

Fight for focus: attention and agency in sight-translated interaction

Randi Havnen

To cite this article: Randi Havnen (2021): Fight for focus: attention and agency in sight-translated interaction, Perspectives, DOI: [10.1080/0907676X.2021.1892785](https://doi.org/10.1080/0907676X.2021.1892785)

To link to this article: <https://doi.org/10.1080/0907676X.2021.1892785>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 12 Mar 2021.



Submit your article to this journal [↗](#)



Article views: 205



View related articles [↗](#)



View Crossmark data [↗](#)

Fight for focus: attention and agency in sight-translated interaction

Randi Havnen 

Department of International Studies and Interpreting, Faculty of Education and International Studies, Oslo Metropolitan University, Oslo, Norway

ABSTRACT

This paper presents an analysis of three roleplayed interpreted institutional meetings in which sight translation is part of the interaction. The analysis is based on multimodal (inter)action analysis and utilises the analytical tool of modal density as indication of attention/awareness. This analytical framework is novel in interpreting studies. The data include filmed material from an experimental setting and participants' reflections about the situation. The findings show variations in sight translation practices and that the shift from interpreting to sight translation affects interactional patterns, particularly social actors' attention and agency. In my discussion of agency in sight-translated interaction, I argue that interpreters, in addition to translating, need to pay attention to interactional issues related to attention and agency caused by the interpreting method.

ARTICLE HISTORY

Received 7 May 2020
Accepted 15 February 2021

KEYWORDS

Sight translation; interaction; attention/awareness; agency; interactional responsibility

Introduction

Sight translation, or interpreting a written text into speech, is a common part of an interpreter's job. In research, however, this practice is not scrutinised from an interactional perspective, as the primary research focus thus far has been linguistic and, to a certain degree, monologist (Havnen, 2019; Vargas-Urpi, 2019). In this study, I draw on a dataset of three interpreter-mediated roleplayed meetings between Serbian-speaking public service users (PSUs) and a public service representative (PSR). During the meeting, the interpreter is asked to do a sight translation; as we shall see, this is an action that significantly influences the interaction. The data also include the participants' reflections about the interaction and serve to supplement the findings.

The analytical framework utilised in this study is that of multimodal (inter)action analysis (MIA) (Norris, 2004, 2019). MIA is based on the concept of mediated action rooted in the works of Scollon (1998a) and Wertsch (1998). An important aspect is that social interaction is co-produced through mediational means/cultural tools in which language is not necessarily the centre of attention. Norris (2004, 2019) has developed analytical tools that are especially suitable for investigating social actors' engagement in simultaneous activities.

CONTACT Randi Havnen  ranha@oslomet.no

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

The activities are analysed as higher-level mediated actions (actions with an opening and a closing, such as a meeting) and lower-level mediated actions (pragmatic meaning units of modes, such as an utterance or gaze shift, which has a start and an end). Pirini (2016, 2017) has demonstrated that the analytical tool of modal density as an indication of attention/awareness (Norris, 2004, 2019) can be further developed to determine intersubjectivity and agency. I will return to this in the analytical framework section.

The motivation to analyse the interaction through the multimodal (inter)actional framework stems from the feedback of the participants, in addition to my own observations. The PSR and the PSU described the struggle to maintain focus during the sight translation; the interpreters, on the other hand, were fully occupied with their own activity. When going through the data, I identified a clear shift in the interactional rhythm when the sight translation started. The PSR and the PSU refrained from talking. Talk and linguistic mediation have traditionally been the centre of attention in interpreting studies. I wanted to take a closer look at interactional aspects beyond language and was curious whether the MIA framework could help me describe and understand what was happening.

Research has identified that interpreters have a significant effect on the coordination of turns and content (Wadensjö, 1998). Contrary to the widespread belief in interpreting research that interpreters play an active role in interactions, both lay people and scholars often treat interpreters as mediational means: 'Translators are solely there to make the interaction between the two politicians possible, i.e., they are viewed, act and react as mediational means' (Norris, 2019, p. 39). The interpreter's position in situated practice is seldom as straightforward as making the interaction possible without interfering; however, static concepts of translation are not uncommon in multimodal studies (and vice versa) (Kaindl, 2020). In this paper, I treat the interpreter as a social actor in the sense that the interpreters' actions beyond the act of translation affect the interactional pattern.

Before further presenting my study, I examine the practice of sight translation and offer a background on previous interactional research including sight translation. Then, I present the analytical framework, followed by the data, the analysis and the discussion. I end the article with the concluding remarks.

Sight translation

Practices

The term 'sight translation' is used to describe the method by which a written text is mediated into speech in another language. Sight translation might be a method for language learning or an exercise for developing interpreting skills (Čeňková, 2015; Chen, 2015). It is also used as a tool for translators who speak their written translations, which are then machine transcribed into written text (Dragsted et al., 2009). Written texts are sometimes used to support simultaneous interpreting (sight interpreting or simultaneous interpreting with text). Even if sight translation is traditionally associated with interpreting training or conference interpreting, it is also frequently used in meetings, both in bilateral negotiations and in public service encounters (Čeňková, 2015; Chen, 2015; Li, 2014; Nilsen & Havnen, 2019; Vargas-Urpi, 2019). The interpreter might not be prepared for the task, rather being asked to translate a document on site. In some assignments, sight translation is always expected, such as written reports at the end of police hearings or documents presented in

court (Maatta, 2015). Sight translation sometimes makes up the main part of the assignment, such as proclaiming a decision or verdict on site or over the phone. Sight-translated texts may also be recorded and handed to the receiver in the form of a spoken document (Biela-Woloniciej, 2015). Sight translation practices are thus multifaceted and far from being standardised, oftentimes not thought or regarded as a specialist translation method but is also treated as one somewhere. Belgian codes of ethics, for example, advise against sight translation (Maatta, 2015). In Norway, sight translation is not mentioned in professional codes of ethics, but it forms part of the degree of BA in Interpreting in the Public Sector (Nilsen & Havnen, 2019). In some countries, interpreters are tested in sight translation for certification, but the criteria for competence assessment are ambiguous (Paez, 2014).

Interpreting research

Documents have been studied from a multimodal perspective as artefacts when they have formed part of an interpreted interaction at a social centre (Ticca & Traverso, 2017) and as part of interpreted parent–teacher meetings (Davitti & Pasquandrea, 2017). These studies found that handling the document reconfigured the participation framework in terms of turn-taking and inclusion/involvement. A case study from a Belgian police hearing described how an on-screen document was used in turn-taking coordination (Defrancq & Verliefdde, 2018). Vargas-Urpi's (2019) exploration of sight translation as dyadic or triadic revealed that the untrained interpreters whom the author studied adapted the translation quite freely to the listeners' needs but excluded the PSR. The trained interpreter in the study included the PSR by translating questions from the listener instead of explaining them (Vargas-Urpi, 2019). A recent exploration of changes in meaning-making in sight translation found that meaning-making is affected not only as a result of the interpreter's strategies when translating but also because of the modal shift from print (writing) to speech (Havnen, 2020). The change was especially evident at the interactional level, which is the focus of the present article.

Interpreting studies applying MIA are scarce even if it is considered particularly useful for such purposes (Kaindl, 2020, p. 56). A pedagogical study of dialogue interpreting using Norris' analytical framework highlights the importance of the interpreter paying attention to including and excluding interlocutors; the framework also proves useful for students' understanding of interaction beyond language (Krystallidou, 2014). This study contributes to the growing body of multimodal interpreting and translation studies (Boria & Tomalin, 2020; Davitti, 2019; Perez-Gonzalez, 2014; Tuominen et al., 2018).

Analytical framework

Norris (2004, 2019) has developed analytical tools to differentiate actions on different levels, with mode being a central theoretical unit. A mode is a system of representations with regularities of use, such as gestures, language or layout in a room. Modes have various materiality, such as visibility, audibility or endurance. The smallest pragmatic meaning unit of a mode can be analysed as a lower-level mediated action; examples are utterances, gesture units or use of gaze. A lower-level mediated action has a start and an end, and it represents an action mediated through psychological and physical mediational means/cultural tools. These can be objects, space, knowledge, body parts

and so on. Lower-level actions are linked together in chains that constitute and are constituted by higher-level mediated actions. The higher-level action of reading, for example, constitutes the use of gaze (if seeing; touch if blind) and print (on screen, paper); this mediation also produces higher-level action. A higher-level action has an opening and a closing and can be studied at different scales. It can be a meeting in an office, translating or giving feedback. Higher-level actions are also found at the level of discourse and practice (Norris & Pirini, 2017). The concept of frozen mediated actions covers previously performed actions that are embedded within an object, such as a document which has been written and printed through someone's mediated actions.

To analyse attention/awareness in a foreground–background continuum, Norris (2004, 2019) utilises the concept of modal density, which embraces modal intensity and complexity. Modal intensity refers to the weight or importance a mode has in lower-level actions and its relevance for the constituted higher-level actions scrutinised. For example, in the process of reading, gaze has a high intensity because if you close your eyes, this higher-level action could not be produced, whereas holding a document 2 cm above a table has a lower density because when you let the document go, it lands on the table and the higher-level action can still be produced. Modal complexity is seen through how intricate or intertwined modes are, such as in sight translation in which gaze, print and spoken language are all needed to execute this higher-level action. Modal configuration can refer to the layout in a room, proximity and body orientation.

In addition to utilising the analytical tool of modal density to analyse attention/awareness towards simultaneous higher-level actions, Pirini has further developed this tool to analyse intersubjectivity (Pirini, 2016) and agency (Pirini, 2017). In his study of high school tutoring settings, Pirini (2014) shows how transitions mark changes in attention and how the social actors produce convergent and divergent actions. In the author's intersubjectivity study based on the same material, the modal density tool is extended to isolate three tiers of material intersubjectivity: stable, adjustable and fleeting. Intersubjectivity is understood as co-construction of joint activity (Pirini, 2016). In my case, the layout, setting and proxemics are stable, the body posture and handling the document are adjustable, and the gaze and spoken language are fleeting. Pirini (2017) relates modal density to agency and demonstrates that an actor's agency, understood as the ability to produce and initiate actions, is related to control over the most relevant means in the co-production of a higher-level action – this actor has primary agency. Building on Jones and Norris (2005), he points out the tension between individual agency and the social and material world, where agency is influenced by professional and institutional practices.

Higher-level actions unfold at different scales of action, from a greeting in a meeting to discourses and practices. Norris (2019), building on Scollon's work (1998a), defines practice as an action with history. To be able to align with one another, social actors must share practices (Norris, 2011). Sight translation can be understood as a practice of reading aloud, which is a social practice with some embedded history not necessarily shared. There is scant literature on adults' practices of reading aloud; however, a British study found that reading aloud is mostly connected to the private sphere (Duncan, 2018). This is probably the case for a literate society, where written documents in institutional settings are read in silence. Historically, however, reading aloud is a more widespread practice in which literates would have to read to illiterates for the sake of sharing or for accessibility (Goody, 1987). Reading as a social practice also affects

interaction (Scollon, 1998b), especially because a reader has primacy over a speaker, i.e., the threshold of interrupting a reader is higher than taking the turn from a speaker.

In the discussion, I will also draw on the concept of rhythm, both at the level of the concrete interaction and the level of practice. Van Leeuwen (2005) points out that balance in visual modes and rhythm in audial modes are basic units in human interaction, categorising them as biological. The author considers rhythm to be a cohesive device, structuring actions and supporting coherence in communication. Rhythm also plays a part in ‘getting the message across’ (Van Leeuwen, 2005, p. 181). Certain actions have a patterning effect on interaction, such as background music, in which interlocutors align with the beat. Alignment as a phenomenon in communication has been studied in linguistics with a focus on verbal alignment and in sociological studies with a focus on proximity (distance) (Norris, 2011). Norris (2011) argues that rhythm can also be seen at the level of practices; when people engage in practices with a mutual understanding of these, they know what they are expected to contribute with and when (Norris, 2011).

Data

The filmed material

The starting point of my analysis was a simulated face-to-face meeting in which a PSR for the Norwegian Labour and Welfare Administration and a Serbian-speaking PSU met in a public office environment. They communicated through an interpreter, and after a brief dialogue framing the meeting, the interpreter was handed a leaflet for sight translation before the PSR and the PSU returned to the dialogue. The experiment was repeated two more times with different actors (the same PSR), and they each lasted for approximately 17 (A), 13 (B) and 8 (C) minutes. The meetings were filmed from two different angles (Figure 1).

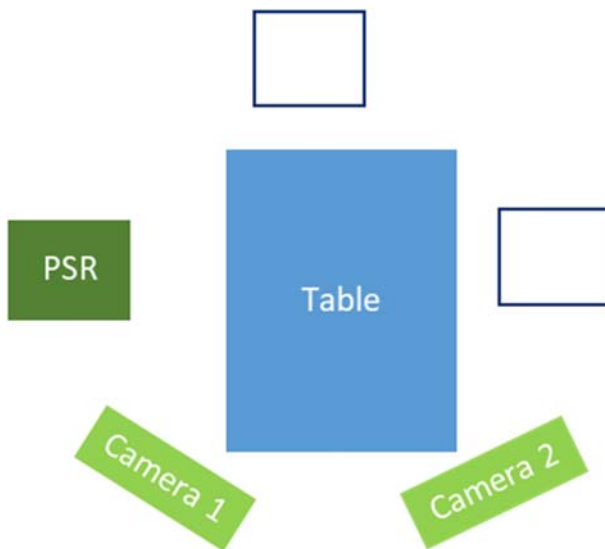


Figure 1. Layout for the roleplay.

The interpreter and the PSU did not sit at the same places in all three situations, hence the two empty boxes in [Figure 1](#). The PSU and interpreter were positioned as follows, as seen from camera 1 ([Figure 2](#)):



Figure 2. Proximity between the interpreter and the PSU and the interpreter's dominant position in relation to the document. Interpreter A and B are positioned to the left, whereas interpreter C is positioned to the right (drawing by Robert Julher based on screenshots in which faces were anonymised).

The PSR, who holds a position in the Norwegian Labour and Welfare administration, had previous experience with roleplaying interpreting exams. She led the meeting. The cameras were turned on before the other participants were guided into the room by me. I was not in the room during the interaction in order not to disturb the interpreters, as I teach and evaluate interpreting. The three interpreters were certified interpreters, who were also familiar with roleplay settings and had no previous knowledge of the other participants. I informed them that sight translation would be part of the experiment and that they should do their jobs as they normally would. The three PSUs were newly arrived working immigrants from Serbia with scant knowledge of Norwegian and with no previous experience in communicating through an interpreter. I told the PSUs that they were about to participate in a meeting to obtain information about what to do in events of illness when in a working relationship. The setup was realistic, with true information concerning a theme that is generally of interest to employees, and there was a real need for interpreting. The main manipulation was that I instructed the PSR to give the responsibility of reading the document to the interpreters and to not interfere unless one of the participants addressed her. This was done because the primary focus was initially how the interpreters would deal with sight translation; the result of this analysis was presented in a previous study (Havnen, 2020). According to students in courses on sight translation, my own experiences as an interpreter and as reported by Felberg (2015), it is common practice for the PSR to leave a document with the interpreter. Sometimes, the PSR excludes themselves from the interaction, occasionally physically leaving the room (Nilsen & Havnen, 2019).

Participants' feedback

The interpreters and the PSR wrote their reflections immediately after each roleplay on a blank document. They were instructed to note whatever came to their mind, and they wrote 150–200 words each. Meanwhile, I interviewed the PSUs. At first, they talked freely, and then I asked them further about the content of the document, whether the interpreter was understandable, their experience of being a listener to sight translation

and their general perspectives on the communication process, if they had not already mentioned it themselves. I will now sum up the issues relevant for this study. As the data are limited in scope, the interviews were around 9, 6 and 4 mins long; they only serve as supplements to the interactional analysis.

The interpreters had a dominant textual focus that was related to their own translation process; however, one interpreter was concerned that there seemed to be a higher threshold for interrupting the sight translation than a speaker in a typical dialogue. This interpreter proposed that maybe the PSR should have asked whether the PSU understood the information before moving to the next topic.

The PSR focused on the difficulty concentrating and expressed uncertainty about the interpreter's solutions and the PSUs' understanding. These doubts were either disproved through the dialogue towards the end of the meeting or were left unaddressed. In setting C ([Figure 2](#)), the interpreter sat on the chair typically meant for the PSU. Here, the PSR also felt that the interpreter was taking over the interaction by focusing solely on the text and paying hardly any attention to the PSU. In terms of posture, the interpreter was positioned towards the PSU, who focused solely on the interpreter. After experiment A, the PSR commented that it would be better to give the information herself orally.

Two of the PSUs mentioned the unusual nature of the situation – not really knowing how to adjust to the interpreting and finding it difficult to remember. One of them argued that he would have preferred to pose questions during reading but was concerned that it was inappropriate, hypothesising that the interpreter purposely did not entertain questions. Another PSU described the experience as follows: '(...) if I am allowed [to say so] – [it was] a little boring'. She described a feeling of being back at school. Engagement was easier when she could relate personally to the content than to abstract information. She said it would be different to read herself. In the experiment in which the PSR felt excluded, the PSU did not comment on any factors other than the interpreted text – the PSU said the interpreter was clear, easy to follow and understandable.

Analysis

I relied on the filmed material to do the analysis, evaluating the interaction *live*, rather than doing detailed transcriptions. This way, the data did not become graphic and linear but instead retained the dynamics that I find are getting lost in transcriptions, especially when focusing on larger scales of actions. As a start, a technician merged the two films into one frame and synchronised them, as shown in [Figure 3](#):

I first identified the relevant modes that were used in the interaction – layout, proximity, posture, gaze, gestures, head, body and hand movements, spoken language and print, and touch (handshakes) – as openings and closings. In my study, I wanted to focus on attention/awareness towards a shared higher-level action over time. Time is not a traditional delineation in MIA, in which the typical starting point is the analysis of the simultaneous production of several higher-level actions at the micro level (Pirini, 2015). In my material, there was an obvious shift from converging to diverging higher-level actions when the sight translation started. These actions were still part of the shared larger-scale higher-level action – the giving and receiving of information. As the production of these higher-level actions appeared stable over time, I decided to



Figure 3. Film from camera 2 with an inserted film from camera 1. On the big screen are the interpreter and the PSU; on the small screen are the PSR and the interpreter. These are synchronised in time, including sound.

focus on the meso level and the different social actors' attention and agency throughout the meeting by utilising the analytical tool of modal density.

To evaluate modal density, I analysed each social actor in time stretches of around 10 s but varying from 5 to 40 s. I used three criteria for delineating the stretches:

- A turn
- If a turn was very short (a question), I included the interpretation and the answer as a turn stretch.
- If a turn was long, I stopped it on a potential turn shifting point (pause or inbreath) or after a meaning unit (guided by the source text document during the sight translation).

I watched one actor in a time stretch and evaluated modal density based on the chains of the lower-level actions producing the higher-level action of interest. Then, I went back to watch the next social actor, rewinding again to evaluate the modal density produced by the third actor. Then, I would rewind again to look at them in relation to each other and double-check when I was in doubt. I plotted the values of low, medium and high modal density into a spreadsheet (Figure 4). Somewhere, there is a low/medium modal density; this is where at the end of the turn, gazes meet, or there is another chain of lower-level action that intensifies in a stretch that had so far been of low density, and there is a mutual exchange of gaze or content feedback (through a nod, for example). The blue marks in the time columns illustrate stretches longer or shorter than average to check whether they affected the pattern (at the meso level, it did not).

Both complexity and intensity were considered when evaluating density. Modes have no a priori density, density relates to the mediation of the mode and the importance in producing higher-level actions. Gaze, which is dominant in face-to-face interaction, takes on various meanings and intensities. I evaluated the gazes as either semiotic means, such

Situation A	PSR	INT	PSU	Action
01.00 - 01.15	High	High	High	Greetings
01.15 - 01.25	High	High	Medium	Introduction
01.26 - 01.36	High	Medium	Medium	
01.37 - 01.46	Medium	High	Medium	
01.47 - 01.57	High	Low	Low	
01.58 - 02.07	Low	High	Medium	
02.08 - 02.11	High	High	High	Q - A (question - answer)
02.12 - 02.17	High	High	High	Q - A
02.18 - 02.24	High	High	High	Q - A
02.25 - 02.36	High	High	High	Q - A
02.39 - 02.50	High	Low	Low	Q
02.51 - 03.01	Low	High	Medium	
03.02 - 03.07	High	High	High	A -R (Answer - response)
03.08 - 03.18	High	Low	Low	
03.19 - 03.32	Low	High	Medium	
03.33 - 03.37	High	High	High	A - R
03.38 - 03.48	High	High	High	Handing over paper
03.49 - 04.00	Low	High	Medium	
04.01 - 04.12	Low	High	Medium	
04.13 - 04.22	Low/Med	High	Low	
04.23 - 04.32	Low	high	Low	
04.33 - 04.48	Low	High	Medium	
04.49 - 05.02	Low	High	Low	
05.03 - 05.21	Low	High	Medium	
05.22 - 05.34	Low/Med	High	Low/med	
05.35 - 05.44	Low	Med	Low	Int reading in silence
05.45 - 06.00	Low	High	Low	

Figure 4. Annotation of modal density in time stretches in the situation including interpreter A for all participants.

as beat actions (structuring), or whether gaze is used pragmatically, such as content ratification or interactional feedback, which takes a different density. Gaze was also used for reading and monitoring, taking on a high intensity because of its importance in the production of higher-level actions. Conversation analysis (CA) has revealed that consecutive interpreter-mediated interaction follows dialogical patterns of typical face-to-face interaction. Even when a listener does not understand a speaker, which is the case for the PSR during the sight translation in my study, gaze is used as a listener response. In this manner, gaze might take on density as feedback when it is not a reaction to another gaze or to content. This phenomenon is described as dual feedback, as gaze is a response to the original speaker and to the interpreter, creating common ground in a triad (Vranješ et al., 2018).

I also identified several instances of deviating gaze shifts, fiddling, jiggling a leg and changing the body posture. These can be beat actions, and/or they can be interpreted as chains of lower-level actions constituting a higher-level action of being bored or trying to stay focused. More fine-grained analysis is needed to establish meaning and function. MIA does not explain internal thoughts and experiences, so what goes on in the mind could only be an assumption, a plausible assumption, though when combining the findings with the participants' reflections. I will return to this in the discussion. Here, I focus on the lower-level action producing the shared higher-level action.

In addition to evaluating the intensity of the modes at play, I evaluated modal complexity, such as the interpreter utilising print, gaze, handling the document, gestures and spoken language when sight translating. In Figure 5, we can see the rough-grained patterns that evolve when highlighting the high modal density for all social actors in the three settings. The figure is based on data from the spreadsheet, as shown in Figure 4. We understand that the higher-level action of sight translation starts when

the yellow narrows into one continuous stretch. There is a shift to new higher-level actions which are opened and closed with the exchange of the frozen action, the document that now guides further actions. At the top and the bottom of the three mid-stretches in the figure, there is a more evenly distributed modal density, representing the higher-level action that I will categorise as *dialogue* in the following.



Figure 5. Density pattern in the higher-level action of giving and receiving information. A (17 min), B (13 min) and C (8 min).

The patterns in the dialogue parts show that density is distributed quite rhythmically/evenly. In the mid-stretch, the PSU and the PSR produce shared higher-level actions with a lower modal density than that produced by the interpreter, and there is little dialogue, although in B and C, we can see that another actor is interacting with the interpreter with a high modal density; this is when the interpreter encounters a challenge (see more detailed transcript in Appendix). The blank breaks in the interpreter column in the sight-translated stretch in situation A represent the interpreter reading in silence.

In the current analysis, the same pattern emerges in all situations (A, B and C); the actors produce similar lower-level actions that constitute a shared higher-level action through the same mediational means, and this is why they can be represented together. There is a pattern in which the modal density is evenly distributed in the dialogue, this is disrupted by sight translation (monologue). During the sight translation, the interpreter controls the mediational means. Analysis at the micro level would reveal the differences

between the three situations; interpreter A, for example, has a more complex modal configuration compared with C, but this is not the focus of this meso-study.

The stable modes that do not change in the three situations are layout and proximity, which foreground the higher-level action of meeting for all the social actors (Figure 1 and 2). In the following, I will present modal density in the dialogue (Figure 6) and the sight-translated part of the meeting (Figure 7) for all three situations together, and I will relate these findings to agency. As explained earlier, Pirini (2017) has developed a tool through MIA to identify primary agency in producing a higher-level action, which is measured through control over the mediational means at play. The actor controlling the mediational means that are most relevant for a higher-level action takes on primary agency, as control influences the possibilities to act and be involved.

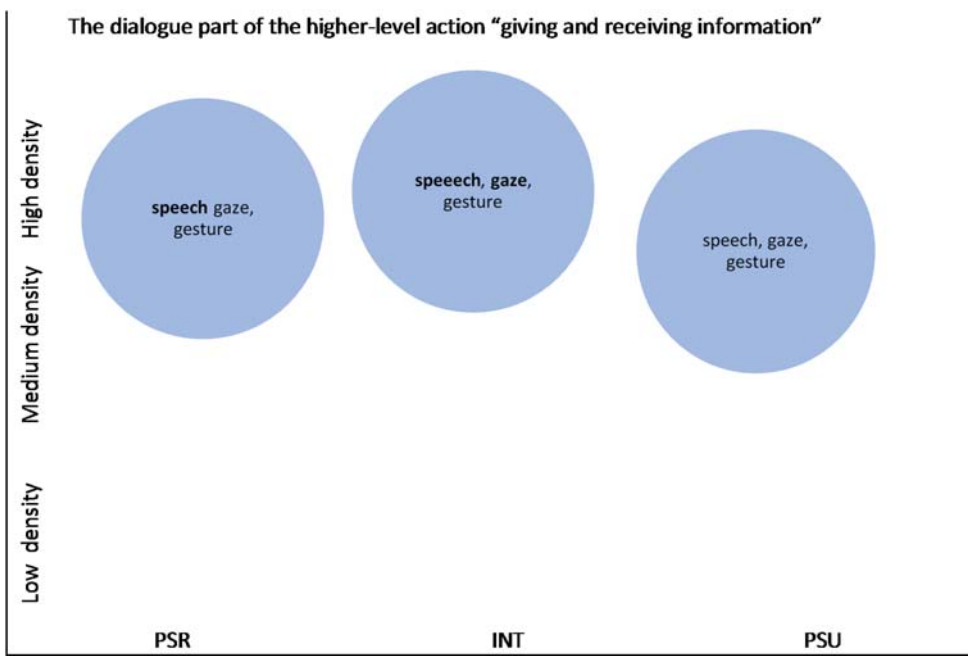


Figure 6. Distribution of modal density in the sight-translated phase.

The modal density is quite evenly distributed, but the intensity of speech, for example, is different (**bold** where the density is higher). The interpreter speaks more (speaking for both), which makes the intensity higher. All social actors foreground the same higher-level action. The PSR leads the meeting through gestures and content of speech. The social actors are related to one another through posture and frequent gaze shifts between the three of them, they mediate their actions through similar mediational means. All in all, in the dialogue, we can say that intersubjectivity is found in all tiers of materiality; stable, adjustable and fleeting, meaning all social actors co-produce this higher-level action through the same means. We shall now see what happens when the PSR gives a written document to the interpreter in Figure 7.

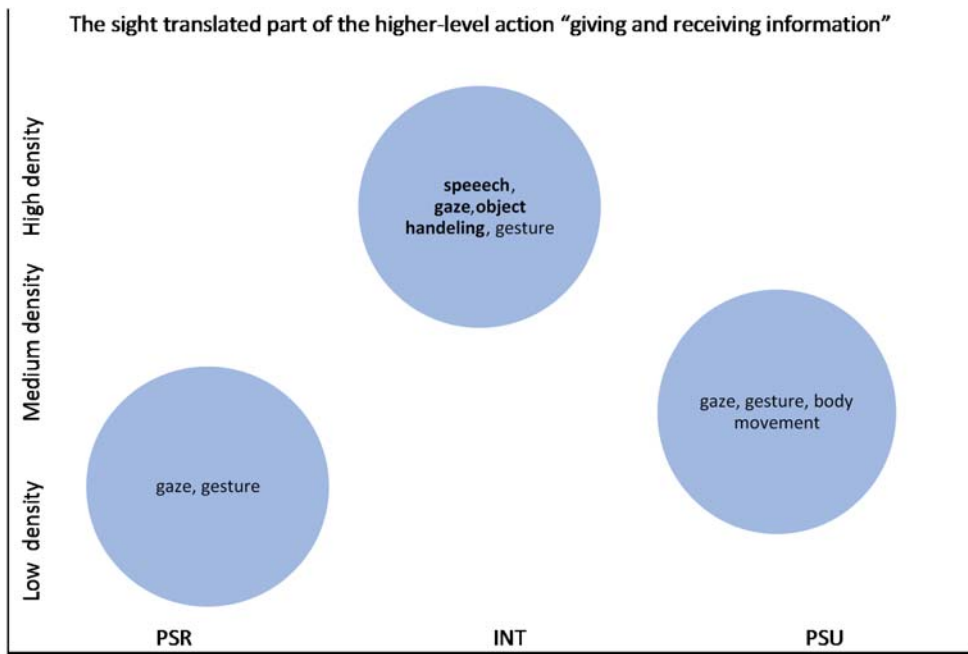


Figure 7. Modal density distribution in dialogue.

Here, we can see that the modal complexity and the modal intensity give a very different picture from that in Figure 6. Both the modal density and the control over the means are focused on the interpreter, which means that the interpreter has primary agency in producing the shared higher-level action. The action is guided by the frozen action, which has now become a mediational mean (object) and a mode (print). Although the PSR and the PSU use similar mediational means, gaze and gestures in the PSU are more often used pragmatically. The PSR is gazing but does not understand the language; however, she is still following the interaction and thereby co-producing the higher-level action of giving and receiving information. Gaze patterns are also affected by the act of reading, as interpreters give primacy to the written text and thereby cannot always react to the other social actors' semiotic means where there are openings for involvement.

The new higher-level actions can be demarcated as sight translation, monitoring and listening. Listening is visible through gaze, nods and expressions of 'mms' as feedback to the interpreter. The PSR's monitoring action is evident through the increased modal density when it looks like the interpreter encountered trouble (hesitations, change in rhythm) and through the turning of the page by the PSR at the same time as the interpreter. Supplementing the analysis with reflections and interviews supports the conclusion that the PSR is monitoring. She commented on one interpreter's explication of the technical abbreviation 'inkluderende arbeidsliv' (IA) in Norwegian, reasoning that the interpreter was trying to figure out how to say it. However, when another interpreter said 'IA' and added a definition, the PSR was not confident that the solution was acceptable (it was adequate). When monitoring the interpreted action without access to the verbal content, the PSR only had fragments of words and the visual to work with, with the latter being proposed to serve as a window into interpreted interaction (Gerwing

& Li, 2019). This is, however, a window without clear sight, so it is a difficult base to act upon and to exercise agency. The intersection of various practices, such as giving information, reading and translating, challenges attention and agency.

Discussion

Attention

As rhythm is an important cohesive device, experiencing a lack of rhythm must affect cognition. When one social actor produces a higher-level action with a higher modal density than the others do, here the interpreter, it seems to create a void or a gap that increases the need for the other participants to do *something*. Involuntary or unidirectional movements, such as jiggling legs, fiddling, rocking and shifting gaze, are typical for people with attention deficits or when they experience cognitive diversion in the same manner as postural changes often indicate some sort of distress (Pirini, 2017). These actions do not necessarily represent a lack of attention, possibly rather a struggle to focus. In this setting, at the meso level, there were no other visible higher-level actions that received midground attention to fill the mentioned gap or void. In the dialogue part of the meeting, when the modal density was more evenly distributed, unidirectional movements were still present; however, they are less salient as the interlocutors were producing similar lower-level actions through the same modes and mediational means in rhythmical alignment with one another, i.e., they share all tiers of material intersubjectivity, stable, adjustable and fleeting (Pirini, 2016), which I propose strengthen each actor's agency.

Attention, in addition to foregrounding a higher-level action, is also related to the unengaged mind – in other words, boredom (Eastwood et al., 2012). Earlier research has mentioned, but not scrutinised, disengagement when listening to sight-translated text (Felberg, 2015; Felberg & Nilsen, 2017). The guidelines for sight translation in asylum hearings cite maintaining attention as being the interpreter's responsibility in order to support listeners' engagement (Spitz & Hlavac, 2017). Eastwood et al. (2012) propose three criteria to define the mental process of boredom: (1) not being able to successfully engage with the internal or external information required to participate, (2) being conscious of the former and (3) attributing the cause of aversiveness to the environment. In short, it is the 'aversive experience of wanting, but being unable, to engage in satisfying activity' (p. 482). Boredom is not at all trivial, the authors argue, as it influences understanding and memory, amongst other things; it thereby creates the potential for agency. In two experiments, the PSUs met the criteria for boredom, with one explicitly using the word 'boring'. In the third experiment, the PSU was not bored with the interpreter, who was faster than the others; interestingly, the interpreter was rocking back and forth. The listener picked up this pace and also started to rock. They were positioned with closer proxemics than the others. This interpreter's modal density was less complex than the others, as the interpreter did not use gaze and gestured less. Rhythm is also related to tempo. It seems that the faster (more rhythmic) interpreter left less space for wandering thoughts compared with the other interpreters; the PSU in this setting did not mention concentration problems and also had less diverted gazes, although his gaze was hardly responded to by the interpreter (or maybe because of that). Despite the rhythm between interpreter C and the PSU, the PSR felt excluded

and was unable to catch any of the other participants' gazes. The interpreter gazed at the document, and the PSU fixed his gaze on the interpreter. As a result, the PSR had little control over the mediational means, which led to a reduction in the PSR's agency. Deciding how and when to contribute in interpreter-mediated interaction is not obvious for the primary interlocutors, especially when practices are not shared.

Agency

The participants did not enter the interaction with the same set of expectations, and this is typical for both institutional and interpreted mediated interaction. In this experiment, the PSR was a trained professional who focused on the PSU. She was engaged, and she controlled her struggle to focus, although she showed signs of restlessness by fidgeting with papers. After the first experiment, which lasted longer, she mentioned that she would prefer to give the PSU the information herself; this would preserve the PSR's agency in the situation by controlling the mediational means. When handing the interpreter, the leaflet, she transfers agency and also foregrounds language in the interaction. Language is further foregrounded by the higher-level action of sight translation, a professional practice that is not an established practice for PSRs, PSUs or interpreters.

The interpreters have different histories embedded in their practice. The less experienced interpreter in sight translation (A) did recognise the lack of routine and felt that some areas could be improved, but overall, 'it went ok'. The other interpreter (B) had extensive experience working with the police and in court and transferred some strategies to this setting, which was unfamiliar to the interpreter. The PSU in situation B commented on the interpreter's formal way of saying 'in parentheses, it says ...' (Havnen, 2020). In police transcripts, parentheses frame comments from the transcriber and are expected to be highlighted; hence, sight translation is not only an interpreting method, but there are also local practices influenced by discourse and text genre.

The PSUs did not have any experience with sight translation, or interpreting, hence they could not experience the practice as an action with history. Moreover, the PSUs had little control over the mediational means and thus could not exercise agency during the sight translation, this lack of control is strengthened by the social practice of reading as readers have primacy over speakers (Scollon, 1998b). Most PSUs will not gain practice in being translated for on a regular basis, and spoken language interpreter users do not constitute a stable user group that can influence and take part in developing practice (sign language interpreter users will probably have more experience than a PSR). There is a need for somebody to act in order to even out this imbalance and to safeguard participants' agency. An obvious actor could be the interpreter, who should be the expert on how the interpreting method influences interaction.

Safeguarding attention and agency

In an institutional setting, agency is, by default, given to the PSR, who has the responsibility to safeguard the participation of the PSU. This responsibility is partly taken over by the interpreter when accepting the translation task. The act of sight translation changes the rhythm of the typical spoken dialogue and this shift challenges the interlocutor's attention and disrupts the turn-taking pattern. The interpreter becomes the social actor who must

take responsibility for balancing the situation in such a way that the interlocutors obtain the opportunity to exercise their agency in the interaction. There are many possible ways to safeguard agency, and since human interaction is always situated, the decisions must be made on the spot by analysing all actions. The interpreter cannot monitor only talk or text, which is the traditional focus; the interpreter needs to pay attention to the whole multimodal communicative environment and the effect of interpreting on interaction.

Professional interpreters work together with other professionals and individuals who base their actions on knowledge different from that of the interpreter. Therefore, interpreters are advised to frame their role in the interaction for the interlocutors at the beginning of a meeting. Sight translation usually occurs in an already ongoing interaction; consequently, it is necessary to frame sight translation in a similar manner before conducting it. How to act when being sight translated for is far from being self-evident or natural; the shift to a new method requires negotiation about communicative preferences, provided that the interpreter understands the other participants' challenges. To know this, interpreters need evidence-based knowledge about the methods they use, the effect it has on interactions and what is at stake. This knowledge is not only required to frame the interaction but to assess situations so that interpreters can choose the best course of action. Knowledge enhances professionals' phronesis – the ability to make decisions in particular situations (Kessels & Korthagen, 2001). My study of sight-translated interaction, with its limitations, contributes novel knowledge about attention and agency that is highly relevant in interpreting education and in professional practice.

CA on reading aloud practices in schools has shown that interaction and attention are affected by several elements, such as whether the pupils have the text in front of them and the degree to which the teacher coordinates via embodied resources and explications (Tainio & Slotte, 2017). Interpreters can coordinate attention strategically by using gaze and/or explicitly asking the listener if the speed is adequate or if the reading is understandable. Such metacommunication enhances agency for the participants who do not control the mediational means. The interpreter can encourage notetaking, maybe asking for an additional copy of the text for the listener or being open for questions during the translation. The interpreter can also encourage/influence (possible) turn-taking through pauses, gaze or prosody, which are common strategies in dialogue interpreting. These strategies might foster involvement and agency.

Interpreters should also be aware that the *excluded* party will make assumptions about the interpreter's actions, for example, when exhibiting trouble. It could be reassuring for the participants to know what the challenge was and to be given the opportunity to provide additional explanations.

As the interlocutors lack a shared practice, they do not have a shared interactional pattern in which they can easily align with one another. Therefore, the interpreter must strive to balance this lack of rhythm by redistributing interactional space and agency.

Conclusion

Through this analysis, I have shown that the act of translating a written text in face-to-face interaction affects the rhythmical patterns of typical interpreter-mediated dialogues. The most salient finding is how the interpreter ends up having primary agency during the sight translation. The interpreter needs to be aware of this and ensure that the participants

get the opportunity to participate, i.e., exercise agency. I argue that this is a responsibility for the interpreter because the shift in pattern is related to the change in the interpreting method.

By exploring the multimodal (inter)actional analytical framework, this study opens avenues for additional and interesting studies of interpreter-mediated interaction, especially those related to the distribution of attention and agency. Interpreters and other social actors' actions beyond the linguistic contribution become very clear when applying MIA. To challenge the results of this study, I am inspired to conduct further investigations of smaller scales of higher-level actions in the material.

This study is not without limitations. If the PSR was not instructed to remain passive when not addressed, she might have intervened. This would show a different practice but would still not undermine the practices scrutinised here. It would be possible to utilise the same framework in order to provide complementary insights; this can be done by repetition of the experiment with different variables. If I were to repeat the experiment, I would also interview all the other participants to gain more exhaustive reflections about their experiences. In the present experiment, I prioritised immediacy. Despite the written reflections being short, they highlight important aspects of the sight-translated interaction, especially in relation to the lack of shared practice.

The variations in sight translation practices that emerged in this study are not exhaustive. Previous studies have shown that PSRs can withdraw totally (Felberg, 2015; Nilsen & Havnen, 2019), the interpreters can align themselves more closely with the PSU than with the other actors (Felberg & Nilsen, 2017; Vargas-Urpi, 2019) and the document can be used more actively by all interlocutors (Davitti & Pasquandrea, 2017; Ticca & Traverso, 2017). This study, however, corroborates previous findings that sight translation affects interactional patterns. The analysis provides new insights into how social actors' attention and agency are affected by the translation method, and it shows that MIA is useful for investigating interpreter-mediated interaction.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Randi Havnen, is a PhD candidate and an assistant professor at Oslo Metropolitan University since 2015. She has over 25 years experience in the field of public service interpreting - as an interpreter and as an interpreter trainer in various institutions. She worked several years for the Directorate of Immigration in the Asylum Department on interpreting issues. She is also a qualified school teacher (Norwegian, Social Science and History). Her MA is from the University of Oslo in Bosnian/Croatian/Serbian (language, literature and area studies).

ORCID

Randi Havnen  <http://orcid.org/0000-0002-0051-9123>

References

Biela-Wolonciej, A. (2015). A-vista: New challenges for tailor-made translation types on the example of the recorded sight translation. *Kalbotyra*, 57(57), 30–39. <https://doi.org/10.15388/Klbt.2007.7555>

- Boria, M., & Tomalin, M. (2020). Introduction. In M. Boria, Á Carreres, M. Noriega-Sánchez, & M. Tomalin (Eds.), *Translation and multimodality: Beyond words* (pp. 1–23). Milton, Routledge.
- Čeňková, I. (2015). Sight interpreting/translation. In F. Pöchhacker, N. Grbić, P. Mead, & R. Setton (Eds.), *Routledge encyclopedia of interpreting studies* (pp. 144–153). Routledge, Taylor & Francis Group.
- Chen, W. (2015). Sight translation. In R. J. Holly Mikkelsen (Ed.), *The Routledge handbook of interpreting* (pp. 144–153). Routledge, Taylor & Francis Group.
- Davitti, E. (2019). Methodological exploration of interpreter-mediated interaction: Novel insights from multimodal analysis. *Qualitative Research*, 19(1), 7–29. <https://doi.org/10.1177/1468794118761492>
- Davitti, E., & Pasquandrea, S. (2017). Embodied participation: What multimodal analysis can tell us about interpreter-mediated encounters in pedagogical settings. *Journal of Pragmatics*, 107, 105–128. <https://doi.org/10.1016/j.pragma.2016.04.008>
- Defrancq, B., & Verliefde, S. (2018). Interpreter-mediated drafting of written records in police interviews: A case study. *Target. International Journal of Translation Studies*, 30(2), 212–239. <https://doi.org/10.1075/target.16141.def>
- Dragsted, B., Hansen, I. G., Vandaele, S., & Bastin, G. (2009). Exploring translation and interpreting hybrids. The case of sight translation. *Meta*, 54(3), 588–604. <https://doi.org/10.7202/038317ar>
- Duncan, S. (2018). Lend me your ears: Mass observing contemporary adult reading aloud practices. *Changing English*, 25(4), 410–425. <https://doi.org/10.1080/1358684X.2018.1504620>
- Eastwood, J. D., Frischen, A., Fenske, M. J., & Smilek, D. (2012). The unengaged mind: Defining boredom in terms of attention. *Perspectives on Psychological Science*, 7(5), 482–495. <https://doi.org/10.1177/1745691612456044>
- Felberg, T. R. (2015). Udfordringer ved prima vista talking i asyLintervjuer. *FLEKS: Scandinavian Journal of Intercultural Theory and Practice*, 2(1), 1–17. <https://doi.org/10.1016/j.socscimed.2019.05.040>
- Felberg, T. R., & Nilsen, A. B. (2017). Exploring semiotic resources in sight translation. *The Journal of Specialised Translation*, 28, 230–249 Retrieved from: https://jostrans.org/issue28/art_felberg.php.
- Gerwing, J., & Li, S. (2019). Body-oriented gestures as a practitioner’s window into interpreted communication. *Social Science & Medicine*, 233, 171–180. <https://doi.org/10.1016/j.socscimed.2019.05.040>
- Goody, J. (1987). *The interface between the written and the oral*. Cambridge University Press.
- Havnen, R. (2019). Multimodal and interactional aspects of sight translation – A critical review. *FITISPos International Journal - Public Service Interpreting and Translation*, 6(1), 91–106. <https://doi.org/10.37536/FITISPos-IJ.2019.6.1.206>
- Havnen, R. (2020). Where did the footnote go? How the change of mode in sight translation affects meaning-making. *The Journal of Specialised Translation*, 34, 78–99. Retrieved from: https://jostrans.org/issue34/art_havnen.php
- Jones, R. H., & Norris, S. (2005). Introducing agency. In S. Norris, & R. H. Jones (Eds.), *Discourse in action: Introducing mediated discourse analysis* (1st ed., pp. 195–215). Routledge.
- Kaindl, K. (2020). A theoretical framework for a multimodal conception of translation. In M. Boria, Á Carreres, M. Noriega-Sánchez, & M. Tomalin (Eds.), *Translation and multimodality: Beyond words* (pp. 49–70). Milton, Routledge.
- Kessels, J., & Korthagen, F. A. J. (2001). The relation between theory and practice: Back to the classics. In F. Korthagen (Ed.), *Linking practice and theory: The pedagogy of realistic teacher education* (pp. 20–31). Lawrence Erlbaum.
- Krystallidou, D. (2014). Gaze and body orientation as an apparatus for patient inclusion into/exclusion from a patient-centred framework of communication. *The Interpreter and Translator Trainer*, 8(3), 399–417. <https://doi.org/10.1080/1750399X.2014.972033>
- Li, X. (2014). Sight translation as a topic in interpreting research: Progress, problems, and prospects. *Across Languages and Cultures*, 15(1), 67–89. <https://doi.org/10.1556/Acr.15.2014.1.4>

- Maatta, S. K. (2015). Interpreting the discourse of reporting: The case of screening interviews with asylum seekers and police interviews in Finland. *The International Journal for Translation & Interpreting Research*, 7(3), 21–35. <https://doi.org/10.12807/ti.107203.2015.a02>
- Nilsen, A. B., & Havnen, R. (2019). Sight translation: Best practices in health care and in training. In I. E. T. de V. Souza, & E. Fragkou (Eds.), *Handbook of research on medical interpreting* (pp. 375–396). IGI Global.
- Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. Routledge.
- Norris, S. (2011). Tempo, Auftakt, levels of actions, and practice: Rhythm in ordinary interactions. *Journal of Applied Linguistics*, 6(3), 333. <https://doi.org/10.1558/japl.v6i3.333>
- Norris, S. (2019). *Systematically working with multimodal data: Research methods in multimodal discourse analysis*. John Wiley & Sons, Inc.
- Norris, S., & Pirini, J. P. (2017). Communicating knowledge, getting attention, and negotiating disagreement via videoconferencing technology: A multimodal analysis. *Journal of Organizational Knowledge Communication*, 3(1), 23. <https://doi.org/10.7146/jookc.v3i1.23876>
- Paez, B. (2014). Performance criteria descriptors for cognitive processing skills used in sight translating. In A. Arnall, & A. Gentile (Eds.), *AUSIT 2012: Proceedings of the “JubilaTion 25” biennial conference of the Australian institute of interpreters and translators* (pp. 27–38). Cambridge Scholars Publishing.
- Perez-Gonzalez, L. (2014). Multimodality in translation and interpreting studies: Theoretical and methodological perspectives. In S. Berman & C. Porter. (Ed.), *A companion to translation studies* (pp. 119–131). Oxford, UK: John Wiley & Sons, Ltd.
- Pirini, J. (2014). Producing shared attention/awareness in high school tutoring. *Multimodal Communication*, 3(2), 163–179. <https://doi.org/10.1515/mc-2014-0012>
- Pirini, J. (2015). Introduction to multimodal (inter)action analysis. In S. Norris, & C. D. Maier (Eds.), *Interactions, images and texts: A reader in multimodality* (pp. 77–92). De Gruyter Mouton.
- Pirini, J. (2016). Intersubjectivity and materiality: A multimodal perspective. *Multimodal Communication*, 5(1), 1–14. <https://doi.org/10.1515/mc-2016-0003>
- Pirini, J. (2017). Agency and co-production: A multimodal perspective. *Multimodal Communication*, 6(2), 109–128. <https://doi.org/10.1515/mc-2016-0027>
- Scollon, R. (1998a). *Mediated discourse as social interaction: A study of news discourse*. Routledge.
- Scollon, R. (1998b). Reading as social interaction: The empirical grounding of reading. *Semiotica*, 118(3/4), 281–294. <https://doi.org/10.1515/semi.1998.118.3-4.281>
- Spitz, K., & Hlavac, J. (2017). Sight translating interview transcript. In U. Austria (Ed.), *Handbook for interpreters in asylum procedures* (pp. 121–136). UNHCR.
- Tainio, L., & Slotte, A. (2017). Interactional organization and pedagogic aims of reading aloud practices in L1 education. *Nordic Journal of Literacy Research*, 3(0), 1–22. <https://doi.org/10.23865/njlr.v3.469>
- Ticca, A. C., & Traverso, V. (2017). Participation in bilingual interactions: Translating, interpreting and mediating documents in a French social centre. *Journal of Pragmatics*, 107, 129–146. <https://doi.org/10.1016/j.pragma.2016.09.008>
- Tuominen, T., Hurtado, C. J., & Ketola, A. (2018). Why methods matter: Approaching multimodality in translation research. *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 17, 1–21. Retrieved from <https://lans.ua.ac.be/index.php/LANS-TTS/article/view/522>
- Van Leeuwen, T. (2005). *Introducing social semiotics*. Routledge.
- Vargas-Urpi, M. (2019). Sight translation in public service interpreting: A dyadic or triadic exchange? *The Interpreter and Translator Trainer*, 13(2), 1–17. <https://doi.org/10.1080/1750399X.2018.1503834>
- Vranješ, J., Brône, G., & Feyaerts, K. (2018). Dual feedback in interpreter-mediated interactions: On the role of gaze in the production of listener responses. *Journal of Pragmatics*, 134, 15–30. <https://doi.org/10.1016/j.pragma.2018.06.002>
- Wadensjö, C. (1998). *Interpreting as interaction*. Longman.
- Wertsch, J. V. (1998). *Mind as action*. Oxford University Press.