speakers of the minority language participating via several different channels, both video and telephone, in the same meeting. The interpreter who has to adapt her conduct to the affordances of the various channels in the setting, might consider this screen interpreting. The medical professional who has set up the meeting might also consider it screen interpreting. However, to get an interpreter to appear at the ward at the given time, they book onsite interpreting. The service provider’s categorization is based on their technological infrastructure and does not reflect the actual event at the hospital or what possibilities and constraints the technological setup may pose to the accomplishment of interpreting or the hospital encounter. It does not refer to what skills this specific event may require of the participants carrying out the interaction. The service provider will book, bill, count, and report this as onsite interpreting.

Interpreting in Norwegian society is a measure and a service that ensures equal public services to a diverse population and simultaneously situated interactional practices carried out in the meeting between individuals and public services. Within the media ecology such as is outlined here, screen interpreting is politicized and frequently compared to onsite interpreting and telephone interpreting. Screen interpreting refers to a way of providing interpreting services and to the accomplishment of interpreting within specific socio-material settings.
4 Remediation in policy documents

Remediation is a process in which people’s understandings and experiences of one medium are inter-twined with those of other media (Bolter & Grusin, 1999; Gershon, 2010a). Screen interpreting is frequently compared to onsite interpreting and telephone interpreting along different dimensions. While the previous section outlined a media ecology as a point of departure, the following section provides examples of remediation in government documents. First, I provide a brief overview of the ONR and demonstrate how it is intertextually connected to other documents. Thereafter I outline some of the main propositions in the ONR and hearing note. Finally, I provide in-depth analyses of some of the propositions in the document and discuss the consequences of these.

The ONR comprises knowledge from various sources. Through more or less explicit intertextual connections between the ONR and other documents, such as project reports, guidelines and the Hearing Note to the draft law, statements from the document gain legitimacy and contribute to construction of knowledge. Remediation of screen interpreting within the intertextual web builds on many of the same claims and assumptions, and the intertextual web reinforces and sediments these claims and assumptions as knowledge.

The figure demonstrates how the Official Norwegian Report (ONR) is connected to documents preceding it and following it, such as the Hearing Note to the draft law (HN). Some of the documents preceding the ONR are guidelines from the Norwegian Directorate of Health (Helsedirektoratet, 2011) and a project in Helse Finnmark (Furskognes et al., 2013). Similarly, proposals from a Danish National Strategy, which in turn is connected to project reports from Danish healthcare, are included and recontextualized in the ONR. The ONR motivates and as such occasions later projects such as the three previously mentioned projects in hospitals (OUH,
Together with the ONR that occasions the HN, the three projects contribute to occasion the argument for increased use of video-mediated interpreting in the Hearing Note to the draft law (2019).

In the ONR, video-mediated interpreting is topicalized in three different sections. First screen interpreting is mentioned in the summary. The summary is a piece of text that rests both on the main body of text and at the same time functions as a separate product as it has been translated into English and Sámi. As such, propositions in the summary must stand alone while at the same time presumably rest upon the knowledge compiled in the rest of the document. Furthermore, screen interpreting is mentioned in the main body of text, specifically sub-sections compiling findings from research and reports organized by sector and country. The statements in this section are explicitly connected to their origins through each sections’ title and use of references. As such, they often have clear intertextual connections. In other parts of the text, these connections might be more implicit. Finally, screen interpreting is addressed in a section dedicated specifically to remote interpreting. This section juxtaposes experiences and findings from research, and defines and discusses screen interpreting.

In both the ONR and HN, screen interpreting is frequently compared to onsite interpreting and telephone interpreting. In the process of remediation, the documents switch between referring to screen interpreting as service provision and as interaction. Referring to service, the ONR and hearing to the draft law compare the channels of communication, finding that remote interpreting is a better option than onsite interpreting building on arguments based on for instance economy and access. For instance, the hearing note suggests that increased use of remote interpreting has benefits compared to onsite interpreting: “increased use of remote interpreting gives several positive effects such as reduced travel costs and better exploitation of interpreters’ capacity” (Det Kongelige Kunnskapsdepartement, 2019, p. 57, my translation). Referring to the interaction, HN suggests that “the dynamics in the interaction may be different” (p. 57, my translation) in interaction when interpreting is provided through video compared to onsite interpreting, and that screen interpreting may cause “wobbly turn taking” (vaklende turtaking) (p. 58). Based on the logic of efficiency and access, remote interpreting is considered to be better than onsite interpreting. Of the remote channels, screen is better than telephone interpreting based on the media’s affordances. Screen interpreting, according to the documents, provides visual access between the interpreter and the other participants which is supposed to promote interpreting and security (Det Kongelige Kunnskapsdepartement, 2019, p. 58). However, remote interpreting may have consequences for the interaction such as challenges regarding turn taking. As such, the documents construct an argument proposing increased use of technology.
The positive aspects of technology are not limited to costs. The three following examples demonstrate how the ONR attributes various positive traits to screen interpreting when comparing screen interpreting to onsite interpreting and telephone interpreting. In the process of remediation, video-mediated interpreting is suggested to improve the setting in different ways and is alternately compared to onsite interpreting and screen interpreting.

Example 1: Professionalism

The first example takes a critical perspective to the construction of knowledge in the ONR and demonstrates how screen interpreting is attributed values in the process of constructing knowledge. The first extract is from the summary to the ONR.

Various pilot projects show that screen interpretation may contribute to professionalization, rationalisation and cost efficiency in the public sector.

(NOU 2014:8, 2014, p. 16; Summary, 2014, p. 6)

Rationalization and cost efficiency are attributes likely to be associated with the organization of interpreting services seeing that the technology affords interaction at a distance. The first proposition in this example, that screen interpreting can contribute to professionalization, is not readily available to the reader. The proposition as presented in the summary does not make clear how mediation can contribute to professionalization or whether this refers to screen interpreting as a way of organizing interpreting services or to screen interpreting as video-interpreted interaction. While this claim is unsubstantiated in the summary, within the main body of text, a section addressing experiences from Denmark gives more information about the source.

According to The National Strategy of Digitalization of the Healthcare System 2013-2017 [A Danish strategic document], use of screen interpreting leads to the interpreting situation being experienced as more professional, among other due to the distance to the interpreter. This contributes to strengthening trust between patient and doctor and thereby creates the foundation for open conversations, more precise diagnoses, and medication and better treatment for patients who might be especially insecure in meetings with the health care system. A consequence of the screen interpreting project is that the number of absences is halved.

(NOU 2014:8, 2014, p 92, my translation)

The ONR refers to a Danish national strategy of digitalization of healthcare services. The paragraph suggests that interpreting situations “are experienced” (oppleves) as more professional due to the
distance to the interpreter. While this excerpt is placed within a section of the ONR describing the situation in different countries, the section is taken from a Danish strategic document (Regeringen, 2013) called “Digitalization with effect”. This is a pamphlet explicitly arguing for increased digitalization. In the preface it suggests that “digitalization is a key tool for creating a more citizen-friendly, coherent and efficient healthcare system” (my translation). While the Danish strategy serves to promote digitalization in the Danish healthcare system, the argument from the Danish document is re-contextualized in the Norwegian report as findings or observations. Whereas the Danish original suggests the video-interpreting project “shows that”, the ONR claims the project “leads to” the interpreting situation being experienced as more professional. This specific paragraph and the causal descriptive construction of the argument is reiterated nearly word by word as it was stated Danish document (Regeringen, 2013, p. 8).

The line of argument proposed in the ONR attributes properties to video-technology as something that contributes positively to the situation. While the first mention of professionalism was unclear as to whether it referred to the service or to the specific situation, this paragraph connects “a professional experience” to “distance” in a way that suggests that distance in the specific setting in fact affords professionalism.

The next referral to the Danish findings is found in the section of the ONR addressing remote interpreting. The idea of the link between video-mediation and professionalism no longer refers only to the Danish document, but suggests that this concurs with Norwegian experiences.

The committee observes with interest that the use of screen interpreting in Denmark leads to the interpreting situation being experienced as more professional, due to distance to the interpreter. This corresponds with experiences from the Norwegian screen interpreting project from 2001 and the Norwegian Labour and Welfare Administrations project of 2014. That the Danish project in addition resulted in halving of the number of absences at the wards, shows how screen interpreting can contribute to both professionalization, rationalisation and cost efficiency.

(NOU 2014:8, 2014, p. 152, my translation)

In the earlier mentions of the two Norwegian projects within the ONR, video-mediated interpreting is not explicitly linked to professionalism. This paragraph makes the connection, however, and the initial statement undergone analysis, the statement from summary, gains increased legitimacy through referring not only to one observation from Denmark but “projects” in plural. The Norwegian Official Report transforms the proposals from the Danish strategic document from observations being referred to, to becoming an object of interest to the committee, and finally, in
the summary, becoming a decontextualized proposition standing on its own. As such, the official report not only compiles knowledge but also transforms strategic arguments into knowledge, which in turn contributes to the report’s political argument. Media ideologies become a part of the knowledge base substantiating political decisions.

Example 2: Neutrality

Professionalism is not the only positive trait attributed to the technology. The following example demonstrates how remote interpreting is suggested to promote professional neutrality. Again referring to service provision, technology that facilitates interaction at a distance is suggested to allow public institutions to gain access to a larger pool of qualified interpreters. The technology affords a different organization of services, which is suggested not just to be a matter of gaining access to qualified interpreters but also a matter of practicing anonymity and confidentiality.

Remote interpreting will also contribute to increased access to interpreters who are neutral, in that they do not have association to the same local area as the minority language speaker.

(NOU 2014:8, p. 147, my translation)

This proposal suggests that gaining access to interpreters from a larger geographical area secures access to “neutral” interpreters. Neutrality here is a matter of practicing impartiality. As such, ways of organizing interpreters, which are afforded by distance, are considered to secure interpreters’ impartiality. Ordinarily, matters of professionals’ impartiality are considerations made by professionals’ themselves. Professionals’ consideration of impartiality is statutory through for instance the Public Administration Act (Forvaltningsloven, 1967). Furthermore, ethical guidelines for interpreters are central in interpreter training and clearly state interpreters must consider whether they are impartial based on certain criteria (Integrerings- og mangfoldsdirektoratet, 2020). The proposal suggests that rather than the practitioners being responsible for their professional conduct, issues of neutrality can be solved by use of technology.

Example 3: The visual affordance

The following example demonstrates how the visual affordance of the media is proposed to increase security and promote the interaction. Arguments for the use of video-technology highlights the
visual affordance of the media not only as an advantage for the interaction, but also for seeing the interpreter’s whereabouts.

Screen interpreting makes it possible for users of interpreting to see the interpreter and the room the interpreter is present in and vice versa. Screen interpreting makes it possible for the interpreter to read the implicated parties’ facial expressions and body language, which is important for the quality of interpreting


The visual affordance of the media is considered to give participants visual access to each other, an affordance clearly distinguishing screen interpreting from telephone interpreting. The richness of the media often forms the basis of arguments proposing that screen interpreting is better than telephone interpreting. The quote above contains two arguments. The first sentence in the paragraph above addresses a different advantage of having visual access to the interpreter - having visual access to the whereabouts of the interpreter. This statement is not a stand-alone idea found only in this document; guidelines produced by the Norwegian Directorate of Health (Helsedirektoratet, 2011) include a nearly identical version of the statement. The sentence is identical, except that the Official Norwegian Report has added a final “and vice versa”, somewhat softening and perhaps even disguising the main implications of the argument. The argument is implicitly structured around the comparison to telephone interpreting. Telephone interpreting and screen interpreting both afford interpreting at a distance. While one of the advantages of video-mediated interpreting is suggested to be that the participants gain visual access to each other compared to telephone interpreting, an argument against telephone interpreting is participants’ lack of visual access to each other and the participants lack of control of the interpreter’s whereabouts and if anyone is eavesdropping (Helsedirektoratet, 2011; NOU 2014:8, 2014). This same claim is found in earlier public documents and reiterated in later documents such as project reports (e.g. Haukeland universitetssjukehus, 2017) and even the hearing for the draft law (Det Kongelige Kunnskapsdepartement, 2019). By being reiterated in a number of documents, the statements become a part of the common knowledge about screen interpreting.

The final sentence constructs an enthymeme not based on facts but on media ideologies. The first premise of the enthymeme is that screen interpreting gives the interpreter access to “implicated parties’ facial expressions and body language”. The second premise is this visual access promotes the quality of interpreting. In a media ecology where screen interpreting is continuously compared to telephone interpreting and onsite interpreting, the conclusion to be drawn here is that screen interpreting is better than telephone interpreting.
Since the ONR was published, the draft law was produced and circulated in a hearing round (Det Kongelige Kunnskapsdepartement, 2019). The hearing note builds on the knowledge accumulated in the ONR and furthers politics of screen interpreting. Other bureaucratic initiatives have also been made since then to increase the use of video-mediated interpreting in society. Whereas the ONR suggested that half of all interpreting assignments should be video-interpreted by 2023, recent estimates indicate that only 0.5% of the assignments in Norwegian public services are currently carried out as screen interpreting (Det Kongelige Kunnskapsdepartement, 2019). The 2020 state budget substantiates the aim to increase the use of screen interpreting in Norway by awarding 10 million NOK to the Directorate of Immigration and Diversity for a national commitment to screen interpreting (Kunnskapsdepartement, 2019-2020). The ONR occasioned work with developing video-interpreting services, while the state budget materializes the incentive to increase use of video-mediated interpreting in Norwegian society.

The knowledge base for promoting political decisions is not merely compiled in the public documents, the documents reinforce claims and construct knowledge through intertextual relations to documents preceding and following the document in question. As such, each of the documents are part of an intertextual web constructing knowledge while at the same time serving as products on their own.

5 Remediation in practitioners’ experiences

Government documents suggest that screen interpreting not only contributes to efficient service provision and increased access but even increases professionalism, neutrality and security through technology’s attributes. However, interviews with medical professionals and interpreters reveal how screen interpreting is also a matter of accomplishing interaction and consequentially accomplishing the medical appointment. The questions in the interview evolve around the participants’ experiences with screen interpreting. For instance, I asked them to tell me about a situation that worked poorly with screen interpreting and one that worked well. The following analytical section draws on the experiences of medical professionals and interpreters and demonstrates how they compare the channels of communication to each other first with consideration to the accomplishment of medical meetings and thereafter with regard to the accomplishment of interpreting.
5.1 Accomplishing medical meetings

The following section demonstrates how comparison between channels of communication may be embedded within narratives about practice and therefore also sheds light on medical practice. I asked Anna, a doctor, about screen interpreting, how it worked.

It worked- what was- it works okay in the first place, it was almost like having the interpreter in the room. Except that the room maybe wasn’t adapted to how one sort of- that it is a little bit of a hassle getting placed so the interpreter sees both of us and that both of us see the interpreter, and one sort of manages to- and that the interpreter gets to see the communicative that happens between. Not just what we say, sort of, but gets all of it and- I found that a bit challenging (laughs) when it came to the screen settings and the distance and turning it into a proper triangle. Which was something that worked. But if it is the equipment or the room or what- both.

Extract 1

Anna compared screen interpreting to having the interpreter present in the room, before she continued to talk about the materialities of the media, describing the setting as difficult to arrange. She had difficulties organizing an interactional space, a space where the interpreter could see what was necessary in order to interpret, and where she, the patient and the screen representing the interpreter were positioned properly in relation to each other. Different wards have different technical setups. Some of them use videoconference units intended for use form a personal workplace, while others use videoconference systems. The issues Anna experienced with the placement of participants and technology reduced her sense of having the interpreter present, the sense of immediacy. Her experience with screen interpreting was connected to the work she was carrying out and the materialities of the media within the specific setting.

To a doctor, using screen interpreting instead of onsite interpreting might be a matter of interacting with and treating a patient in a satisfactory way. Asked about a situation that did not work well, Anna told me about a situation where the transition from having the interpreter present to having screen interpreting did not work out as planned.

Yes, um we quit and went back to onsite interpreting again. In a way [I] didn’t- [I] got the impression that the patient directed the utterances to the interpreter and didn’t get- that one sort of didn’t achieve good contact, in a way. That the contact was more limited. But this is also because- [I] became unsure of the patient’s cognitive level- because there were a lot of misunderstandings and I didn’t get the responses I was expecting sort of. Relating to what- the questions one wa- yes, it didn’t flow as well. The communication didn’t flow as well and the patient didn’t hear- that is, didn’t hear as well. I got the impression, with screen interpreting, that he had bigger problems hearing than I had- but we should have had onsite interpreting. We started with onsite interpreting and then went over to screen interpreting. And then we went back to onsite interpreting. I got the impression that the contact with the patient became worse in a way. At the same time as I became unsure of whether one heard okay. So it [screen interpreting] kind of resulted in the interpreter and me having a different understanding of what we were talking about in a way than the patient ((laughs)).

Extract 2
To Anna, exchanging one channel of communication, onsite interpreting, with another, screen interpreting, did not work well. She became unsure of the patient’s cognitive levels and hearing. Furthermore, she felt that she lost contact with the patient and that the patient seemed to direct their utterances more to the interpreter than to her. She experienced that screen interpreting caused a change of participation frameworks (Goodwin, 2007), how the participants’ physical orientation to each other in interaction frames the interaction, in a way that made her unsure of the patient’s condition. Comparing screen interpreting to onsite interpreting, she found that the interaction did not flow as well and that her contact with the patient was reduced. Finally, they switched back to onsite interpreting. Screen interpreting, to the doctor, was not just a matter of transferring a message through a different channel; the use of screen interpreting was relevant to the accomplishment of the medical meeting.

To the medical professionals Beata and Charlotte, how well screen interpreting works is also a matter of what interactional activities they are carrying out.

The more it is like information, like pure information, the better it works. The more it is sensitive conversational topics that are going to be talked about, the worse it works with screen interpreting. You don’t always know this in advance. Because it is not necessarily stated clearly in the referral. This is stuff that just comes up.

Extract 3

To Beata and Charlotte, the channel of communication is not neutral. Whether screen interpreting works satisfactory as a channel of communication in their meetings with patients, depends on what they are doing in their meeting. They do not always know what will come up in an encounter in advance.

Talking about a situation that worked well, Diana suggests that explaining medicinal adjustments can be difficult using screen interpreting. She describes how they in a meeting drew on a number of resources in the interaction.

DIANA: It can be difficult to explain increase and decrease of medication dosages, for instance, through a screen interpreter, in a way. But I had a setting where the parents also understood some Norwegian. So it was a little in Norwegian and a little through the screen interpreter and some drawing and explanations and stuff like that. That time at least we had a very good experience of the use of the interpreter. Because [the interpreter] understood, sort of, everything we wanted to convey with all the medicinal adjustments and we experienced that when [the family] came back next time, they had done what ((laughter)) we had agreed on during this conversation. So one got a confirmation that, um, the experience we had, that they understood, that it was right at least.

JESSICA: increasing and decreasing medication is complex to explain probably through interpreting, is there something that makes it extra complex through technology, do you think?
DIANA: Um, maybe more because one can, when one is the same room it is easier to use body language, sort of. That is, you can demonstrate more. Often when we have screen interpreting we sit in a room far away from the ward, so it is not so easy to, for instance, to go and get the pills and demonstrate, in a way, when we are away from the ward. When we are close to the ward then we do that a bit more.

Extract 4

To Diana, a measure of whether the screen interpreting works well is whether the interpreter understands what they are doing and the patient and family understand and act accordingly. She described a situation where they used a number of resources to explain what they wanted the family to do. She mentions initially that medicinal adjustments can be difficult to explain through a screen interpreter. Diana’s story implicitly compares screen interpreting to having the interpreter present. When I asked her what makes this difficult when using technology, she responds that it has to do with body language and access to resources they would ordinarily use if they were in the ward.

Anna also mentions embodied resources in interaction when she talks about a situation that worked well for her work:

Yes, it was it was more like concrete in that case. It was more, a consultation that was more like a medicinal evaluation. Where we began, and then sessions where we talked- parts of sessions, and those sessions where the focus was on medication and doses and effects and side effects, where it was more concrete, what we talked about, because It was more about numbers, and yes, no, in a way, than the good relation and contact. There was less room for misunderstanding there in a way. You did not need to see so much body language in a way to feel that you had contact ((laughs))

Extract 5

Anna works at a different hospital ward and has a different field of specialization. To her, medicinal adjustments are easier to talk about using screen interpreting. She talks about this as something concrete. Although the video technology affords the participants at the ward and the interpreter with visual access to each other, the medical professionals express that they lack some of the resources they are used to using in the interaction.

To medical professionals the experience of screen interpreting, and how it functions as a channel of communication, depends on the specific work they are carrying out in the situation. Medical professionals meet with patients in meetings with various purposes in a specialized material settings. They might be concerned with assessing a patient, giving or receiving information, adjusting medication, discussing treatment options and treating patients through interaction. They have different technical setups and different facilities. They use different resources in their
management of the interaction and utilizing these same resources in the screen interpreted setting with varying success.

5.2 Accomplishing interpreting

The organization of talk in turns is fundamental to social interaction (Sacks, Schegloff, & Jefferson, 1974) and turn taking is considered to be central to interpreters’ work (Englund Dimitrova, 1997; Frøili, 2001; Skaaden, 2013). While problems with turn taking in video-mediated interpreting was reduced to matters of “wobbly turn taking” in the hearing to the draft law (Det Kongelige Kunnskapsdepartement, 2019), this is not only relevant to the work of the interpreter. It is relevant to the accomplishment of the medical appointment. Beata and Charlotte told me that they found that the specific interactional organization poses a burden on the patient in the screen interpreted event:

Yes, and it isn't because the interpreter does a bad job. It is because it is the whole situation, screen and sound, and that [the patients] are not able to relate to that they have to stop. It is difficult to stop talking at the right place so the interpreter can interpret. This affects a lot of the conversation. (…) Then there is not, then I experience at least, that even though the patient has to relate to the interpreter in the room, it is less choppy than having the screen over there and then they should look naturally enough at the screen and at us. It becomes a lot more exhausting for them. Or so it seems.

Extract 6

To Beata and Charlotte, screen interpreting, compared to having the interpreter present, causes an extra burden for the patient. The materiality of the video-technology and the physical arrangement in the room makes it difficult for the patients to coordinate their talk with the interpreter. The patient has to present their utterances in chunks so the interpreter can interpret, and although this is the case also when the interpreter is present, this becomes more pronounced when the interpreting is mediated, they say. Beata and Charlotte attribute this in part to the materiality of the media. The patient has to look at the interpreter on the screen and look at the professionals at the ward alternately. This is exhausting for the patient. They suggest that this causes the interaction to become choppy or disjointed. While the official documents acknowledge that there is something about the turn taking, this is not connected to the accomplishment of for instance medical appointments. Several of the medical professionals find it to be a problem that the interaction is choppy with screen interpreting compared to onsite interpreting.
Responding to a question about what recommendations medical professionals should receive before a screen interpreted event, an interpreter, Farzana, suggests that the participants at the ward should look at the interpreter.

It is simply that they look at the screen, both sides, patient and healthcare professional. Look at the screen. Look at me. When they talk. It is not to look at each other. Because it becomes a completely different type of assignment and then they can see my signals. Okay, I have to interpret now. Stop. Take it easy. Like all these hand signals one uses. Just take it easy. Because it becomes something completely different than to look at- they have to pay attention to the screen. To me. And that is the opposite of ordinary interpreting assignments, where we ask them to look at each other. So here, they have to look at me on the screen to get – or not all the time, but at least all, don't forget me. They can't forget me. To get signals from me. Okay, stop now. So I don't have to go completely - use my voice to interrupt the conversation. They can do that now and then, makes it easier for me to interpret.

Extract 7

Farzana suggests that screen interpreting needs to be organized in a different way at the ward than what the participants are used to in interpreted encounters. In order for the doctor and patient to perceive interpreter’s more subtle signals to take the floor, the participants need to be seated so they are all looking at the screen with the interpreter. Then the interpreter does not have to interrupt the conversation.

In addition to the organization of talk, both medical professionals and interpreters mention how the media causes the transmission to drop out and bits to go lost. To the interpreter this might cause worry. Talking about the difference between onsite and screen interpreting an interpreter suggests that things disappear in the mediated setting.

The difference is that one can get a lot more with onsite than one does with screen interpreting. Screen interpreting has its disadvantages and that is that one does not get all the details, then. Mimicry. And it depends on the sound quality too. If it is good, one can get everything that is said. If not, then pieces can drop out or one does not register it at all.

Extract 8

Several interpreters mention how elements of the interaction get lost in screen interpreting: words, details, nuances, emotions. In the work with accomplishing interpreting within the mediated setting, dropouts and loss of information are not necessarily minor details. They may have consequences for the work of the interpreter. Ghalib, an interpreter, talks about the complexity of asking the participants to repeat what they said, initiating repair, in the specific setting.

But when you do not get what is said in an interpreting setting, then one has to ask for repetition as the interpreter, right. You cannot just let it pass. So, what often happens then, with- with applicants
that are at the interview, is that when you ask them to repeat, you rarely get the same information back. They will often continue on their story, building on what they have already said that never was translated. And then it might make little sense to the interviewer or then they say it in a different way than what was originally said. I am kind of just give me that sequence, just like you just said it. I don’t want more. It was enough, a lot that is. It was a lot of information there. But unfortunately it doesn’t go that way. And then a lot of information gets lost.

Extract 9

The complexities of conducting their work in the mediated setting is not just a matter of convenience to the interpreters, it is a matter of the outcome of the meeting. When video-mediation is found to cause loss of content, emotional and informational, this is a concern when there is a lot at stake. For instance when someone is having their asylum case tried. If this one piece of information is the one piece that gets lost, what might the consequences be?

Medical professionals and interpreters in interviews engage in the process of remediation when talking about their experiences with video-mediated interpreting. The comparison is tightly connected to their practice.

5.3 Materialities

Medical professionals and interpreters differ in the way they reflect upon remediation. Opposed to remediation in government documents, practitioners’ statements are similar to one another in that they consider the concrete material setting and conditions of using screen interpreting. Attention to the materiality as a property or a feature of video mediated interpreting occurs in the data as a focus distinguishing between those who deal with the media to accomplish interaction and those who deal with the media as a matter of funding, politics or policy making. The materialities occur in practitioners’ narratives with consideration to how the material setting of the media affords the conduct or smooth interaction. In the analysis of remediation, materialities emerge as a sub-aspect that occurs from a specific institutional perspective rather than other institutional perspectives. While the medical professionals talk about accomplishing meetings, the materialities of the media and the setting are prominent. This reflects on the complex spatial arrangements they coordinate in their work.

In government documents, the comparison between screen interpreting, telephone interpreting and onsite interpreting is often based on the idea of interpreting as a service. Remote interpreting is suggested to be an option because, compared to onsite interpreting it increases access to qualified interpreters, is cost efficient as it can be used to cut down interpreters travel time, and
it is considered to be a way to ensure person security. Screen interpreting is suggested to be the better of the two remote options based on arguments regarding the technological affordances. While the ONR mainly focuses on video-technology as a way to provide services, those whose work will be altered by the implementation of technology, respond to the suggestions by addressing how the nature of their work may be changed. This suggests that moving the interpreter out of the facilities not only affects the work of the interpreter, but in fact also the work of those who work with interpreters.

6 Discussion: Practice meets policy

While some medical practitioners also mention being able to access qualified interpreters as a part of the benefits of video-technology, even in the urban Southern Eastern Norway area, when talking about screen interpreting the practitioners talk about the accomplishment of interaction. Interpreting is different than many services in that is achieved within the interaction it facilitates. Interpreting is interactionally achieved by participants in situ, and video-mediation is therefore not just a matter of service provision but also a matter of changing the socio-material setting and consequently a matter of accomplishing interaction. To medical professionals the materiality of the media becomes relevant to the accomplishment of the interaction. Accommodating medical practices with videoconference systems at the ward may not always be easy. Ensuring that the interpreter has a view supporting the activity of interpreting while at the same time positioning participants at the ward can be challenging. To interpreters it is a matter of being able to conduct their work in the video-mediated interactional space. To the medical professionals it is a matter of accomplishing the medical purpose of the meeting and accomplishing the interpreter with the interpreter at a different location. In the hearing and the draft law, concerns regarding trouble with the interaction, such as turn taking, are not connected to the achievement of the meetings or the socio-materiality of the setting. Rather, the problems that are reported of are restricted to affect the interpreting or interaction without connecting it to the quality of public services.

   Government documents do not simply reiterate knowledge from various sources. Through more or less transparent intertextual links to other documents, the documents generalize, reinforce and construct knowledge. ONRs appear as coordinated texts resting upon both epistemic, social and political authority, which in turn makes them powerful instruments for governing (Tellmann, 2011,
Furthermore, such reports combine the voice of researchers, organized stakeholders, representatives from public authorities and other stakeholders. This tends to blur the boundaries between science and politics (Tellmann 2011, 124). In the construction of a political rationale for increased use of video-technology for providing interpreting, the boundaries between science and politics have been blurred. What is presented as knowledge, constructs political arguments to increase the use of video-technology for screen interpreting.

The media ecology such as is has been described here, is not static nor is it universal, it is local and temporary. For instance, the COVID-19 pandemic has called for changes in technological constellations in provision of public services causing an increased number of interpreted-mediated meetings to be carried out with all participants at separate locations. Although the channel itself, video-technology, is not new to the media ecology, the introduction of new arrangements of the mediated event occasion new debates and comparisons.

This article has demonstrated how media ideologies as a concept can contribute to an understanding of tensions between positions held by stakeholders regarding implementation of technologies to a media ecology. The concept of remediation has provided insights into a different perspective on video-mediated interpreting in Norwegian society. Finally, this article demonstrates how ideologies drawn upon by government authorities do not necessarily correspond to the reality of practitioners’ work. This article has demonstrated that stakeholders’ media ideologies differ. Media ideologies go beyond aspects of new media and social media, and can be found in authorities’ treatment of processes of digitalization. The understanding of how media ideologies are included in institutional and professional discourses has not yet been extensively researched.

7 References


---

9 This article does not expand on this as the data was collected before the pandemic.


