Social Interaction

Video-Based Studies of Human Sociality

_____

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

_____

Jessica P. B. Hansen

University of Oslo, Multiling - Center for Research on Multilingualism

in Society across the Lifespan

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, in order to 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 participating 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 s/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 they 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 brief consultations in an outpatient clinic. 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.
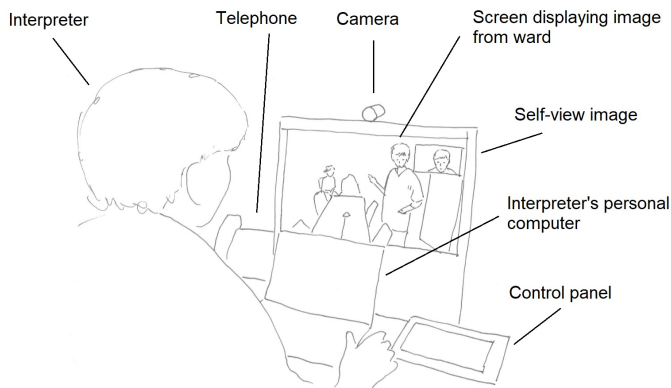

Illustration A: An interpreting studio

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.
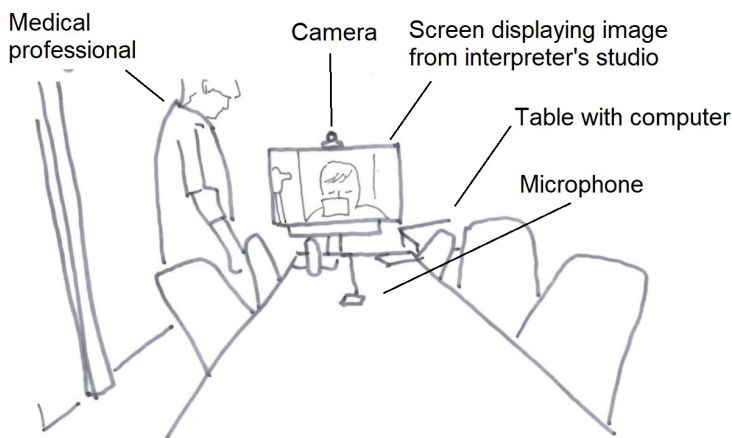

Illustration B: A hospital meeting room

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).
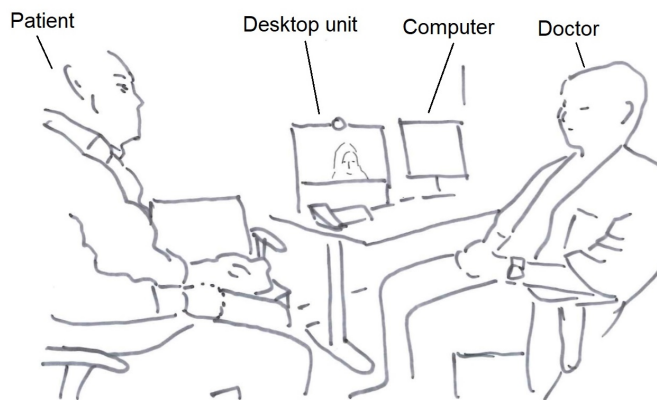

Illustration C: Meeting room in an outpatient clinic

The 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
      mp:            ¤facing SCR/INT, pointing to SCR/INT and RES
3    MP:    *så da bru*ker ¤e:-
              so then uses   e:-
      mp:        -->        ¤turns to doorway-->
      int:   *nods       *
4              (0.6)¤(0.5)
      mp:           ¤gestures to PAT/NOK-->
5    MP:    vær så go:
             please
6              #¤(0.8)*(0.6)#
      mp:       ¤pulls out chair and gestures to chair
      pat:   walks into the room toward seats -->
      int:           *nods
      fig:   #fig1.1     #fig1.2
7    INT:   bardzo proszę zapraszam
             please welcome
8              &(1.1)
      pat:   &takes a seat -->>>
```


Figure 1.1

Figure 1.2

```
9      MP:    så da e bruker e:hm (0.3) melissa litt tid fø:rst,
               so then  melissa will use a little time first
10            (1.1)
11            til å inform#ere de,
               to inform them,
       fig:              #fig 1.3
12            (0.4)
13     RES:   ja,
               yes
14            (0.4)
15     INT:   e::m najpierw, e: parę słów powie Melissa,
               uh:m first     uh Melissa will say a few words
16            żeby państwa poinformować,
               in order to inform you (polite, plural)
17            (0.4)
18            e: men før dere bynner samtalen
               but before you begin the conversation
19            plikter tolken å orientere om sin rolle,
               the interpreter is obliged to inform about her role
20     MP:    vær så snill,
               please do
21     INT:   [e::] (0.4) minne om sin rolle,
               [e::] (0.4) remind of her role
22     MP:    [ja ]
```

Figure 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.
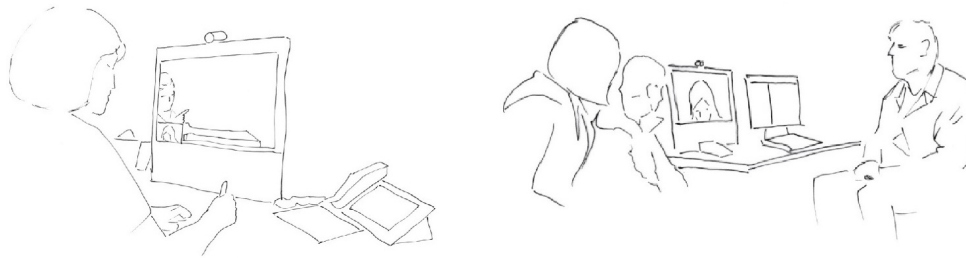
Figure 2: Interpreter's point of view to the left. Ward's point of view to the right.

Extract (2A)

Figure 2.1

```
1              (.)
2      DR:     hvordan 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 Ê  (.) dạ,  chị ra sao?
               .h uh (.) yes, how is big sister
9              Chị cảm thấy trong người khỏe hay không?
               does big sister feel healthy or not
10     NOK:    không,¤ ảnh chứ hỏng có chị đâu.#
               no,      it is big brother and not big sister
       nok:         ¤leans forward and turns to screen--->
       fig:                            #fig 2.1
11     NOK:    Ch&ị đi theo:      D&ợ thôi.
               Big sister followed. I am only the wife.
               (0.7)
12     INT:    .h e:: å&nei 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
```

11

```
15    DR:    #jehe:g vet det
              i::h   know
      Fig:   #fig 2.2
16    DR:    [hehe]he[hehehe]
17    NOK:   [ja::]
               yes
18    PAT:           [ja    ]
                      yes
19    NOK:   %hhh
      dr:    %leans toward PAT --->>
20    DR:    og du føler deg frisk?
             and you feel healthy
```


Figure 2.2

Gazing at the patient, the doctor asks the patient how he is feeling (line 2–4). 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 *chị* (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 thắc mắc gì 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.

Extract (3)

```
1     DR:    e:::m
             uh:m
2            (0.6)
3            er det greit for deg å fortsette me: 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)
      dr:         ¤turns to screen
      dr:               ¤stretches out arm to screen
9     DR:    hm¤
                ¤retracts arm
10    PAT:   hehe
11    DR:    hehe vi må bare bruke 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 take this up, 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.

Extract (4)

```
1    INT:    .hhhh laew kaw thuksing thukyang thi rao phud theenee
             .hhhh and in addition to every single thing that is said here
2            rao kep pen khwamlap duay. (.) kha
             I (we) will keep silent    (.) prt
3            (1.1)
4    PAT:    kha
             yes
5    INT:    .hh
6            (0.9)¤(0.2)¤(0.2)#
     int:         ¤nods ¤palms facing down, out to side and back to centre->
     fig:                   #fig4.1
7    INT:    to-
8            &(0.3)¤
     int:     -->  ¤
     dr:     &smiles
9    DR:     &okay?
             &turns to SCR/INT-->
10   INT:    ja
11           (0.6)
12   DR:     var det greit?
             was that okay
```
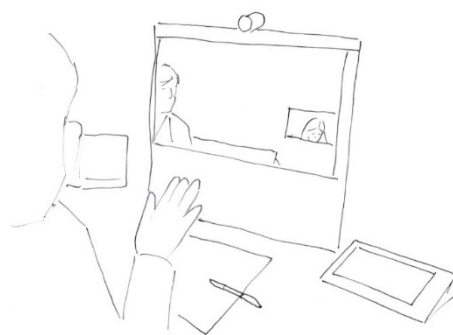

Figure 4.1

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 4.1). 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.
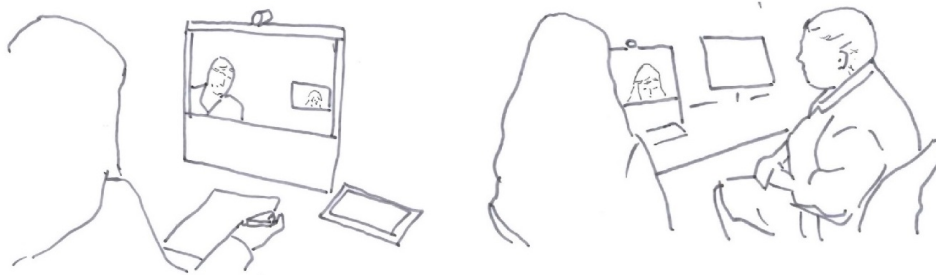
Figure 5: The interpreter's visual access to the left. The ward's perspective to the right.

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 intere¤ssert,¤
              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 in¤trodusere 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
```

18

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

References

Angermeyer, P. S. (2007). 'Speak English or what?' Codeswitching and interpreter use in New York Small Claims Courts. *International Journal of Speech Language and the Law, 14*(1), 147. doi:10.1558/ijsll.v14i1.147

Arminen, I., Licoppe, C., & Spagnolli, A. (2016). Respecifying Mediated Interaction. *Research on Language and Social Interaction: Orders of Interaction in Mediated Settings, 49*(4), 290-309. doi:10.1080/08351813.2016.1234614

Balogh, K., & Hertog, D. (2012). AVIDICUS comparative studies - part II: Traditional, videoconference and remote interpreting in police interviews. In S. Braun & J. L. Taylor (Eds.), *Videocoference and Remote Interpreting in Criminal Procedings* (pp. 101-117). Antwerp/Cambridge: Intersentia.

Bolden, G. (2000). Toward Understanding Practices of Medical Interpreting: Interpreters' Involvement in History Taking. *Discourse Studies, 2*(4), 387-419. doi:10.1177/1461445600002004001

Bolden, G. (2018). Understanding interpreters' actions in context. *Communication & Medicine, 15*(2), 135-149. doi:https://doi.org/10.1558/cam.38678

Braun, S., & Taylor, J. L. (2012). AVIDICUS Comparative Studies - part I: Traditional interpreting and remote interpreting in police interviews. In S. Braun & J. L. Taylor (Eds.), *Videoconference and Remote Interpreting in Criminal Procedings* (pp. 85-101). Antwerp/Cambridge: Intersentia.

Davitti, E. (2019). Methodological explorations of interpreter-mediated interaction: novel insights from multimodal analysis. *Qualitative research : QR, 19*(1), 7-29. doi:10.1177/1468794118761492

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

Deppermann, A. (2013). Multimodal interaction from a conversation analytic perspective. *Journal of Pragmatics, 46*(1), 1-7. doi:10.1016/j.pragma.2012.11.014

Det Kongelige Kunnskapsdepartement. (2019). *Høring Forslag til lov om offentlige organers ansvar for av tolk mv. (tolkeloven)*. Retrieved from https://www.regjeringen.no/no/dokumenter/horing/id2626660/

Dourish, P., Adler, A., Bellotti, V., & Henderson, A. (1996). Your place or mine? Learning from long-term use of Audio-Video communication. *Computer supported cooperative work, 5*(1), 33-62. doi:10.1007/BF00141935

Gavioli, L., & Baraldi, C. (2011). Interpreter-mediated interaction in healthcare and legal settings: Talk organization, context and the achievement of intercultural communication. *Interpreting : international journal of research and practice in interpreting, 13*(2), 205-233. doi:10.1075/intp.13.2.03gav

Hansen, J. P. B. (2018). Tolking som felles aktivitet: Med lupe på kommunikative ståsted i tolkede sykehussamtaler. In H. Haualand, A.-L. Nilsson, & E. Raanes (Eds.), *Tolking - språkarbeid og profesjonsutøvelse* (pp. 123-142). Oslo: Gyldendal Akademisk.

Hansen, J. P. B., & Svennevig, J. (forth.). Creating space for interpreting within extended turns at talk. *Manuscript in preparation*.

Hazel, S., Mortensen, K., & Rasmussen, G. (2014). Introduction: A body of resources – CA studies of social conduct. *Journal of Pragmatics, 65*, 1-9. doi:10.1016/j.pragma.2013.10.007

Heath, C., & Luff, P. (1993). Disembodied conduct: interactional asymmetries in video-mediated communication. In G. Button (Ed.), *Technology in working order. Studies of Work, Interaction, and Technology* (pp. 35-54). London: Routledge.

Jefferson, G. (1979). A technique for inviting laughter and its subsequent acceptance / declination. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology*. New York: Irvington.

Li, S. (2015). Nine Types of Turn-taking in Interpreter-mediated GP Consultations. *Applied Linguistics Review, 6*(1), 73-96. doi:10.1515/applirev-2015-0004

Li, S., Gerwing, J., Krystallidou, D., Rowlands, A., Cox, A., & Pype, P. (2017). Interaction—A missing piece of the jigsaw in interpreter-mediated medical

consultation models. *Patient Education and Counseling, 100*(9), 1769-1771. doi:10.1016/j.pec.2017.04.021

Licoppe, C., Luff, P., Heath, C., Kuzuoka, H., Yamashita, N., & Tuncer, S. (2017). Showing Objects: Holding and Manipulating Artefacts in Video-mediated Collaborative Settings. In In *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Vol. 2017-, pp. 5295-5306). https://doi.org/10.1145/3025453.3025848

Licoppe, C., & Verdier, M. (2013). Interpreting, video communication and the sequential reshaping of institutional talk in the bilingual and distributed courtroom. *The international journal of speech, language and the law, 20*(2), 247. doi:10.1558/ijsll.v20i2.247

Licoppe, C., Verdier, M., & Veyrier, C.-A. (2018). Voice, Power, and Turn-Taking in Multilingual, Consequtively Interpreted Courtroom Proceedings with Video Links. In J. Napier, R. Skinner, & S. Braun (Eds.), *Here or There : Research on interpreting via video link* (pp. 299-322). Washington DC: Gallaudet University Press.

Licoppe, C., & Veyrier, C.-A. (2017). How to show the interpreter on screen? The normative organization of visual ecologies in multilingual courtrooms with video links. *Journal of Pragmatics, 107*, 147-164. doi:10.1016/j.pragma.2016.09.012

Luff, P., Heath, C., Kuzuoka, H., Hindmarsh, J., Yamazaki, K., & Oyama, S. (2003). Fractured Ecologies: Creating Environments for Collaboration. *Human-Computer Interaction: Talking About Things in Mediated Conversations, 18*(1-2), 51-84. doi:10.1207/S15327051HCI1812_3

Meeuwesen, L., Twilt, S., ten Thije, J. D., & Harmsen, H. (2010). "Ne diyor?" (What does she say?): Informal interpreting in general practice. *Patient Educ Couns, 81*(2), 198-203. doi:10.1016/j.pec.2009.10.005

Mondada, L. (2007). Operating Together through Videoconference: Members' Procedures for Accomplishing a Common Space of Action. In S. Hester & D. Francis (Eds.), *Orders of Ordinary Action. Respecifying Sociological Knowledge* (pp. 51-68). London: Routledge Taylor & Francis Group.

Mondada, L. (2014). The local constitution of multimodal resources for social interaction. *Journal of Pragmatics, 65*, 137-156. doi:10.1016/j.pragma.2014.04.004

NOU 2014:8. (2014). *Tolking i offentlig sektor. Et spørsmål om rettssikkerhet og likeverd.* Oslo: Barne-, likestillings- og inkluderingsdepartementet Retrieved from https://www.regjeringen.no/no/dokumenter/NOU-2014-8/id2001246/

Paananen, J., & Majlesi, A. R. (2018). Patient-centered interaction in interpreted primary care consultations. *Journal of Pragmatics, 138*, 98-118. doi:10.1016/j.pragma.2018.10.003

Pappas, Y., & Seale, C. (2010). The physical examination in telecardiology and televascular consultations: A study using conversation analysis. *Patient Education and Counseling, 81*(1), 113-118. doi:10.1016/j.pec.2010.01.005

Rintel, E. (2010). Conversational management of network trouble perturbations in personal videoconferencing. In *ACM International Conference Proceeding Series* (pp. 304-311): ACM.

Schegloff, E. A. (2007). *Sequence organization in interaction.* Cambridge: Cambridge University Press.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The Preference for Self-Correction in the Organization of Repair in Conversation. *Language, 53*(2), 361-382. doi:10.2307/413107

Schutz, A. (1953). Common-Sense and Scientific Interpretation of Human Action. *Philosophy and Phenomenological Research, 14*(1), 1-38. doi:10.2307/2104013

Skaaden, H. (2001). *On-Screen Interpreting.* Paper presented at the 18th International Symposium on Human Factors in Telecommunication Bergen.

Skaaden, H. (2013). *Den topartiske tolken : lærebok i tolking.* Oslo: Universitetsforl.

Streeck, J. (2009). Forward-Gesturing. *Discourse Processes: Projection and Anticipation in Embodied Interaction, 46*(2-3), 161-179. doi:10.1080/01638530902728793

Wadensjö, C. (1998). *Interpreting as interaction*. London: Longman.