Katholieke Universiteit Leuven
Campus Antwerpen
Subfaculteit Letteren

On the role of eye gaze in the coordination of interpreter-mediated interactions

an eye-tracking study

Proefschrift ingediend tot het behalen van de graad van

Doctor in de Taalkunde

door

Jelena Vranjes

Promotor: Prof. dr. Geert Brône
Co-promotoren: Prof. dr. Kurt Feyaerts, Dr. Hanneke Bot

2018

Promotiecommissie

Promotor:	Prof. dr. Geert Brône
Co-promotoren:	Prof. dr. Kurt Feyaerts & Dr. Hanneke Bot
Leden:	Prof. dr. Peter Auer
	Prof. dr. Franz Pöchhacker
	Prof. dr. Heidi Salaets
	Prof. dr. Dorien Van De Mieroop
Voorzitter:	Prof. dr. Elke Peters

The research reported in this thesis was supported by the Louis Meerts Chair for Specialized Communication, KU Leuven.

Za moju ljubav

Acknowledgments

First and foremost, I want to express my gratitude to my supervisors. Thank you for lifting up my mindset and for taking me to new dimensions of linguistic research. Geert, it was a great privilege to have you as a supervisor. Thank you for introducing me to eye-tracking and for being a constant source of encouragement and inspiration throughout this process. I also want to thank you for being a patient and attentive mentor, and for always making the time for me when I needed your advice on any issues I had. Kurt, it happened on numerous occassions that just one remark you uttered, just one reference you named, was enough to redirect my thoughts in a new direction. Thank you for your trust, enthusiasm and your wit which contributed much to the warmth and liveliness of our meetings. I have also had the fortune to be supervised by Hanneke Bot. Hanneke, your work has greatly inspired the way I approached this topic. Your directness, critical mind and deep understanding of the practical issues of working with interpreters have been a constant reminder that theory is nice, but practice is better..

This work has been enriched by the input of many colleagues and scholars whom I had to chance to discuss my work with. I am grateful to Elwys De Stefani for his feedback and insightful comments on some of my work. In the last year of my PhD, I also had the privilege to spend one month at the Centre of Excellence on Intersubjectivity in Interaction in Helsinki. I would like to thank Jan Linström, Melissa Stevanovic, Anssi Peräkylä, and Tuija Kinnunen for the opportunity to take part in data sessions and to present my ongoing work in several seminars. I am also thankful to Katariina Harjunpää, Giovanni Rossi and Aurora Guxholli for their warm welcome and fruitful discussions.

Next, I would like to thank my colleagues, and especially Annelies Jehoul and Inez Beukeleers for the good times at our office and for their help during the recordings of my dataset. I am also grateful to Stijn De Beugher for his invaluable technical support

during the first stages of my data collection. Many thanks go to Isabelle Heyerick for her encouragements and for reminding me of the deadlines for the administrative parts of this PhD. I extend my thanks to all the other colleagues who have, in one way or another, contributed to this journey: Bert Oben, Paul Sambre, Liesbeth Van Vossel, Roel Coesemans and Charlotte Maekelberghe. Finally, I would like to thank my dear friends for all their support and unforgettable moments during this journey: Marija Djeranović, Milana Marić, Gvozden Dušanić, Sandrine Debecker, Margot D'Hertefelt, Johanna Laurent, Darya Holemans, Laura Uwase, Marko Topalović and Marija Stojiljković. My talented friend Marija Djeranović deserves special thanks for making the cover illustration of this dissertation. Sincere thanks are also due to all the people who agreed to participate in this study and who made this work possible.

Most importantly, I want to thank my parents, the most selfless people I know and my biggest heroes. Thank you for being a constant source of inspiration, love and support in my life. I am also grateful to my sister Ivana for her help with some of the statistical analyses presented in this dissertation. Most importantly, I want to thank her for helping me to always put things into perspective and for being the wonderful person she is.

Above all, I want to thank my love Tibor. *Ljubavi*, your support over the past four years, and especially over the past months has been invaluable to me. Thank you for your constant encouragements, words of wisdom and for your rigorous reading of the last versions of this text. Most of all, thank you for your limitless love throughout this journey.

Contents

Acknowledgments i

Contents iii

Introduction 2

Chapter 1. Literature overview 6

1.1. Defining the object of study 6

1.1.1. Interpreter-mediated interaction as 'communicative pas de trois' 7

1.1.2. Interpreter's participatory role 9

1.2. The role of gaze in interaction 11

1.2.1. Gaze and participation 12

1.2.2. Regulatory functions of gaze 13

1.2.3. Gaze as a social act 14

1.3. The role of gaze in interpreter-mediated interaction 15

1.4. Multimodality and social interaction 19

1.4.1. Sequentiality 19

1.4.2. Simultaneity 20

1.4.3. Reciprocity 21

1.5. Present study 22

Chapter 2. Data and method 24

2.1. Methodological considerations 24

2.2. Participants and corpus 25

2.3. Seating arrangements 27

2.4. Recording set-up and technical equipment 29

2.5. Recording procedure 32

2.6. Data processing 32

2.7. Notes on the multimodal transcription conventions of gaze 36

Chapter 3. Distribution of gaze orientation 40

3.1. Gaze and participation in face-to-face talk 40

3.2. Gaze and participation in interpreter-mediated talk 42

3.3. Primary participants' gaze while speaking 44

3.3.1. Counsellor's gaze while speaking 44

3.3.2. Student's gaze while speaking 46

3.4. Primary participants' gaze while listening 47

3.4.1. Counsellor's gaze while listening 48

3.4.2. Student's gaze while listening 50

3.5. Interpreter's gaze orientation 53

3.5.1. Interpreter's gaze while speaking 53

3.5.2. Interpreter's gaze while listening 54

3.6. Conclusion 56

Chapter 4. Gaze and turn-taking 58

4.1. Turn-taking and gaze in face-to-face interaction 59

4.2. Turn-taking in interpreter-mediated interaction 61

4.3. Extended tellings 64

4.4. Analysis 68

4.4.1. First-pair parts of adjacency pairs: gaze at the selected speaker 68

4.4.2. Chunking of multi-unit turns 70

4.4.2.1. Immediate gaze shift at turn-end: close collaboration 70

4.4.2.2. No gaze shift at turn-end: incongruity in action formation 78

4.4.2.2.1. The selected participant does not take the turn 78

4.4.2.2.2. The selected participant takes the turn 83

4.5. Conclusion 88

Chapter 5. Gaze and backchannel responses 90

5.1. Backchannel responses in face-to-face interaction 91

5.1.1. Defining the phenomenon 91

5.1.2. Backchannel responses and grounding 95

5.1.3. Backchannel responses and multimodality 96

5.1.4. Backchannel responses within a multiperson participation framework 98

5.2. Backchannel responses in interpreter-mediated interaction 99

5.2.1. Interpreter's participation status 99

5.2.2. Lack of direct communicative contact 103

5.3. Analysis 108

5.3.1. Distribution of backchannel responses in the dataset 108

5.3.2. Backchannel responses and speaker's gaze 113

5.3.3. Directionality of backchannel responses 116

5.3.3.1. Interpreter-directed backchannel responses 117

5.3.3.2. Primary-directed backchannel responses 118

5.3.4. Dual feedback 121

5.3.4.1. The phenomenon 121

5.3.4.2. Backchannel responses and gaze shifts in turn-medial position 124

5.3.4.3. Dual feedback and triadic participation framework 127

5.3.4.4. Receipting newsworthy information 129

5.3.4.5. Displaying agreement or affiliation 136

5.4. Conclusion 141

Chapter 6. A case study in psychotherapeutic context 144

6.1. Presenting the case 144

6.2. The setting: interpreting in psychotherapy 146

6.3. Distribution of backchannel responses 148

6.4. Mutual gaze and interpreter's backchannel responses 149

6.4.1. Gaze as response inviting cue 149

6.4.2. Sustained mutual gaze and increased affiliative strength 153

6.5. Therapist's dual feedback 155

6.6. Conclusion 159

Chapter 7. Conclusion 162

7.1. Summary of the findings 162

7.2. Implications for research and practice 165

7.3. Limitations and suggestions for future research 166

Appendix 168

Appendix A: Transcription conventions 168

Appendix B: Specifications of the technical equipment 170

Appendix C: Background information about the interpreters 171

Appendix D: Coding scheme for backchannels 172

Appendix D: Distribution of gaze orientations in Chapter 6 173

References 176

Samenvatting 194

Introduction

> *The union and interaction of individuals is based upon mutual glances. This is perhaps the most direct and purest reciprocity which exists anywhere*. (Simmel 1920: 358)

This work focuses on the role of gaze direction in the interactive organization of interpreter-mediated interactions. Traditionally, the interpreting process was considered as the production of 'texts' and the interpreter as a 'linguistic conduit' who renders messages from one language into the other. More recently, detailed examinations of actual interpreter-mediated encounters have contributed to a view of the interpreting process as 'interaction' and the interpreter as an *active* participant within the exchange (Wadensjö 1998, Davidson 2002). Moreover, since participants in interpreter-mediated interactions have asymmetrical access to each other's language, it is likely that visual resources will play an important role in this type of encounters (Pasquandrea 2011). Indeed, studies have shown that participants in interpreter-mediated encounters collaboratively manage their talk through both verbal and visual resources (such as gesture, posture and gaze direction) (Pasquandrea 2011, Mason 2012). Especially gaze direction is starting to attract increasing interest as an important resource in signaling conversational attention and in negotiating participation (Bot 2005, Mason 2012, Davitti & Pasquandrea 2017). Yet, it is unclear *how* gaze and other modalities interact and contribute to the management of interpreter-mediated encounters (Mason 2012). Research in this field has been hampered for a long time by the difficulties in collecting video data that would allow a systematic study of gaze in ongoing interpreter-mediated encounters. With this dissertation, I hope to contribute to the research on the organization of gaze in the context of face-to-face interpreter-mediated interaction.

By adopting an interdisciplinary approach that combines insights from Interpreting Studies, (multimodal) Conversation Analysis (CA) and joint action theory, I focus on two key phenomena that are manifest in any type of interaction and that are typically realized multimodally, i.e. through the tight interaction of different semiotic resources including speech, gesture, and gaze (Goodwin 1981, Bavelas et al. 2002, Rossano 2012, Brône et al. 2017, Auer 2017):

(a) turn-taking, or the participants' regulation of speaking turns (Sacks et al. 1974).

(b) backchannel responses, such as *mm hm*, *yeah* and head nods, which build a central mechanism for the establishment of reciprocity, rapport and common ground between the interlocutors (Clark & Brennan 1991, Gardner 2001, Bavelas et al. 2002);

I will relate these phenomena to the management of the dynamic participation framework and to the notion of 'speakership' in this particular setting. Moreover, by revealing some systematic practices in the use of gaze in relation to these two phenomena, this study will help answer some basic questions concerning the coordinating role of the interpreter in the interaction and the primary participants' orientations towards the interpreter. The analyses will draw on a dataset of interpreter-mediated interactions (Dutch-Russian) that were recorded with mobile eye-tracking glasses. This innovative approach allows for a highly detailed analysis of interlocutors' gaze behavior in relation to speech and other modalities in ongoing interaction. Consequently, this dissertation responds to Mason's (2012) recent call for an in-depth exploration of the concurrent interaction between speech, gesture and gaze "of all participants" in interpreter-mediated talk (p.182).

The study is structured as follows: **Chapter 1** starts with the introduction of triadic, interpreter-mediated interaction as focus of analysis. This is followed by a brief theoretical discussion on the multimodal nature of face-to-face interaction, characterized by the notions of sequentiality, simultaneity and reciprocity.

Chapter 2 presents the dataset used for this study, namely interpreter-mediated encounters (Dutch-Russian) in a university setting, which were recorded with mobile eye-tracking glasses. It discusses the methodological preliminaries associated with the use of eye-tracking in the context of interpreter-mediated interactions, how the data generated by eye-trackers were processed and analyzed.

In order to get a sense of the general visual dynamics in the dataset, **Chapter 3** provides a quantitative analysis of the participants' gaze distribution during these encounters. It shows how participants' participation status during the talk (as speaker or listener) and the overarching asymmetries in social positions (relative power, social distance, professional status) affect the employment of gaze in this type of encounters. Moreover, it sheds new light on some of the previous claims in relation to the organization of gaze in face-to-face (interpreter-mediated) interaction.

Chapter 4 examines the organization of turn-taking in interpreter-mediated interaction. It starts with a brief literature review on the turn-taking 'machinery' as described by Sacks, Schegloff and Jefferson in their seminal study from 1974, followed by a discussion on the role of gaze in this process. More specifically, the chapter focuses on the interpreter's role in the stepwise production of extended multi-unit turns (such as long descriptions and story tellings) in interpreter-mediated talk. In this process, a multi-unit turn is 'chunked' into smaller units in order to give the interpreter the opportunity to render the talk piecemeal. Responding to recent discussions by Rossano (2012) and Auer (2017) on the role of gaze for turn allocation, I will argue that the interpreter's gaze direction contributes both to the local management of turn-taking (regulation of turn transitions and next-speaker selection) and to the accomplishment of the action in progress. Moreover, these findings will allow me to discuss the interpreter's 'speakership' in the interpreting process.

Chapter 5 presents a detailed analysis of the backchanneling behavior of the participants in the dataset. Given the interpreter's particular position within the exchange, how does the interpreter take part in the interaction as listener? Moreover, how do primary participants orient towards the interpeter while listening to the latter's rendition? After presenting a literature overview on the use and function of backchannel

responses in human interaction, I discuss how the specific nature of interpreter-mediated interactions affects the participants' production of backchannel responses. The analytical section of the chapter is divided into three parts: the first part provides a quantitative analysis of the participants' production of backchannel responses in the dataset. The second part examines the *inter*-personal coordination of backchannel responses, and more specifically the relationship between interlocutors' mutual gaze and the production of backchannel responses in the dataset, thus extending previous research on this topic from a two-person (Bavelas et al. 2002) to a three-person participation framework. The third, and most extensive part of the analysis focuses on the the *intra*-personal coordination of gaze and other verbal and visual modalities in the production of backchannel responses. In response to the general view of interpreter-mediated talk as two dialogues that are interconnected through the interpreter's interventions (Davidson 2002), this chapter shows how (verbal and visual) backchannel responses and gaze together constitute 'composite signals', employed to maintain a triadic participation framework during the interaction. Moreover, through fine-grained qualitative analyses of conversational data, I will demonstrate how listeners, by shifting their gaze from one interlocutor to the other while producing backchannel responses, display their orientation to the differences in the participation status or 'speakership' of the interlocutors.

In order to respond to the question of the general applicability of the current approach and the generalizability of the reported findings, **Chapter 6** presents a case study on the role of gaze in the production of backchannel responses in a naturally occurring interpreter-mediated therapeutic encounter that was recorded by using mobile eye-tracking technology. The case study reveals some patterns that were observed in the previous chapters, while delving deeper into the specificities of this setting.

Finally, **Conclusion** presents the overall implications of the present study for our understanding of interpreter-mediated interaction and for the study of gaze in face-to-face talk. In addition, it examines the implications of this study for the interpreting practice and for the novel forms of (remote) interpreting. It concludes by addressing the limitations of the current approach and by providing suggestions for future research.

Chapter 1

1.1. Defining the object of study

Last decade has witnessed an increase in demand for interpreting services around the world. Especially with the rapidly growing globalization and international migration, interpreters are employed more and more to enable communication between people who do not speak each other's language (Tebble 2012). The variety of settings in which interpreters are employed include medical consultations, social welfare interviews, parent-teacher meetings, immigration hearings, business encounters, police interviews and other social contexts. Parallel with the growing need for interpreting services, there has been an increasing scholarly interest in the interactional organization of such events: How do participants achieve mutual understanding with the aid of an interpreter? What is the level of the interpreter's involvement within the exchange? How does interpreting affect the interlocutors' participation during the talk? Within the field of Interpreting Studies, interpreter-mediated interactions are typically researched under the designation 'dialogue interpreting' (Mason 1999, Baraldo & Gavioli 2012). According to Mason (2009), dialogue interpreting has four fundamental characteristics: (i) dialogue, entailing bi-directional translation, (ii) spontaneous speech, (iii) face-to-face exchange, and (iv) consecutive interpreting mode (i.e. after the speaking turns of the primary participants)[1].

[1] Mason's definition, however, does not take into account signed language interpreting, which is performed in a simultaneous interpreting mode, nor does it include remote interpreting (e.g. via the telephone) (Merlini 2015). Furthermore, Pöchhacker (2012) points out that "the notion of dialogue interpreting should not be understood as referring only to one-on-one interactions", since "the constellation of interactants in an event may of course be much more complex" (p. 46).

Most importantly, what distinguishes interpreter-mediated interaction from "monologue-based communication' that is inherent in conference interpreting settings, is the *interaction* itself (Merlini 2015). The present study focuses on the prototypical, face-to-face constellation of three-person interaction between two primary participants, who have no understanding of each other's language, and an interpreter, who renders the primary participants' utterances in both languages in consecutive turns.

1.1.1. Interpreter-mediated interaction as a 'communicative pas de trois'

Traditionally, the view of interpreting was determined by the 'transfer' model of communication (see e.g. Reddy 1979), in which the interpreting process was conceived as the production of 'texts' and the interpreter as a mere linguistic 'conduit' or translating machine (Wadensjö 1998, Bolden 2000, Davidson 2002, Bot 2005). This model starts from the idea of an "invisible interpreter" (Wadensjö 1998) who provides precise and complete renditions of everything that is being said, without engaging meaningfully in the conversation (Davidson 2002). Although still present in many professional codes of conduct for interpreters (Angelelli 2004)[2], recent contributions within the domain of Interpreting Studies have shown that such theoretical conceptualization of the interpreting process does not correspond to the intrinsically complex reality of an interpreted event (e.g. Wadensjö 1998, Angelelli 2000, 2004; Davidson 2002, Bot 2005, Mason 2012, Pöchhacker 2012, Davitti 2013). Detailed analyses and observations of actual interpreter-mediated exchanges have shown that interpreters are faced not only with the cognitively demanding task of listening, processing and translating the ongoing talk, but also with the "added constant (re-) negotiation of role, turn management and general monitoring of the unfolding of the talk exchange" (Mason 1999: i). Furthermore, primary participants may have to adapt their behavior to the interactionally atypical situation of conversing with the aid of an interpreter (see e.g. Bot 2005). Far from being a "trivially modified version of same-

[2] For instance, the professional Code of Conduct for community interpreters in Belgium states that an interpreter "always remains the interpreter and therefore does not take part in the conversations at any moment " (http://www.agii.be/sites/default/files/bestanden/deontologische_code_sociaal_tolken.pdf)

language [talk]" (Davidson 2002), interpreter-mediated interaction warrants an investigation of the complex nature of such communicative events.

In her seminal work *Interpreting as Interaction* (1998), Wadensjö proposed a new direction for research on interpreter-mediated exchanges "based on a *dialogic*, rather than *monologic* view on language and language use" (p.7). She argued that, in order to understand the conversational dynamics of interpreted encounters, it is more useful to look at them as joint activity, or 'communicative *pas de trois'*, rather than production of texts. This implies that, within such a complex activity, coordination and collaboration between participants - interpreter included - is key for a successful accomplishment of the joint action.

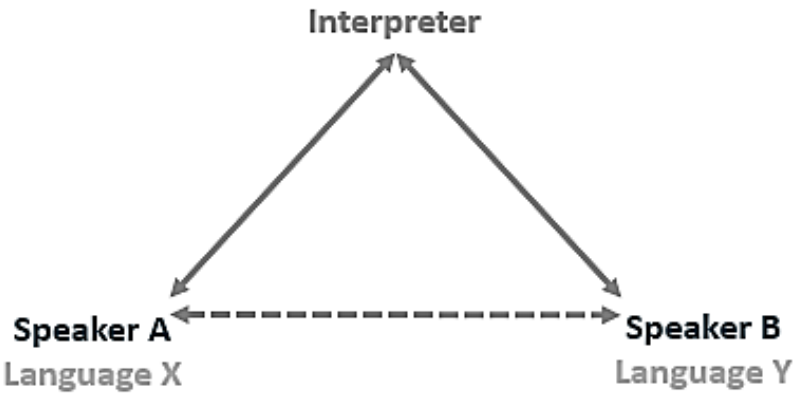


Figure 1.1: Interpreter-mediated interaction as triadic exchange.

A conceptualization of interpreter-mediated interaction as a 'communicative *pas de trois*' leads to the following basic questions: within the 'communicative triangle' constituted by the two primary participants and the interpreter, what are the responsibilities and contributions of the participants within their respective positions? How do they relate to each other and to the ongoing talk? What contributes to a successful accomplishment of joint action? Over the past two decades, a considerable body of empirical studies has analyzed micro-level processes through which the primary participants and the interpreter collaborate in the construction of social action (Bolden 2000, Davidson 2002, Bot 2005, Pasquandrea 2011, Gavioli 2012, Baraldi & Gavioli 2012, Davitti 2013, Raymond 2014). Given the interpreter's central role in establishing mutual understanding between primary participants and high level of responsibility for

the overall success of the exchange, most research has focused on the interpreter's participatory role and level of involvement within the encounter, as will be discussed below.

1.1.2. Interpreter's participatory role

The interpreter's distinct role within the interaction has been the main theme in research on interpreter-mediated interaction over the past two decades (Mason 2012). Rather than just rendering the primary participants' utterances as a 'linguistic parrot', the interpreter is an active participant who is critically engaged in the negotiation of meaning and in the coordination of the interaction (Englund Dimitrova 1997, Wadensjö 1998, Bolden 2000, Angelelli 2004, Bot 2005, Davidson 2002, Gavioli 2007, 2012; Van De Mieroop & Mazeland 2009, Pasquandrea 2011, Mason 2012, Pöchhacker 2012, Davitti 2013, Raymond 2014). As stated by Pöchhacker, paraphrasing Watzlawick et al. (1967), "the interpreter cannot *not* participate" (2012: 50); the question being how and at what level does the interpreter contribute to the communicative event.

Bolden (2000), for example, showed through detailed examination of the history-taking part of medical consultations that interpreters may initiate monolingual dyadic sequences with the patient in order to pursue diagnostically relevant issues. Bolden concluded that "interpreter's actions are primarily structured by their *understanding* of the ongoing activity and only secondarily by the task of translation" (Bolden 2000: 347, my emphasis). These findings were supported by Davidson (2002) in his 'collaborative model' of interpreted discourse in which the interpreter has the central role in the organization of turn-taking and the construction of conversational common ground within the exchange. His model provides space for 'mini-dialogues' between the interpreter and each of the primary participants (Speaker 1 and Speaker 2 in Figure 1.2 below), which allow the interpreter to ensure comprehension with each of the primary participants.

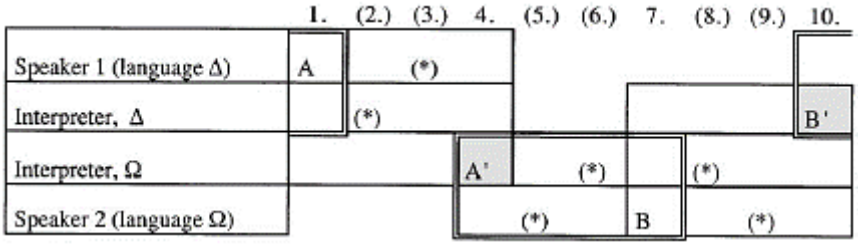


Figure 1.2: Davidson's (2002) 'collaborative' model of interpreted discourse.

Apart from the asymmetries in language knowledge, interpreter-mediated encounters usually take place in institutional settings that are characterized by asymmetries in knowledge or power relations between professionals (e.g. teachers, therapists) and nonprofessionals (e.g. patients, students), between more 'powerful' and 'weaker' participants. An interpreter can sustain or balance out those asymmetrical relationships (Mason 2009, Raymond 2014). For example, by suspending translation provision and giving the 'weaker' participant the opportunity to come out with his/her telling, the interpreter can promote that participant's position within the exchange (Gavioli 2012). This means that the primary participants unavoidably share control over the communicative event with the present interpreter (Davidson 2002). In sum, the outcome of an interpreter-mediated interaction "is the result of a subtle combination and balancing out of strategies adopted and choices made by the interpreter and by the participants involved in a constant process of interactional negotiation" (Davitti 2012: 24). The question is thus no longer "whether" but in what measure and how interpreters engage in interpreter-mediated interaction (Davitti 2012).

Notwithstanding these observations, our understanding of interpreter-mediated events "still seems to be colored by the traditionally dominant understanding of interpreters, answering to a similar orientation to all participants in interpreter-mediated encounters, first and foremost as producers of talk" (Wadensjö 2015). Interpreters are, in the words of Davitti,

> "constantly confronted with a difficult balancing act, i.e. to what extent should they identify with or distance themselves from one party or the other, where is the boundary to be traced, what are their tasks and socio-epistemic rights and responsibilities in interaction" (Davitti 2012: 173)

On the other hand, primary participants are faced with the difficult issue how they should orient towards the interpreter (Wadensjö 1998, Davitti 2012). Thus, whereas theory stresses the importance of the interpreter's 'invisibility', the interpreter's involvement within the exchange is characterized by a high level of responsibility and *visibility*. The latter notion calls for the integration of interactionally relevant communicative resources such as gaze, head movements, facial expressions and gesture, "layers that largely remain to be systematically explored in studies in interpreter-mediated interaction" (Wadensjö 2015). Interpreter-mediated interaction can benefit from such an approach, since

> "by its own nature, it is a multiparty interaction, involving a complex participation format that is likely to be managed multimodally, rather than merely verbally. In addition, it takes place in a multilingual environment, in which the asymmetrical access of the participants to each other's language limits the amount of verbal communication they are able to employ" (p. 456).

One communicative resource, that is of particular importance for human communication and that has fascinated scholars from various disciplines for decades, is gaze. This dissertation will focus on the role of gaze as a continuous and dynamic visual signal within the specific setting of face-to-face interpreter-mediated encounters.

1.2. The role of gaze in interaction

There has been a long-standing interest in the role of gaze in social interaction (for an overview, see Rossano 2013). Apart from allowing us to perceive the world around us, gaze fulfills a whole range of communicative functions in interaction. As pointed out by Goodwin (1981: 30) "gaze is not simply a means of obtaining information, the receiving end of a communicative system, but is itself a social act". Early studies on gaze focused

primarily on the role of gaze in displaying and maintaining engagement with our interlocutors during the talk (Goffman 1964, Nielssen 1962). In his pioneering study of gaze direction in social interaction, Kendon (1967) identified specific patterns associated with the interlocutors' participation roles and found evidence for the regulatory functions of gaze in interaction. Subsequent research aimed at separating out various functions of gaze, including the display of participation, signaling of understanding, (dis)agreement, affiliation and interpersonal relationships, and regulation of speaking turns (e.g. Argyle & Cook 1976, Beattie 1978, Goodwin 1980, 1981; Heath 1984, Bavelas et al. 2002, Lerner 2003, Haddington 2006, Rossano et al. 2009, 2012; Auer 2017, Brône et al. 2017). In a recent overview of research on gaze in social interaction, Rossano (2013) presents the following three topics that have received most scholarly attention (especially within a conversation analytic-tradition): (1) gaze and participation, (2) regulatory functions of gaze and (3) gaze as a social act. I will briefly present each of these topics in the following sections.

1.2.1. Gaze and participation

There has been a lot of interest in the differences in gaze behavior of the interlocutors in relation to their participation roles (i.e. being the speaker or listener) in the exchange. Studies have found that interlocutors tend to gaze more at the other participant when listening than when speaking[3] (Kendon 1967, Argyle & Cook 1976, Goodwin 1981, Bavelas et al. 2002). Goodwin (1981) showed through detailed analyses of interactional data how participants display sensitivity to the (lack of) gaze of their interlocutors during the talk and proposed a set of 'rules' that guide participants' gaze behavior (Goodwin 1981, see also Chapter 3). Heath (1984) demonstrated how gaze is used by the listener as a 'display of recipiency' and availability for the ongoing activity. Rossano (2012, 2013), on the other hand, argued that interlocutors' gaze is not only determined by their

[3] Rossano (2013) noted that normative gaze patterns during listening versus speaking may be influenced by factors such as race and cultural differences. For example, he shows how within a Mayan community in Mexico people do not look at speakers' faces when listening to them (p.310).

participation role, but also by the type of social action that the participants are engaged in (e.g. longer tellings versus turn-by-turn talk).

Furthermore, in interactions involving three or more participants, one can distinguish between different 'types' of listeners. An important concept in this respect is the notion of '**participation framework**'. This notion refers, in its original sense (Goffman 1981), to the different positions that hearers can occupy while listening to an utterance, as either 'ratified' or 'unratified' hearers. Among 'ratified' hearers, Goffman distinguishes between 'addressed' (i.e. those to whom the speaker addresses his visual attention and who is expected to take the turn) from 'unaddressed' hearers (cf. 'side-participants', Clark 1996), and among 'unratified' hearers he distinguishes between 'overhearers' (who are not intended to 'take part' in the current course of action) and 'eavesdroppers' (who are secretly listening without the speaker being aware of it). Speaker's gaze is an important means for the speakers to distinguish between who is being addressed at any moment during the talk. In more recent work, 'participation framework' is viewed as encompassing the totality of all participants, both listeners and speakers, involved in the ongoing talk (Goodwin & Goodwin 2004, Deppermann 2013). Deppermann (2013) defines it as "participants' mode of presence in the interaction (…) their spatial line-up and orientation (vis-à-vis each other), their availability as recipients in terms of auditory and visual perception" (Deppermann 2013:3). In this dissertation, I will mainly adopt this second view of the 'participation framework'.

1.2.2. Regulatory functions of gaze

A second line of research has focused on the role of gaze in turn-taking (the management of speaking turns and turn transitions) and in the mobilization of backchannel responses during the talk.

Kendon (1967) was among the first to observe specific patterns in participants' gaze direction that were associated with different phases of the speaker's utterances or turns. For example, he found that speakers in two-person interactions tend to look away

just before beginning a long turn[4] and gaze back at the interlocutor as the end of the turn approaches, in this way signaling that they are 'open' for some response from the listener (see also Duncan 1972). He argued that speaker gaze at the end of the turn functions as a 'turn-yielding' cue. Kendon also found that turns that ended with gaze at the recipient were more likely to be followed by an immediate response than those that did not. Subsequent research found both supporting and contradicting evidence for this pattern, depending on the number of participants (e.g. Lerner 2003, Brône et al. 2017) and the interactional situation (e.g. Beattie 1978).

In another line of research, Bavelas et al. (2002) examined if speaker gaze can solicit a backchannel response (such as 'mm hm', 'right' and head nod) from the listener. They argued that speakers and listeners create moments of mutual gaze (or 'gaze windows'), during which backchannel responses are most likely to occur. Furthermore, Stivers & Rossano (2010) found that mutual gaze increases response relevance during actions that normally do not require a response from the recipient (e.g. announcements and noticings). The regulatory functions of gaze will be discussed in more detail in Chapter 4 and Chapter 5 of his dissertation.

1.2.3. Gaze as a social act

Scholars have also addressed the role of gaze as a 'social act' in conversation (Rossano 2013). For example, Haddington (2006) argued that gaze patterns are important resources for stance taking in the production of assessments. Through fine-grained analysis of assessment sequences in ongoing talk, he observed that interlocutors engage in *mutual gaze* while producing agreements about an assessable, in this way displaying their like-mindedness or 'convergent stance'. Similar observations on the role of mutual gaze were made by Kendon (1967) who noted that (extended) mutual gazes "appear to be indicative of an intensifying of the direct relations between the participants" (p.48) and by Argyle & Cook (1973) who observed that mutual gaze can fulfill 'affiliative'

[4] Gazing away also appears to reduce the 'cognitive load' of the moment and help the speaker in planning his utterance (Kendon 1967, Argyle & Cook 1976, Beattie 1978).

needs in cooperative situations. Furthermore, Haddington (2006) found that gaze aversions occur when the participants produce a divergent stance (i.e. a disagreement) from the one proposed by the prior speaker. Kendrick & Holler (2017) provided quantitative support for these claims by showing how preferred responses to polar questions are produced with mutual gaze with the questioner, while dispreferred responses are produced with gaze aversion.

Others have shown that gaze can be actively employed to direct others' attention to particular objects in space, in acts of 'deictic pointing' (Stukenbrock 2018) or 'gaze-cueing (e.g. Frischen et al. 2007). These studies stress the importance of gaze cueing for the establishment of a joint attentional frame in interaction.

There are undoubtedly many other functions of gaze that could not be discussed here. The main aim of this section was to provide a brief overview of the prevalent themes regarding the role of gaze in social interaction that served as a starting point for this dissertation. In the following, I will focus on interpreter-mediated interaction and discuss what is known about gaze in this type of conversational setting.

1.3. On the role of gaze in interpreter-mediated interaction

Lang (1978) pioneered research on gaze in interpreter-mediated interactions with his case study of a consecutively interpreted court session in Papua New Guinea. He described gaze behaviors of all participants in a five-minute sequence and found some regularities in their gaze (and posture) behavior in relation to their display of attention and regulation of turn transitions. Furthermore, he observed that the interpreter held his gaze averted throughout the proceedings, which seemed to signal his neutrality or non-involvement. This, however, caused the interpreter to miss some important turn management cues during the exchange. Although Lang's observations were intriguing, there was no follow-up research on this topic for over two decades after his study. It is only in recent years that we are witnessing a growing number of - mainly qualitative - studies on the role of gaze direction as a communicative resource in interpreter-mediated

exchanges (Bot 2005, Pasquandrea 2011, 2012; Mason 2012, Krystallidou 2013, Davitti 2012, Davitti & Pasquandrea 2017, Vranjes et al. 2018).

Bot (2005) was the first to offer a detailed account on the role of gaze in a dataset of nine authentic interpreter-mediated therapeutic encounters. She mainly focused on the types of engagement that the participants in these encounters display through their gaze conduct (see also Chapter 3). Furthermore, Bot observed that gaze can have a regulatory role during such exchanges, in managing the turn transfer from the therapist to the interpreter; she described one particular case in which the therapist appeared to use gaze to "drag the interpreter into the dialogue" (p. 136). Pasquandrea (2011, 2012) described the role of gaze (and body posture) in the management of interpreter-mediated doctor-patient interactions. Drawing on insights from Conversation Analysis, which studies interaction in its emerging, co-constructed context (Gardner 2001), he analyzed doctor's visual behavior during dyadic exchanges between the interpreter and the patient in the other language. Pasquandrea demonstrated how an interpreter's lack of immediate translation after the patient's utterance could be directly linked to the absence of mutual gaze with the doctor. His study stresses the importance of gaze in the establishment of mutual involvement between the doctor and the interpreter at specific moments during the exchange.

The most detailed analysis of gaze direction in interactions with an interpreter was offered by Mason (2012). His study scrutinizes gaze patterns of two participants (the immigration officer and the interpreter) during immigration interviews. Note that those immigration interviews were recorded for broadcasting purposes (television documentaries) and offer "a far-from-complete record of the dialogues" (ibid. 182). For example, mutual gaze could not be studied since the camera focused on one person at a time. Drawing on Kendon's (1967) seminal work, Mason primarily analyzed participants' gaze while listening and speaking, focusing on their displays of mutual engagement, and presented some patterns in gaze behavior that seem to be specific to interpreted encounters. In addition, Mason described some regulatory functions of gaze; he noted that intepreter's gaze shifts (and co-occurring head movements) are a "very visible marker of turn organization and helps to ensure that transitions are managed smoothly" (p. 192). Furthermore, he noted that immigration officers displayed a

tendency to gaze at the interpreter towards the end of the asylum seeker's turns, "signalling either a wish for or an expectation of the [interpreter's] turn" (p.188).

In line with Mason (2012), Davitti (2012) offered a detailed description of gaze patterns in interpreted parent-teacher meetings. Furthermore, she adopted insights and methods from Conversation Analysis to show how gaze (coupled with other verbal and visual resources) is used as a 'contextualization cue' and as a means of establishing or maintaining a triadic participation framework. Davitti also presented some observations on how gaze functions in regulating turn transitions between the interpreter and the primary participant(s). Finally, drawing on insights from Norris' (2006) analytical framework and Goodwin's notion of 'participation framework', Krystallidou (2012, 2014) discussed the role of gaze direction and body orientation as a means of participant 'inclusion' and 'exclusion' in the context of interpreter-mediated doctor-patient interactions. Table 1.1. offers a brief summary of these studies.

Altogether, these studies suggest that there is much more to be learned about the role of gaze in interpreter-mediated conversation. Mason (2012), pointed out that there is still "no full record of the concurrent interaction between speech, gesture, posture and gaze of all participants" in interpreted interactions (p.182). Moreover, previous research mostly focused on *macro*-processes such as 'participation', 'engagement' and 'inclusion/exclusion' in interpreted interactions (Davitti 2018). Further systematic research remains to be done on the role of gaze as part of specific interactional practices (at the *micro*-level) in such conversational settings.

This dissertation aims to contribute to this growing line of research by providing a fine-grained, multimodal analysis of the role of gaze in interpreter-mediated interaction, by focusing on two specific interactional phenomena: the organization of **turn-taking** and production of **backchannel responses** during the talk. These two phenomena are prevalent in any type of human communication and are typically realized multimodally, i.e. through simultaneous use of various resources such as speech, gaze, gesture, posture and facial expressions. The following sections will discuss face-to-face interaction as a *multimodal* accomplishment and present the main theoretical assumptions guiding this research.

Table 1.1: An overview of research on gaze in interpreter-mediated interaction.

Reference	SETTING	CORPUS	PARTICIPANTS STUDIED	APPROACH	ANALYTIC FOCUS	GAZE FUNCTIONS
Lang (1978)	court	1 session (Enga-Tok Pisin)	all	qualitative	gaze, posture, gesture	-participation -regulatory
Bot (2005, Ch.8)	mental healthcare	9 recorded sessions (Dutch-Dari)	(all)	qualitative	turn management, engagement	-participation -regulatory
Pasquandrea (2011,2012)	medical consultation	6 consultations (Chinese-Italian)	doctor, interpreter	qualitative	dyadic sequences	-participation -regulatory
Mason (2012)	immigration hearing	5 interviews (German-…)	immigration officer, interpreter	quantitative qualitative	gaze distribution	-participation -regulatory -(social act)
Davitti (2012,2013)	parent-teacher meeting	3 encounters (Italian-English)	all	qualitative	evaluative assessments, gaze	-participation -regulatory
Krystallidou (2014)	medical consultation	19 consultations (Flemish-Russian/…)	all	qualitative	participation, engagement	-participation

1.4. Multimodality and social interaction

People interact with each other by mobilizing various semiotic resources or modalities, such as speech, gaze, hand gestures, body posture and facial expressions (Streeck et al. 2011, Mondada 2016). Within a multimodal view of social interaction, language is seen as "integrated within this plurality [of modalities] as one among other resources" (Mondada 2016: 338). Especially within Conversation Analysis (CA) there has been a long-standing interest in "the ways in which visual orientation, gesture and other forms of bodily comportment inform the production of social action, in particular a turn-at-talk" (Heath & Luff 2013: 283). Conversation Analysis has become the most widely used qualitative approach for a systematic study of social interaction, as it allows researchers to "identify and describe the practices that interactants use in talk-in-social interaction and uses these results to understand and describe the underlying structural organization of social interaction" (Stivers 2015: 1). Conversation Analysis is not primarily concerned with language *per se*, but with actions (e.g. requesting, agreeing, telling etc.) that are accomplished through language and other resources and how participants themselves recognize and respond to each other's actions (Heath & Luff 2013: 286).

Below, I will briefly present three assumptions on the constitution of face-face interaction that emerge from Conversation Analysis and related approaches: (1) sequential organization of talk, (2) simultaneous use of verbal and visual resources and (3) reciprocity and collaboration between interlocutors in the accomplishment of joint action. These assumptions will guide the current study of the organization of gaze in interpreter-mediated interaction.

1.4.1. Sequentiality

A basic tenet of CA is that conversations are *sequentially* organized (Stivers 2013). This means that participants' turns, and actions accomplished by them, are related to what came before and what comes next (ibid.: 191). Typically, an answer follows a question, an acceptance follows an offer, a denial follows an accusation and so on. The most basic

instance of sequential organization are such adjacency pairs (Sacks et al. 1974). The production of a first pair part (e.g. a question) of an adjacency pair makes a second pair part (e.g. an answer) expectable or 'conditionally relevant' (Heritage 1984)[5]. This means that the recipient's absence of an expected second pair will be treated as problematic either by the producer of the first pair part (who will pursue a response) or by the recipient himself (who will account for the absence of an answer). Furthermore, through the production of a relevant next turn, the responder displays his/her understanding or appreciation of the prior speaker's turn (Heritage 1984). According to Heritage (1984), "linked actions, in short, are the basic building-blocks of intersubjectivity" in interaction (p. 256).

1.4.2. Simultaneity

An analysis of sequential organization of talk is complemented by an additional focus on *simultaneity* of behaviors (Streeck et al. 2011; see also Bavelas et al. 2002, Deppermann & Schmitt 2007). When people communicate with each other, they usually employ various semiotic resources such as speech, gaze and gestures simultaneously. The multimodal organization of such resources is increasingly recognized as essential to the analysis of human interaction. As Streeck et al. (2011) put it, "talk and embodied behavior co-occur as interdependent phenomena, not separable modes of communication and action**"** (p.7). The simultaneity or co-occurrence of modalities in human interaction has received a lot of scholarly interest from different disciplines, ranging from linguistics (Enfield 2009, Brône et al. 2017), gesture studies (Kendon 2004), psycholinguistics (Bavelas et al. 2002 'integrated message model') and

[5] There has been some discussion in the literature with reference to the question whether visual behavior (such as gestures) is also sequentially organized. According to some researchers, a gesture alone can constitute a relevant next turn (Heath 1984, Levinson 2013), whereas others (Deppermann 2013) reject such a view. Depperman (2012) argues that "there are no such things as 'nonverbal turns', although (…) turn-construction is multimodally informed" (p. 4). Without going into this issue any further, I want to point out that a multimodal approach to language and interaction may challenge some of the basic notions in Conversation Analysis and other disciplines studying talk in interaction.

conversation analytic-approaches (Mondada 2007, Deppermann & Schmitt 2007). These studies have focused on the ways in which semiotic resources work together to form 'composite signals' (Clark 1996, Bavelas et al. 2002), 'composite utterances' (Enfield 2009), 'ensembles' (Kendon 2004), 'micro-phenomena' (Brône et al. 2017) and 'multimodal Gestalts' (Mondada 2014). At the same time, integrating visible conduct may pose methodological challenges for Conversation Analytic approaches, since various resources do not necessarily follow a sequential arrangement during conversation (Mondada 2016).

This dissertation will explore the organization of gaze in interpreter-mediated interaction both in terms of sequentiality (e.g. Chapter 4) as in terms of simultaneity with other resources (e.g. Chapter 5).

1.4.3. Reciprocity

Another important aspect in the organization of face-face interaction is the high degree of *reciprocity* or inter-personal collaboration between interlocutors (Bavelas et al. 2002). Goodwin (1984) illustrated how a seemingly simple action such as telling a story is not accomplished by the speaker alone, but rather with the collaboration of a recipient who is expected to display his/her engagement in the storytelling process though gaze and listener responses. As argued by Clark (1996) "speaking and listening are not independent of each other. Rather they are participatory actions, like the parts of a duet, and the language use they create is a joint action, like the duet itself" (p. 20). These *joint actions* are then performed within a specific participation framework (Goffman 1981) in which persons display their constantly changing alignments with reference to the ongoing conversation.

1.5. Present study

The main aim of this dissertation is to contribute to a growing body of research on the multimodal organization of gaze within the context of ongoing interpreter-mediated interactions. The contributions of this study will be situated on the following levels:

- First, the study will single out two specific interactional phenomena as the focus of analysis, namely backchannel responses and turn-taking, that have not yet been studied systematically in interpreter-mediated interactions.
- Second, the study will introduce an innovative approach to the study of gaze in interpreter-mediated interaction, namely mobile eye-tracking.
- Third, the study will adopt a mixed-methods approach, that includes both qualitative analyses and quantitative measures of distribution of discussed phenomena.

On the basis of these findings, I shall offer an account of gaze in interpreter-mediated interaction that will strengthen our view of an interpreted event as 'joint' action.

Chapter 2

Data and method

A systematic study of gaze direction in interaction obviously requires a video corpus with specific characteristics. However, the study of gaze in interpreter-mediated talk has been faced with various challenges, ranging from gaining access to interpreted events, to dealing with impoverished data and far-from ideal recording conditions. This chapter presents the dataset used for this dissertation and introduces *mobile eye-tracking* as a novel methodological approach in the study of interpreter-mediated interaction. Furthermore, it discusses some central factors regarding the collection of the data (the setting, the participants' seating arrangements), data processing and representation of the analytically relevant gaze information for the reader (in the form of transcripts).

2.1. Methodological considerations

As argued by Mason (2012), systematic research on the role of gaze in interpreter-mediated interaction is faced with various challenges. First and foremost, it is extremely challenging to gain permission to video record naturally occurring interpreted encounters due to their often sensitive and confidential nature (Mason 2012, Pöllabauer 2004). Second, even when the recording is made by using one or even two (see Pasquandrea 2011, 2012; Davitti) video cameras, it remains difficult to register detailed gaze information (for instance, rapid gaze shifts) of all participants and at the same time capture other semiotic channels, such as gesture and posture. Third, videos do not always allow a detailed study of interlocutors' mutual gaze and errors can occur when making judgements of gaze from video recordings (Argyle & Cook 1976, Streeck 2009).

Mason (2012) points out the limitations of using video recordings for the study of gaze in ongoing interpreter-mediated talk, by observing that it is difficult to capture detailed gaze information for each participant and, in particular, mutual gaze: "If you adopt a wide angle to include all participants, detailed gaze information for each single participant is hard to detect and only one of the participants is likely to be directly facing the camera" (2012: 181). Therefore, the starting point of the current investigation was to record a dataset of interpreter-mediated interactions that would enable a detailed study of gaze of all participants in this type of encounters.

2.2. Participants and corpus

The data for this study consist of nine encounters between a Russian-speaking foreign student, a Dutch-speaking university counsellor and an interpreter at the University of Leuven in Belgium. In total, three sets of recordings were made, each consisting of three consecutive sessions with the same counsellor and student, but another interpreter (Table 2.2). Recordings were made between December 2015 and April 2017.

Table 2.1: An overview of the participants.

Participants	Gender	Age
Students	female (2), male (1)	21-26 years
Counsellors	female (2), male (1)	>35 years
Interpreters[6]	Female (3)	29-59 years

The students were contacted to come to consultations with the counsellor regarding their study program, their stay in Belgium, integration into the local university and other issues, questions or concerns they had. The students and the counsellors were previously unacquainted and had very limited to no knowledge of each other's language. Moreover, there were clear asymmetries in authority between the primary participants,

[6] See also Appendix C for further information about the interpreters.

relating to their different roles in this institutional setting (the counsellor as the institutional representative and the student as lay person), different knowledge of institutional procedures (see Drew & Heritage 1992: 49) and their age differences. The institutional setting of the encounter and the asymmetry between the primary participants approximates to a certain extent the institutional discourse in which interpreters are usually employed. It was crucial that the interactions are as natural as possible, which is essential for the study of natural feedback and gaze behavior in conversation (see also Gerwing & Bavelas 2013)[7].

Table 2.2: An overview of the recordings.

Session	Student (S)	Counsellor (C)	Interpreter (I)
1	S1: female	C1: male	I1: female
2	S1: female	C1: male	I2: female
3	S1: female	C1: male	I3: female
4	S2: male	C2: female	I3: female
5	S2: male	C2: female	I1: female
6	S2: male	C2: female	I2: female
7	S3: female	C3: female	I2: female
8	S3: female	C3: female	I3: female
9	S3: female	C3: female	I1: female

[7] Mason (2012) stresses the importance of natural conversations and ascribes limited value to the use of scripts, as "interactional goals can be acted out, convincingly even, but the deployment of paralinguistics and kinesics is unlikely to be as vital as it would be if the outcomes really mattered to the participants (cf. Bot 2005:172)" (p. 180).

Each conversation was interpreted consecutively[8] by one of the three qualified (i.e. certified) and experienced interpreters, who agreed to take part in this study (all three of them were female). All three interpreters are originally from Russia and speak Russian as their mother tongue. Note that none of the counsellors and the students who participated in this study had prior experience with talking with the aid of an interpreter; nevertheless, we can assume that the counsellors, who work at the faculty and are acquainted with the interpreter training that is being taught there, knew what is expected from an interpreter.

Each session was around 20 minutes in length, which amounts to ~180 minutes of recorded sessions. The participants gave their written informed consent prior to the conversations, in which it was stated what the study was about and how the data were going to be used.

2.3. Seating arrangement

Previous studies have noted that the seating configuration of the participants, and especially of the interpreter, can have an impact on the conversational dynamics of the encounter (Wadensjö 2001, Bot 2005, Llewellyn-Jones & Lee 2014). Although the codes of conduct for community interpreters in Belgium (and in the Netherlands) outline that interpreters in conversational settings should position themselves at a more or less equal distance from the primary participants (thus forming a triangular shape, see Figure 2.1)[9], this may vary according to the context and personal preferences of the professional parties. For instance, in her study of interpreting in the context of psychotherapy in the Netherlands, Bot (2005) described a case in which a therapist instructed the interpreter to sit slightly behind him, so that the patient would maintain his orientation at him and

[8] From time to time, interpreters rendered in (almost) simultaneous interpreting mode, especially in the case of short turns, when they could easily project the content of the turn.

[9]http://www.agii.be/sites/default/files/bestanden/deontologische_code_sociaal_tolken.pdf

not specifically at the interpreter[10]. According to Bot, such seating arrangement stems from the idea of the interpreter as a 'translation-machine'. In this way, however, the interpreter and the therapist form one party and the patient the other. A different case was reported by Wadensjö (2001), who described an encounter where the interpreter was instructed to sit behind the patient, in order to reduce visual contact between the interpreter and the patient and "minimize the risk of her getting involved in processes of transference and counter-transference" (p.72). Nevertheless, the interpreter reported being "caught in an undesired relationship with the patient, sitting physically close to her" (ibid.: 72).

Wadensjö (2001) argues for the importance of the interpreter being within the 'communicative radius' shared by all participants present, which maximizes their opportunity to perceive one another. In this way they form, what Goffman has called, an 'eye-to-eye ecological huddle' (Goffman 1963) or an F-formation (face-formation), i.e. orientation in space so that "each is facing inward around a space to which each has immediate access" (Ciolek & Kendon 1990). Wadensjö further observes that in talk where the interpreter is outside of the communicative radius formed between the primary participants, the talk seemed less synchronized. She lists the following potential advantages of the interpreter's placement within a shared communicative radius (p. 83-84):

- it facilitates the interpreter's coordination of interaction (turn-taking) by making it visible to both primary interlocutors;
- it allows the interpreter to mark the necessary distinction between "the speaking self" and "the meaning other (primary participant)" through (non-)verbal resources;
- it also allows the interpreter to promote the affinity between primary interlocutors.

[10] Llewellyn-Jones & Lee (2014) note in this regard that "if the interpreter takes up the position that would be expected in any triadic interaction, i.e. more or less equidistant from the interlocutors" the interpreter can simply "re-direct" the speaker's gaze to the addressee with his/her own gaze (p.44).

In the present dataset, the participants were always seated in a triangular formation in line with the code of conduct for community interpreters in Belgium (see Figure 2.1), with the interpreter in the middle and on more or less equal distance from the primary speakers, who were seated opposite to each other.

Figure 2.1: Seating arrangement of the participants.

Such seating arrangement maximizes mutual eye contact between interlocutors and the opportunity to monitor each other's actions, allowing us to highlight some systematic regulatory features of gaze in interactions with an on-site interpreter. Or in the words of Llewellyn-Jones & Lee (2014), the result of such a seating arrangement "is a genuine opportunity for a triadic communicative event, one in which there is a shared *situation*, a shared construction of meaning." (p. 44)

2.4. Recording set-up and technical equipment

In order to get detailed information on the gaze behavior of the participants, each participant was wearing *mobile eye-tracking glasses* during the conversation (see Appendix B for specifications on the types of eye-trackers that were used). Mobile eye-tracking glasses resemble spectacle frames with typically two integrated cameras: (i) a

scene camera[11] that captures the visual field of the participant and (ii) an eye camera (corneal reflex camera) that records the movements of the pupil (see Gullberg & Kita 2009, Brône & Oben 2015, De Beugher 2016, Conklin et al. 2018). The participants are thus able to move their bodies freely while being eye-tracked. An additional camera was set up (*Sony HDR-FX1000E*) to provide a profile shot of the interaction (Figure 2.1) and a microphone (Zoom H2 or iPhone Microphone) to record the audio track.

In comparison to video recordings, the eye-tracking systems allow for a fine-grained analysis of gaze direction (e.g. rapid gaze shifts) in ongoing, spontaneous interaction while at the same time capturing other semiotic channels, such as gesture and posture. Thus, these eye-tracking systems provide objective gaze information for analyses (Jokinen 2010)[12]. Moreover, with the recent advances in mobile eye-tracking technology (viz. the eye-trackers becoming more light-weight, unobtrusive and easier to use), it is now possible to study eye gaze not only as a measure of cognitive processing (which is the traditional focus of eye-tracking research) but also as a communicative instrument that co-participants actively employ and engage with in interaction. Mobile eye-tracking technology is increasingly being used within cognitive interaction research (e.g. Brône & Oben 2015, Oben & Brône 2015, Holler & Kendrick 2015), conversation analytic research (e.g. Weiß & Auer 2016, Auer 2017, Kendrick & Holler 2017, Stukenbrock 2018) and recently also in face-to-face dialogue interpreting research (Vranjes et al. 2018)[13]. Altogether, these studies show that, by measuring multiple

[11] Since the scene camera moves with the head, it is much easier to register slight movements of the head, such as head nods.

[12] Note that eye-tracking is directed at **overt attention** and not at **covert** attention that is directed at some point in the surrounding space (Holmqvist & Andersson 2017: 26). However, according to Holmqvist & Andersson (2017), this does not threaten the validity of studies that focus on eye gaze as the starting point of attention, since it would require a lot of control "to inhibit gaze from saccading to the position of covert attention" (p.26). However, gaze movements that fall outside of the scope of the scene camera (for instance, when the participant gazes down without moving his head) remain invisible for the analysis.

[13] Eye-tracking technology has already been used for the study of simultaneous interpreting, where pupil dilation was used to measure cognitive load (see e.g. Seeber 2013, 2015). Those studies used remote eye-tracking technology instead of mobile eye-tracking technology.

participants' eye gaze simultaneously, we can get a rich insight into the interaction dynamics of gaze distribution.

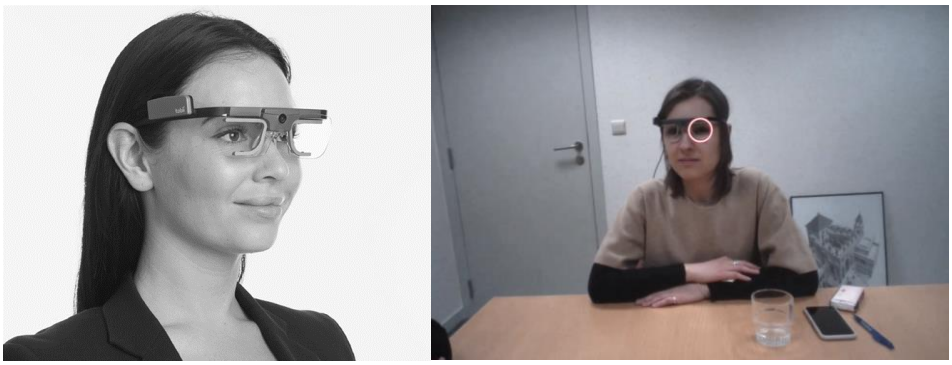


Figure 2.2: An example of eye-tracking glasses (*Tobii Pro Glasses 2*) (left picture) and the eye-tracking recording (right picture). Participant's gaze fixation is indicated with a colored dot (also called *gaze cursor*).

What also needs to be considered here is the potential influence of the eye-trackers on the participants' (gaze) behavior. Some participants reported being 'almost not aware' of the fact that they were wearing an eye-tracker during the conversation; other participants reported that they were thinking about them from time to time. Interestingly, some participants mentioned after the second (or even third) session that they had been more aware of the eye-trackers than the first time they were wearing them. This implies that an increased familiarity with the recording equipment does not to guarantee a decrease in the participants' awareness of it. Moreover, the participants could at all times see the mobile eye-trackers worn by their interlocutors. Altogether, it is difficult to determine the level of intrusiveness of the mobile eye-trackers on the ongoing talk (see also Vranjes et al. 2018). We can assume that the participants were primarily oriented at achieving some interactional goals during the talk: the students and the counsellors were oriented towards exchanging information and establishing an interpersonal relationship, whereas the interpreters was focused on her task of rendering the talk. As noted by Wadensjö (1998) on the 'naturalness' of recorded data "[o]ccasionally, subjects may feel as if they are in something like an examination

situation. In that case, it should perhaps be expected that they are trying to do what they understand to be best" (p.95).

2.5. Recording procedure

The recordings took place either in the counsellor's office or in another room at the campus. The recording equipment was set up prior to the conversations, so that the eye trackers could be calibrated[14] immediately upon the arrival of the interpreter and the student. The calibration was performed by myself and one or two colleagues experienced in eye-tracking. During this preparation phase, the participants were also asked to sign the informed consent, in which it was stated what the study is about and how the data were going to be used. This whole procedure took up several minutes, allowing the participants to get accustomed to wearing the mobile eye-trackers before the actual start of the recording. During the recording, the participants were alone in the room. Each session lasted approximately 20 minutes. At the end of each session, I conducted a short (semi-structure) interview with each participant individually to learn about their background and their impressions about the encounter.

2.6. Data processing

The output of each conversation consisted of four video recordings (from the three eye-trackers and one regular video camera), which were then synchronized in Adobe

[14] Calibration consists of displaying a series of points on a screen or wall, that the participant must fixate. In this way, the eye-trackers are adjusted to each individual participant. This procedure is necessary in order to optimize data recording (Conklin et al. 2018: 75).

Premiere Pro into one single video or a *quadvid*, which displays all four perspectives at once (see Oben 2015 for a detailed description of the synchronization process), as shown in Figure 2.3. This step is crucial, as it enables the researcher to analyze synchronous gaze information of all participants and to keep track of all three participants' gaze orientations at each step of the analysis. Note that, due to technical issues, more than 90% of gaze data from Student 1's eye-tracker were lost and were therefore left out from further analysis.

These synchronized videos were imported into ELAN (EUDICO Linguistic Annotator), an open-source tool for transcription and annotation of video and audio data, developed at the Max Planck Institute for Psycholinguistics, (Nijmegen), Netherlands (Wittenburg et al., 2006). ELAN allows researchers to conduct both qualitative and quantitative research on large amounts of video recordings in an efficient and systematic way. First, the tool displays annotations and transcriptions on a timeline under the video player. In this way, one can easily detect, for instance, silences or moments of overlapping talk, whilst at the same time having direct access to the video recording (see Figure 2.4). Also, the advanced playback options (e.g. playing a selection, playing in slow motion) enable the researcher to examine in detail any interactional phenomenon of interest. Second, ELAN allows the researcher to make annotations on different layers or 'tiers', which makes it easier to determine if some behaviors co-occur with others (for instance, silence and a certain type of gesture). This is of particular interest for any type of multimodal analysis. Third, ELAN has an advanced 'search'-function, which makes it possible to search multiple annotation files at once.

Figure 2.3: Four synchronized perspectives in one video.

All recordings were transcribed and annotated for gaze direction for each participant separately. Talk of each participant was segmented into intonation units, viz. stretches of speech with a single intonation contour (Selting 2000, Brône et al. 2017) and transcribed on separate tiers in ELAN (Figure 2.4).

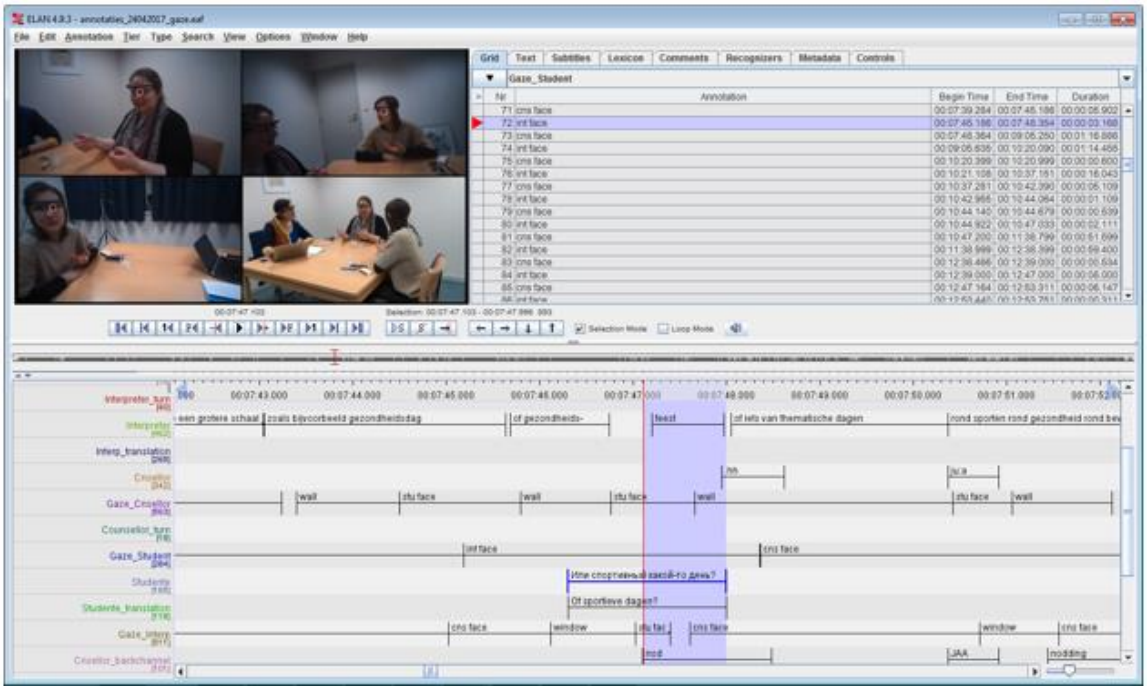


Figure 2.4: An example of an ELAN file, with the synchronized video (upper left corner) and different 'tiers' with annotations (below).

Annotations of gaze were also done manually which resulted in 13.362 annotations for gaze for the nine recordings. Although this is a very time-consuming process, it was necessary for the systematic approach envisioned in this study. In the meantime, new tools are being developed that will help speed-up this process (see De Beugher 2016) and facilitate the use of eye-tracking for future studies.

Since eye-trackers provide a continuous stream of gaze data, it was necessary to segment the data in some way (Brône et al. 2017). In line with previous studies (Oben 2014, Brône et al. 2017) the annotations were segmented on the basis of *gaze fixations*[15], viz. the intervals between the eyes' movements when the eyes stand still (Holmqvist & Andersson 2017: 14). According to the 'eye-mind hypothesis', when we measure a fixation, we also measure attention to that position (ibid.: 14). A commonly used value for gaze fixations in interaction research, which was also adopted in this dissertation, is 120 milliseconds (Vertegaal et al. 2001, Oben 2015). In the present dataset, the participants were typically fixating each other's faces, the background or objects in space during the interaction. Hence, the coding of gaze fixations was based on a limited tag set (see also Brône et al. 2017), such as 'interpreter's face', 'student's face', 'wall', 'paper' and so on.

The analyses were based on close examinations of the synchronized videos in ELAN that allowed me to define the phenomena of interest. After determining the basic interactional features of the phenomena I was focusing on, I coded them in ELAN. In order to further support my analyses, I used transcriptions in conversation analytic style (see Chapters 4-6). As for the statistical analyses, the software SPSS was used.

[15] Holmqvist & Andersson (2017) point out that the term 'fixation' is a bit misleading because the eye is not completely still, but exhibits distinct types of micro-movements (e.g. tremor, micro-saccades and drifts) (p.14).

2.7. Notes on the multimodal transcription conventions for gaze

As described in the introduction of this dissertation, I will combine a quantitative and a qualitative, CA-inspired approach for the analysis of gaze in interpreter-mediated interaction. Since the use of transcripts as a means of discovering and describing "orderly practices" in social interaction (Hepburn & Bolden 2013: 58) is central to a CA-approach, it seems at place to discuss here some of the issues associated with the transcription of gaze. How do we transcribe gaze (and other (non-)verbal conduct) in social interaction? Within CA, there are well-established standards for the transcription of talk, of which Jefferson's transcription conventions (e.g. Jefferson 2004) and GAT 2 (Selting et al. 2011) are most commonly used. This is, however, not the case for the transcription of gaze or, more broadly, nonverbal conduct (Deppermann 2013, Mondada 2018). Gaze information can be provided in the form of transcriptionist's comments (inserted in the transcript in brackets), specialized notational systems (e.g. Rossano 2012) and visual representations (e.g. drawings and video frame grabs) (Hepburn & Bolden 2013: 70). Currently, the most known transcription systems for gaze are those developed by Goodwin (1981, see also Haddington 2006 for an adapted version) and by Rossano (2012, 2013). An example of both transcription conventions is shown and discussed below.

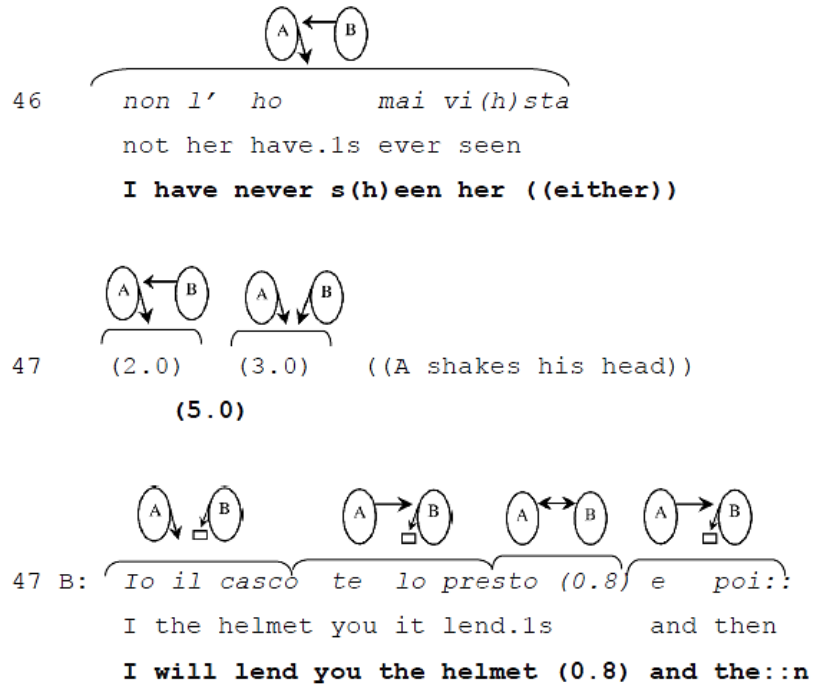
Extract 1 - Goodwin's transcription system (1981: 80).

(31) MARSHA: ..... X_____
[
An*o*ther interesting group were the one s from
[
DIANNE: ..... X_____

In Goodwin's notation system, speaker's gaze towards the recipient is marked above the utterance, whereas the gaze of the recipient(s) is marked below it. A series of dots marks gaze movement and a capital X shows at which point in the talk gaze reaches the other (Goodwin 1981: vii). However, as pointed out by Rossano, Goodwin's notation

system only shows whether participants are looking at each other's face and when there is a gaze shift (Rossano 2012: 49). Rossano devised a new notation system, an example of which can be found below.

Extract 2 - Rossano's transcription system (2012: 274).



Rossano lists the following main properties (and advantages) of his notation system (Rossano 2012: 49-50):

- it is more iconic than any other system;
- it is combinatorial, in the sense that it allows adding any number of possible symbols;
- it shows various gaze configurations (e.g. looking away, looking up, looking at an object etc.);
- it represents all participants' gaze in one symbol;
- duration of gaze is represented with brackets.

Rossano's system has been adapted to three-person interaction by Davitti (2012) and Vranjes et al. (*forthcoming*), who focused on triadic interpreter-mediated

interactions, and by Auer (2017), who analyzed triadic same-language interactions between peers. Although Rossano's transcription format has a number of obvious advantages, it also has some limitations. First, such symbolic encoding of gaze behavior is very time-consuming. Second, and related to the first point, it may not always be necessary to transcribe all interlocutors' gaze in such detail (for instance, when focusing only on unaddressed participant's gaze, see Holler & Kendrick 2015). Indeed, the level of granularity for the transcription of gaze will depend on the analyst's research question. Third, it does not reveal much about the co-occurrence and configuration of different resources in the construction of 'composite signals' (Clark 1996) and 'multimodal Gestalts' (Mondada 2014); only gaze is represented visually in Rossano's transcription system, whereas for other semiotic channels he still uses the 'traditional' transcription method with comments (in brackets). Because of this, some multimodal patterns do not clearly come forward in this notation system. Fourth, it does not show the characteristics of the 'ecology' of the social activity (the surrounding environment, the positions of objects and of participants' bodies in space), which occasions the employment of specific resources by the interlocutors (Mondada 2018). Therefore, the use of frame grabs is particularly useful, as they provide a *synthetic* view on the social action (Mondada 2018). As argued by Mondada (2018), "images are not only contributing to the representation of the movements textually described in the transcript but also to their holistic composition and the ecology in which they happen" (p.90).

The transcription system used in this dissertation is loosely based on Mondada (2013) (See Appendix A for a description of transcription conventions used). An example is shown in Extract 3. Gaze is transcribed on a separate line under the transcription of speech. This transcription format has a number of important features: (1) it allows the analyst to represent the different modalities separately and in relation to each other. As shown below, the modalities of interest (speech, gaze, head movements) are represented on separate lines, which allows us to study their co-occurrence and co-ordination during talk; (2) the numbered frame grabs in the transcript provide a synthetic view of the participants and of the local ecology of action. Moreover, frame grabs taken from the eye-tracking recordings also have the advantage of displaying the gaze direction (represented in green) of the person wearing them; (3) it

allows us to focus on analytically relevant phenomena. For the sake of the reader (and to increase the readability of the transcripts), I will mostly work with frame grabs to show most prominent gaze behaviors. However, when needed for analytical purposes, I will provide a more detailed transcription of gaze behavior of each of the participants.

Extract 3

1 **INT** **#maar meestal is het in de wee:k+#**
but mostly it is during the week
Cns gaze at Stu----------gaze at wall--->
cns **+double nod--->**
fig #fig 1 #fig 2

figure 1 (CNS perspective)

figure 2

2 **INT** **en rond vijf uur 's a:vonds+**
and around five in the evening
Cns ----------gaze at Stu-->
cns ----------+

Thus, although multimodal transcripts are a "means to enable the reader to see what matters analytically" (Deppermann 2013: 3), they are necessarily selective and a reflection of the analyst's research questions. As pointed out by Deppermann (2013), "[w]hich phenomena are captured in which way depends on what participants make relevant in their activities and on the research question and analytical interests which in turn need to be adapted to the participants' orientations" (p.3). Moreover, I have shown that a combination of different representations (descriptive and/or symbolic and visual) is sometimes needed.

Chapter 3

Distribution of gaze orientation in interpreter-mediated talk

The aim of this chapter is to provide an overview of the quantitative distribution of gaze in the recorded dataset, which will serve as a backdrop for the analyses presented in subsequent chapters. In line with previous research, I will investigate how interlocutors' participation status (speaker versus listener) and their respective social positons or roles within the conversation affect their gaze behavior. First, the focus will be on the specificities of gaze direction of the primary participants (counsellors and students), followed by an analysis of the interpreter's gaze during the interaction. I will compare the observed patterns with findings from previous studies on gaze in (interpreter-mediated) interaction and discuss some limitations of an exclusively quantitative approach to the analysis of gaze.

3.1. Gaze and participation status in face-to-face talk

Early work by Kendon (1967) revealed an asymmetrical gaze pattern in face-to-face interaction, linked to the participation status of the interlocutors: while listening, participants sustain their gaze at their interlocutor for a longer period of time than while they are speaking. By gazing at the speaker[16], recipients show that they are acting as listeners of the ongoing talk. Kendon also found that the participant's glances "during speech tend to be shorter than those observed during listening". By looking away from listeners during longer utterances, the speaker 'shuts out' visual input from the

[16] 'Speaker' is defined as the person who performs a particular action (Levinson 2013), such as requesting, inviting, complaining etc. during a turn-at-talk.

interlocutor and concentrates on planning what s/he is going to say. It is through this practice that the speaker "signals his intention to continue to hold the floor" (Kendon 1967: 42). However, in interactions with more than two participants, speakers usually divide their visual focus between several co-participants. In those cases, the speaker has to make clear to whom s/he is addressing his or her utterance (Goffman 1981, Goodwin 1981, Auer 2017). Moreover, experimental research has shown that speakers "compensate" for divided visual attention by increasing their total amount of gaze at interlocutors (Vertegaal et al. 2001).

Kendon's findings were further elaborated by Goodwin (1980, 1981), who analyzed gaze orientations of hearers and speakers in *relation* to each other. He formulated a set of basic rules, stating that (a) a speaker should obtain the gaze of his recipient during the course of a turn at talk, and that (b) a recipient should be gazing at the speaker when the speaker is gazing at the hearer[17]. If the speaker finds that the recipient is not looking at him and thus displaying diminished engagement in the talk, he has different ways (such as interruptions, pauses and restarts) to solicit the recipient's gaze (Goodwin 1980)[18]. Goodwin also notes that "gazing at a hearer is inappropriate" (Goodwin 1981: 45), since one should be gazing at the speaker[19]. There are thus normative specifications associated with gaze conduct of the participants in face-to-face interaction (for an extensive overview, see Rossano 2013)[20]. However, as noted by Rossano (2013), we still do not know whether factors such as social relationships,

[17] Goodwin (1980: 287) also stipulates that "[a] recipient can look at the speaker when the speaker is not looking at the recipient without the rule being violated. However, if the speaker gazes at a non-gazing hearer, the rule is violated."

[18] These observations pertain to communicative activities in which interlocutors are primarily oriented at each other. Situations where the recipient is engaged "in a competing activity" (such as eating) provide "a ready account for looking away rather than at the interlocutor's face and thus makes it less sanctionable" (Rossano 2013: 313).

[19] Experimental research has shown that interlocutors gaze more at people they speak or listen to than at other people (Vertegaal et al. 2001).

[20] Note that these are also defined by culture, gender and race. For instance, studies have reported that Black Americans look at the recipient more while speaking than while listening, whereas White Americans tend to follow the opposite pattern (Rossano 2013: 312).

hierarchy and power asymmetries "modify the norms suggested in the works by Kendon [and] Goodwin" (p. 325).

3.2. Gaze and participation in interpreter-mediated talk

In the specific setting of interpreter-mediated dialogues, interlocutors do not only employ their gaze to confirm their participation status (as listener or speaker) (Mason 2012), but also to *position* themselves and others within the encounter (Mason 2012, Davitti 2012). Previous studies have reported that some interpreters display a tendency to avert their gaze from interlocutors in order to signal their 'neutrality' (Lang 1978, Mason 2012), whereas others adopt a more involved communicative style during the talk (Bot 2005, Pasquandrea 2011, Davitti 2013). Mason (2012) observed that interpreters may display different orientations towards primary participants within the same encounter: while interacting with immigration officers, interpreters in his data tend to "signal their quasi-non-participant status via their deflected gaze", whereas when interacting with asylum seekers, they look at them and "pay close attention to what [the latter] is saying (p.191). Thus, as noted by Davitti (2012: 164), gaze "may be considered an indicator of the different positioning taken up by the [interpreter]" in ongoing interaction.

Primary participants are often instructed to gaze at each other "instead of looking at the interpreter" (Englund Dimitrova 1997: 157), despite the fact that they do not understand each other's language[21]. Such recommendations stem in part from the idea

[21] Bot (2005: 132) notes in her study on gaze in interpreter-mediated therapeutic encounters that "In the Netherlands, primary speakers in interpreter-mediated dialogue are specifically instructed - through brochures of the *TVCN* - to look at each other and not at the interpreter. This instruction is communicated explicitly to therapists, sometimes also to patients". However, she also notes that such behavior "goes against the grain of 'ordinary conversation'" and "the users of interpreter services have to make an effort to behave in the prescribed way" (p. 132).

of 'normal' gaze behavior, viz. gazing at the interlocutor "to whom the words are meant", and in part also from "the idea of the interpreter as a non-person" (Bot 2005: 132). In her analysis of gaze patterns in interpreter-mediated therapeutic encounters, Bot (2005) reported that therapists tend to gaze at the patient and not at the interpreter while speaking, thus displaying their "disengagement" from the latter. A similar observation was made by Mason (2012) in his systematic study of gaze in interpreted immigration hearings. A deviation from this pattern was reported in Davitti (2012), who found in her study of interpreter-mediated teacher-parent meetings that both teachers and parents gaze more at the interpreter than at each other during their own turns[22]. Such gaze behavior confirms the interpreter's status as full-fledged participant in the exchange, but "excludes the linguistically-different participant" (De Pedro Ricoy 2009). Krystallidou (2014) noted in that regard that the interpreter can either be 'fully' ratified (verbally and through gaze) or 'partially' ratified (only verbally) as addressee during the talk.

It has also been observed that 'institutional parties' (e.g. therapists) and 'non-institutional parties' (e.g. patients) may display different orientations toward the interpreter. For instance, Bot (2005) found that, whereas therapists tend to maintain gaze at the patient while listening to the interpreter, patients tend to gaze at the interpreter. According to Davitti (2012), the professional's gaze behavior "may be the result of training guidelines or, more simply, of a regulated behavior which has gained consensus within their community of practice" (Davitti 2012: 167). Moreover, such gaze behavior may also reflect asymmetries in power relationships within the interaction (Mason 2012).

In the following sections, I will give a systematic overview of the overall distribution of gaze in the recorded dataset in order to determine the specificities in gaze behavior of these interactions. There are different ways of measuring participants' gaze orientations in ongoing interactions. One way of measuring individual and mutual gaze

[22] According to Davitti (2012), "[t]his may be due to the lack of an established common practice, which, on the contrary, characterizes other settings" (p.143).

is "the percentage of time spent looking" (Argyle & Cook 1994: 848). The percentage of time provides a useful representation of specific variables associated with gaze in face-to-face interaction (e.g. the proportion of mutual gaze between interlocutors). The other way of measuring gaze includes "the number and length of glances" (ibid.: 848, For instance, a higher number of fixations serves as an indication of increased visual attention towards an object or a person. In this work, I am interested in gaze as a noticeable and potentially significant event or *social act* (cf. Goodwin 1981) within interaction. Therefore, the present chapter provides an analysis of frequencies and length of participants' gaze, in relation to their social positions within the exchange and their participation status (speaker versus listener) during the talk.

3.3. Primary participants' gaze while speaking

This section presents an analysis of the primary participants' (counsellors' and students') gaze during their own turns[23]. I focus on three relevant regions of interest (ROI's) for each participant, namely, (i) the face of speaker A, (ii) the face of speaker B, (iii) elsewhere in the room (see also Vranjes et al. 2018). Mean durations (M) of fixations were measured in seconds. One-way ANOVA analysis was performed to test whether there was any statistically significant difference between the means of percentages of gaze fixations at different ROI's. In addition, a Bonferroni Post-hoc analysis was performed to discover which pairs of means of gaze fixations at particular ROI's were significantly different (Vanhoomissen 2012).

3.3.1. Counsellor's gaze while speaking

Figure 3.1 shows the counsellors' distribution of visual attention while speaking. The analysis reveals that the counsellors gaze significantly more at the student (46%,

[23] This does not include the production of backchannel responses (see Chapter 4), such as 'mm hm' or 'yeah' during another speaker's talk.

M=2.257 s), who is the primary addressee of their talk, than at the interpreter (17%, M=1.052 s) during their own turns. This is in line with what was reported in previous studies (Bot 2005, Mason 2012). Nevertheless, through their occasional glances at the interpreter, the counsellors appear to engage the interpreter in the talk. As these counsellors were not accustomed to talking with the aid of a consecutive interpreter, we cannot say that this behavior arises from their experience or specific training. Their gaze direction towards the student can be seen as a manifestation of their 'professional' attitude within the exchange. Moreover, although they are aware of the fact that the student cannot understand their utterance, they employ their gaze to visually ratify the latter as their addressee. The student becomes in this way the 'addressed ratified participant' (Goffman 1981) of the counsellor's words, viz. the hearer to whom the speaker's words are particularly oriented. It is through this visual ratification process that counsellors present themselves as 'professional' and/or 'powerful'[24] participants within the interaction.

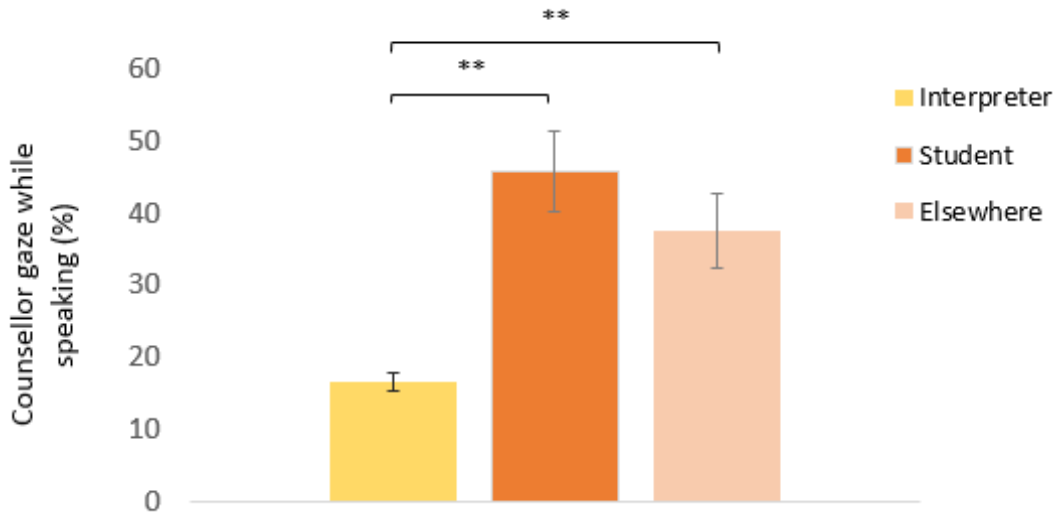


Figure 3.1: Overall proportion of gaze fixations for the three ROI's, averaged across all three counsellors. Error bars show the standard deviation. One-way ANOVA test ($F(2,6) = 34{,}57$, $p = 0.005$) with Bonferroni Post-hoc analysis (**$p < 0.01$).

[24] Previous studies have also argued that gaze reflects the power relationships within the interaction (see also Mason 2012). According to Argyle & Cook (1977) more dominant speakers look more towards their addressees in competitive situations.

Furthermore, the relatively high proportion of counsellor's away-gaze (38%, M= 1.052 s) is more in line with what was reported for two-person interaction (see Kendon 1967). According to Vertegaal et al. (2001) "speakers need not to look away when addressing a group because they can easily avoid prolonged eye contact by looking at other people" (p. 306). In this case, the counsellor does not seem to address the interpreter and the student as a 'group'; the high proportion of counsellor's away-gaze is a result of his primary orientation at the student.

3.3.2. Student's gaze while speaking

In contrast to the counsellors, the two students[25] in the recorded sessions address their utterances primarily at interpreters (Figure 3.2), with whom they share a common language. Moreover, they scarcely look at the counsellor (only 2% of their gaze fixations), who is temporarily excluded from the interaction, both linguistically and visually. Thus, the students orient towards the interpreter as a full-fledged participant in the interaction (Bot 2005, Davitti 2012). The contrast in gaze conduct between students (Figure 3.2) and counsellors (Figure 3.1) may result from the fact that the students (given his/her age and non-professional status) exert less authority in these exchanges. There is also a matter of hierarchy between the student and the interpreter who is older and obviously more knowledgeable of the local language and culture than the student.

[25] As mentioned earlier in Chapter 2, the eye-tracking data from Student 1 were lost for the most part due to technical issues. Therefore, I do not include her gaze behavior in the overview.

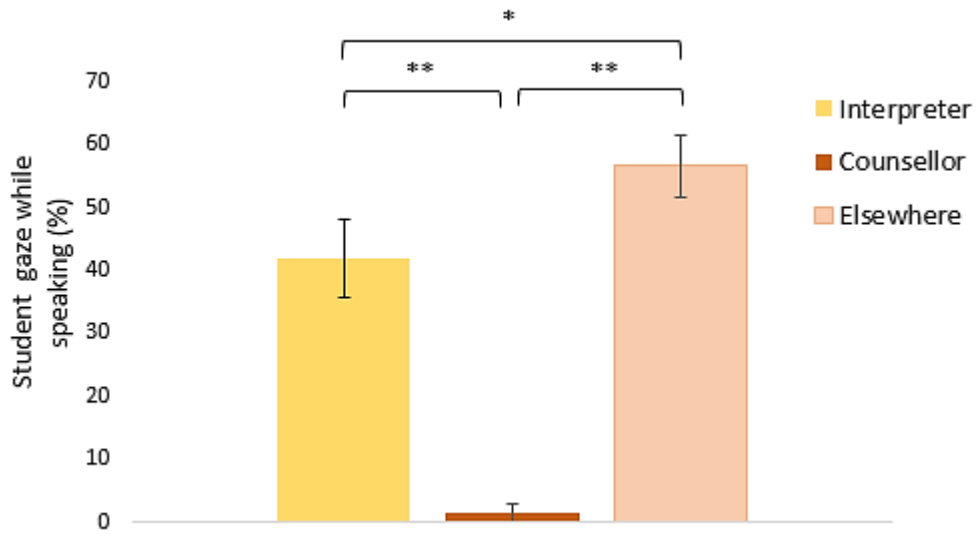


Figure 3.2: Overall proportion of gaze fixations for the three ROI's, averaged across the two students. Error bars show the standard deviation. One-way ANOVA test ($F(2,6) = 114.15$, $p < 0.001$) with Bonferroni Post-hoc analysis (*$p < 0.05$, ***$p < 0.001$).

Altogether, these results show that students gaze at the interpreter while speaking, whereas counsellors gaze at students, which is illustrative of their 'professional' status within the exchange. These differences are also indicative of their different orientations towards the interpreter (for a similar observation, see Bot 2005: 140): counsellors display some disengagement from the interpreters and seem to orient towards the latter as 'conduit'. Students, on the other hand, orient towards the interpreter as a full-fledged participant. Thus, the primary participants' gaze patterns reveal much about their engagement with the interlocutors, their own position, and the power-relations within those exchanges.

3. 4. Primary participants' gaze while listening

The following section focuses on primary participants' gaze orientations while listening to the interpreter or to the other primary participant. I made an additional distinction between the 'Interpreter's turn in Dutch (NL)' and the 'Interpreter's turn in Russian

(RU)' in the assumption that the understanding of the language will have an influence on the listener's gaze direction.

3.4.1. Counsellor's gaze while listening

The distribution of counsellors' gaze fixations while listening is represented in Figure 3.3. During **student's turns**, the counsellor maintains significantly more gaze at him/her for long periods of time (73%, M=4.795 s) when compared to the interpreter (16%, M=0.621 s), although s/he has no understanding of the student's language. This pattern is in line with Goodwin's (1981) above-mentioned maxim that listener should be gazing at speaker. The counsellor occasionally glances at the interpreter during student's turns, which may be explained as the counsellor's way to monitor the interpreter's potential reactions or any nonverbal displays of incipient speakership.

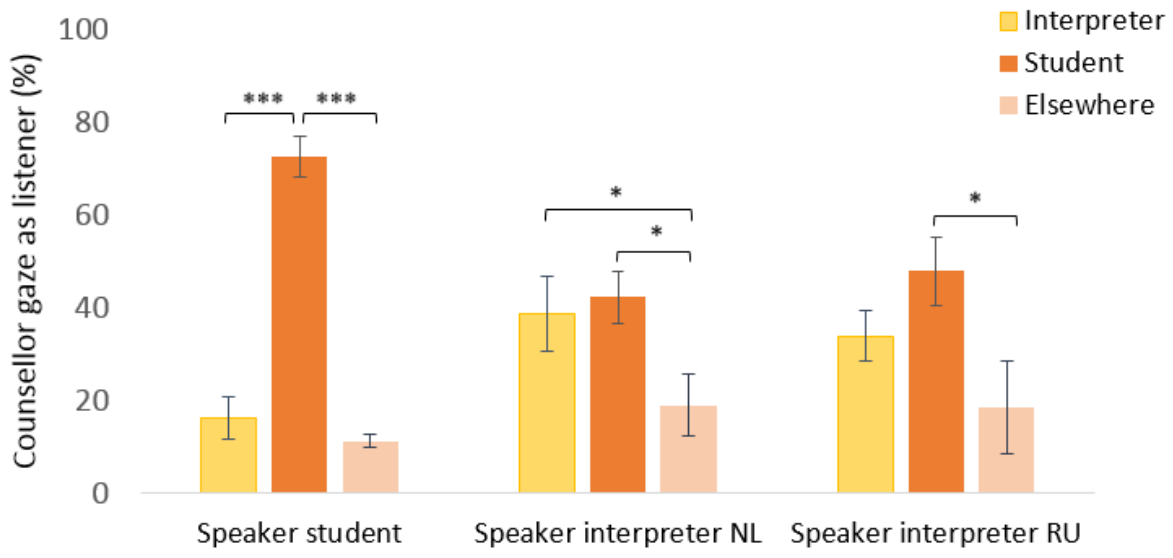


Figure 3.3: Overall proportion of gaze fixations for the three ROI's, averaged across all three counsellors. Error bars show the standard deviations. One-way ANOVA test (speaker student: $F(2,6) = 248.90$, $p = 1.68e\text{-}06$; speaker interpreter NL: $F(2,6) = 9.95$, $p = 0.0124$, $p < 0.001$; speaker interpreter RU: $F(2,6) = 10.49$, $p = 0.0110$) with Bonferroni Post-hoc analysis (*$p < 0.05$, ***$p < 0.001$).

Although counsellor gazes almost equally at the interpreter and the student (39% and 42%, respectively) while listening to the interpreter's rendition in Dutch ('**interpreter NL'**), his gaze fixations at the interpreter are of longer duration (M=3.18 s) when compared to those at the student (M=1.105 s). This goes against Goodwin's rule that listener should be gazing at speaker, when speaker is gazing at hearer (on interpreter's gaze direction, see Section 4.1. below). Counsellor's gaze seems to acknowledge that the words he hears are the student's (Bot 2005: 140) and, consequently, *includes* the student in the ongoing interaction from which s/he is temporarily excluded. We will examine this gaze pattern in relation to the therapist's backchanneling behavior (what we refer to as 'dual feedback') and in relation to the ongoing discourse in Chapter 4.

Finally, during the interpreter's renditions in Russian ('**interpreter RU'**), the counsellor tends to gaze more at the student (48%, M=2.068 s) than at the interpreter, who is the current speaker (34%, M=1.432 s). A similar observation was made by Bot (2005: 140), who - drawing on Goodwin 1981 - noted that this "may be felt as intrusive, as 'being watched' more than necessary". Yet, it is difficult to establish the level of 'intrusiveness' in this setting; at least in the present dataset, none of the participants expressed nor displayed any discomfort about being 'watched' (based on the post-interviews). This can be attributed to the fact that, during the interpreter's rendition of the counsellor's turn, the interpreter and the counsellor both act as a 'party' of speakers with a clear division of labor between them; the interpreter acts as spokesman for an action (such as telling, request or agreement), for which the counsellor is the 'principal' (see Goffman 1981). Thus, the counsellor is in a sense still involved in the interpreter's turn, which warrants his gaze at the addressee. By gazing at the student, the counsellor monitors the student's reactions and presents himself as 'principal' of the ongoing turn (see also Chapter 5). Thus, as will be shown in Chapter 5, 'hearer' and 'speaker' are no mutually exclusive notions, but oftentimes overlap and are implicated in one another in triadic (interpreter-mediated) interactions.

3.4.2. Student's gaze while listening

Students 2 and 3 display large differences in gaze orientations while listening to the ongoing talk. I will therefore discuss each of their gaze patterns separately in the following paragraphs.

As shown in Figure 3.4, Student 2 gazes significantly more at the counsellor (42%) during latter's turns than at the interpreter (5%). However, Student 2 displays a rather atypical gaze pattern while listening to the interpreter (NL and RU), as he usually looks away during her turns.

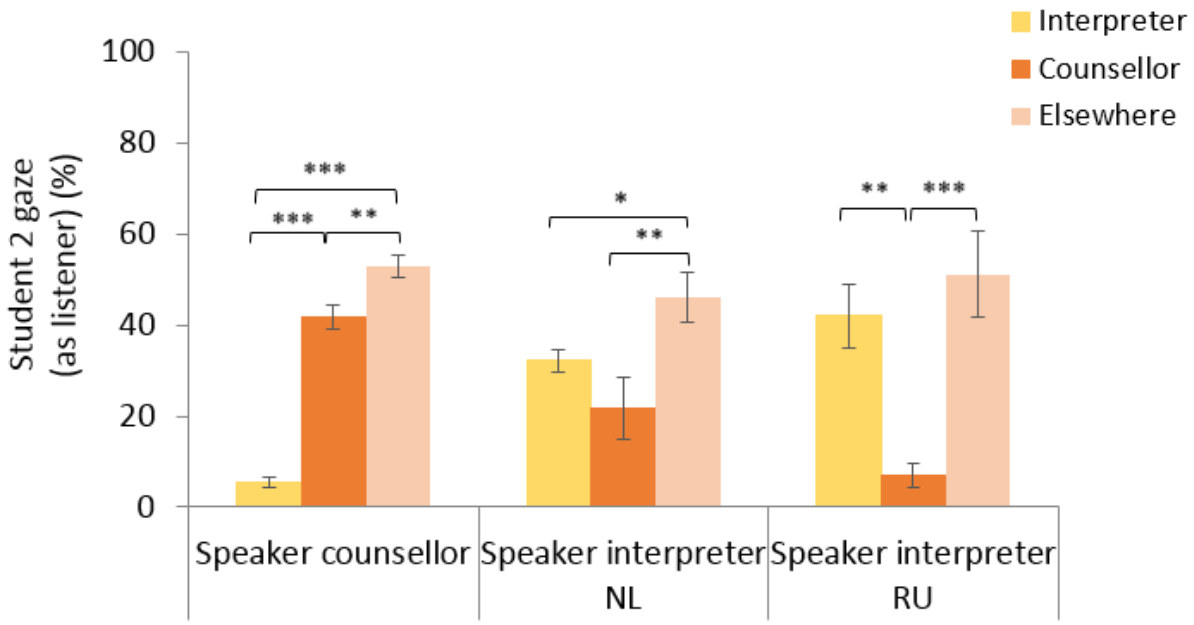


Figure 3.4: Overall proportion of gaze fixations for the three ROI's, averaged across the three conversations. Error bars show the standard deviations. One-way ANOVA test (speaker counsellor: $F(2,6) = 400.40$, $p < 0.001$; speaker interpreter NL: $F(2,6) = 16.72$, $p = 0.0035$; speaker interpreter RU: $F(2,6) = 34.35$, $p = 0.0005$) with Bonferroni Post-hoc analysis (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

For instance, while being addressed by the interpreter in Russian ('interpreter RU'), Student 2 tends to gaze elsewhere (51%). Furthermore, while the interpreter is rendering his turn in Dutch, Student 2 shows a low level of attention (32%), as almost

half of his gaze fixations (46%) are directed elsewhere in space[26]. In contrast to the counsellor (Figure 3.3), Student 2 displays no need to present himself as 'principal' of the ongoing talk nor to monitor the counsellor's nonverbal reactions, but seems to relegate the 'speakership' to the interpreter.

Figure 3.5 below shows the distribution of gaze fixations of Student 3. Student 3 displays much more involvement with the interlocutors than Student 2[27]. During the counsellor's turns, Student 3 looks significantly more at the counsellor (67%) in comparison to the interpreter (29%). She looks at the interpreter most often in cases of overlapping talk between the counsellor and the interpreter. For instance, there are numerous instances in which the interpreter anticipates the completion of the counsellor's turn by producing a rendition slightly in overlap with the counsellor's turn-final TCU's (Sacks et al. 1974), which attracts Student 3's gaze towards her (on the role of gaze in the organization of turn-taking, see Chapter 4)

During the interpreter's renditions of her previous turn ('interpreter NL'), Student 3 gazes almost equally at the interpreter (M=2.916 s) and at the counsellor (M=2.431 s). When the interpreter is addressing the student in Russian ('interpreter RU'), the student divides her visual attention between the interpreter and the counsellor. However, her gaze fixations at the interpreter are of longer duration (M=7.817 s) than those at the counsellor (M=0.946 s). A similar pattern was observed for the counsellors above (Figure 3.3), where it was noted that such gaze behavior has a role in maintaining a *triadic* participation framework (for further discussion on this topic, see Chapter 5).

[26] In the post-interview, the student remarked that when the interpreter was translating his utterance in Dutch, he said that his mind was "switched off", he was "not thinking" about what the interpreter was doing at that moment.

[27] Judging from the scene cameras, Student 1's gaze appears to follow a similar pattern as described for Student 3.

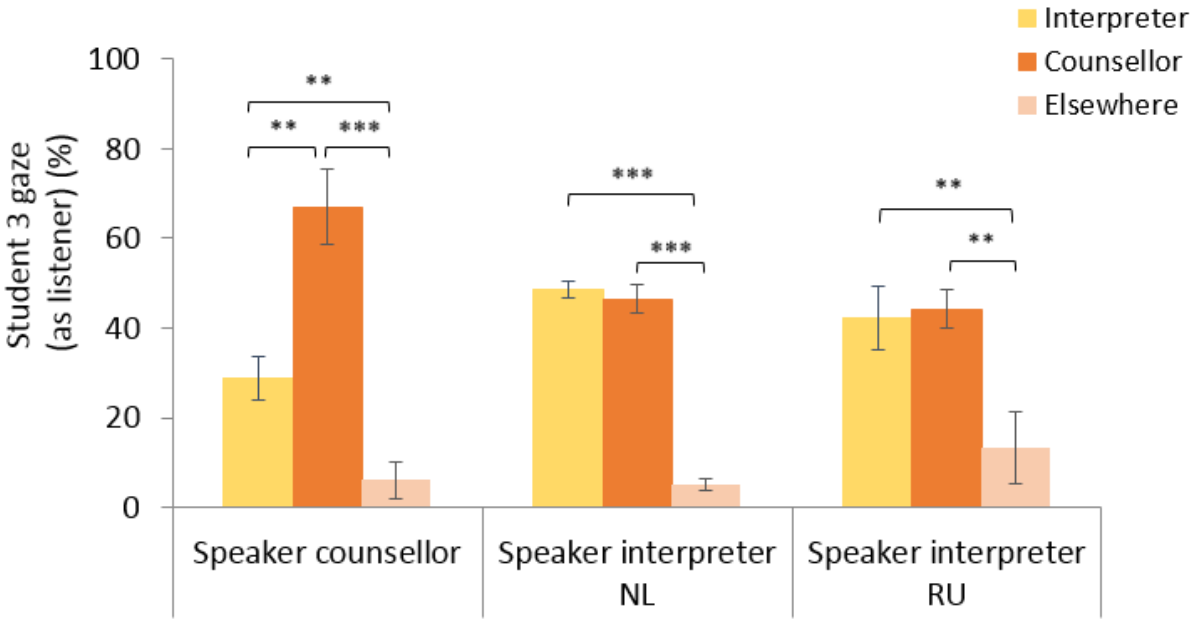


Figure 3.5: Overall proportion of gaze fixations for the three ROI's, averaged across the three conversations. Error bars show the standard deviations. One-way ANOVA test (speaker counsellor: $F(2,6) = 58.14$, $p = 0.0003$; speaker interpreter NL: $F(2,6) = 376.75$, $p < 0.001$; speaker interpreter RU: $F(2,6) = 20.23$, $p = 0.0022$) with Bonferroni Post-hoc analysis (**$p < 0.01$, ***$p < 0.001$).

Altogether, these gaze distributions seem to reflect the primary participants' engagement during the talk and how they position themselves within the exchange. Furthermore, we have seen how Goodwin's distinction between 'hearer'-gaze and 'speaker'-gaze is far from clear-cut in interactions where 'hearers' are somehow implicated or involved in the production of a turn.

3.5. Interpreter's gaze orientation

In the following paragraphs, the focus will be on gaze direction of the interpreters in the recorded dataset.

3.5.1. Interpreter's gaze while speaking

The interpreters in the recorded sessions gaze significantly more at the participant to whom they address their rendition (45% when speaking in Russian with the student and 37% when speaking in Dutch with the counsellor) than to the non-addressed primary participant (10% towards the counsellor when speaking in Russian and 13% towards the student when speaking in Dutch) (Figure 3.6). Thus, the data indicate that interpreters are involved in one dyad at the time. The interpreters look at the currently non-addressed participant in several sequential positions:

- when rendering a very short utterance and projecting more to come from the previous speaker;
- mid-turn, when displaying hesitation about some aspect of their rendition. Sometimes, the gaze shift is combined with a deictic gesture and appears to emphasize to whom the words attest. This is an interesting phenomenon that deserves further investigation in future research.
- before the end of their rendition, when projecting that the previous participants will continue with their extended telling. This pattern will be examined in more detail in Chapter 4.

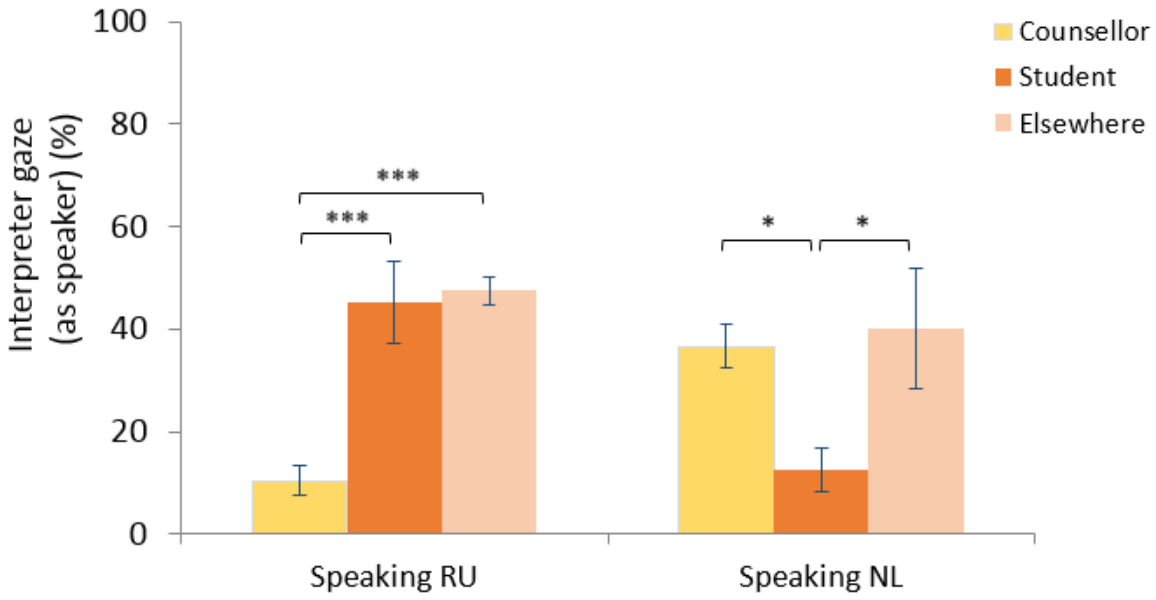


Figure 3.6: Overall proportion of gaze fixations for the three ROI's, averaged across all three interpreters. Error bars show the standard deviation. One-way ANOVA test (interpreter speaking RU: $F(2,6) = 48.84$, $p = 0.0002$; interpreter speaking NL: $F(2,6) = 11.73$, $p = 0.0084$) with Bonferroni Post-hoc analysis (*$p < 0.05$, ***$p < 0.001$).

3.5.2. Interpreter's gaze while listening

The interpreters in the current dataset gaze significantly more at the current speaker than at the other primary participant while listening (Figure 3.7). More precisely, 59% of their gaze is directed towards the student and only 10% towards the counsellor when listening to the Russian-speaking student. On the other hand, the interpreters' gaze is more often directed towards the counsellor (43%) than towards the student (18%) when listening to the Dutch-speaking counsellor. These observations are in line with Goodwin's rule 'hearer should be gazing at the speaker'. However, a relatively high proportion of interpreters' 'away'-gaze (elsewhere) could be noticed during counsellor's turns (40%) and student's turns (31%) (Figure 3.7). This is largely due to the fact that interpreters tend to gaze away from the current speaker just prior to taking the turn (this will be illustrated in more detail in Chapter 4). This observation is further supported by previous research, which argued that the practice of averting gaze can be linked to the activity of searching for a word (Kendon 1967, Goodwin 1980).

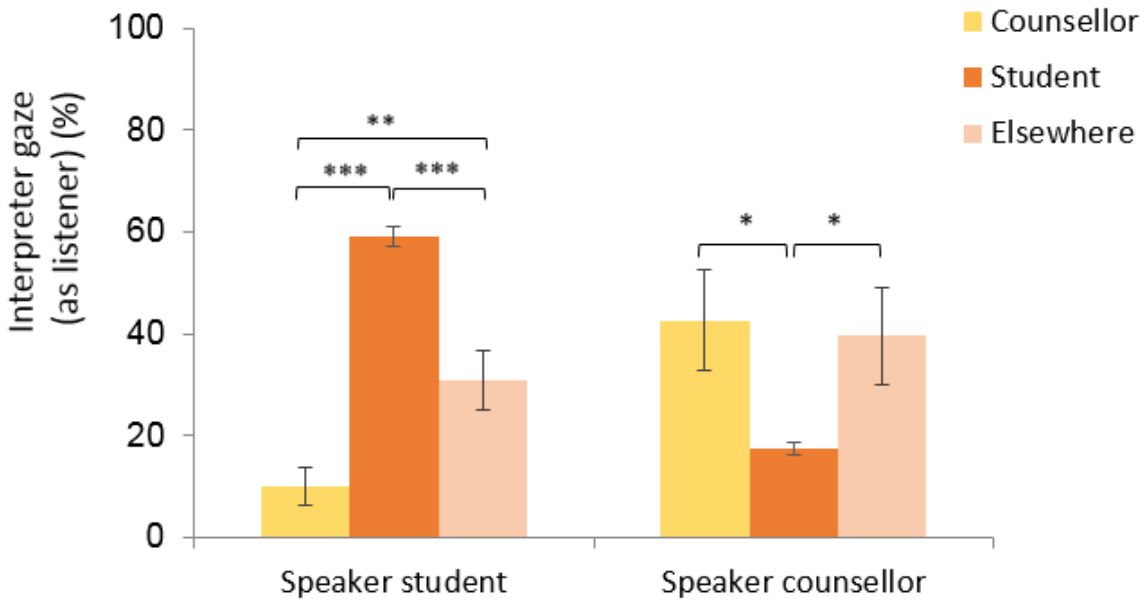


Figure 3.7: Overall proportion of gaze fixations for the three ROI's, averaged across all three interpreters. Error bars show the standard deviation. One-way ANOVA test (interpreter listening to the student: $F(2,6) = 106.56$, $p < 0.001$; interpreter listening to the counsellor: $F(2,6) = 8.96$, $p = 0.0158$) with Bonferroni Post-hoc analysis (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

We also find some instances of the interpreter gazing at the hearer. In those cases, the interpreter shifts her gaze to the addressee before the end of the current speaker's turn, thus anticipating turn completion and displaying readiness to start rendering the talk**.** In this way, they reduce the transition space between the current speakers' and her own turn (see also Chapter 4).

To summarize, the interpreter's gaze patterns in the current dataset are illustrative of (i) the interpreter's engagement with both primary participants during the exchange and of (ii) her pivotal role in shifting from one conversation to the other (see Chapter 4).

3.6. Conclusion

This chapter has provided an analysis of gaze orientations in interpreter-mediated interactions in relation to the interlocutors' participation status (listener or speaker). We have established some general tendencies in the gaze behavior of the participants in the recorded interpreter-mediated exchanges and compared them with findings reported in earlier studies (Bot 2005, Mason 2012, Davitti 2012). The analyses have shown that gaze is strongly linked to the position and professional status of the participants within the exchange (student or counsellor). The more 'powerful' primary participant orients his/her gaze less at the interpreter and much more at the primary addressee. Furthermore, the data have also shown deviations from Goodwin's 'rule' on the organization of speaker's and hearer's gaze, which were accounted for by considering the specificities of this type of exchange. Such global analysis of participants' division of visual attention reveals something about their engagement, positioning and social relations within the exchange.

Nevertheless, the presented analysis tells us very little about *how* interlocutors organize their gaze in relation to the sequential unfolding of the talk and in correlation to other (non-)verbal resources (such as speech and gesture). In order to learn more about the moment-by-moment *organization* of participants' gaze in ongoing talk, we need to examine how gaze is organized in relation to specific interactional practices through detailed analyses of particular cases. In the words of Goodwin & Goodwin: "when the analytic focus shifts to organization of situated activities (...) it becomes possible to recover the cognitive life of the [participant] (2000:225). This will be discussed further in the following chapters, where I will examine gaze as part of turn-taking (Chapter 4) and interlocutors' backchanneling behavior (Chapter 5).

Chapter 4

Gaze and turn-taking in interpreter-mediated interaction

The present chapter examines the relevance of gaze in turn-taking (Sacks et al. 1974) and action projection in interpreter-mediated talk. In particular, it will investigate the process of 'chunking' of extended multi-unit turns and the role of gaze in this process[28]. On the basis of detailed single case analyses it will show that (a) the interpreter's gaze at the end of her turn is employed as a mechanism for next-speaker selection (Auer 2017) and (b) that the interpreter's gaze orientation is guided by her understanding of the ongoing courses of action. Moreover, the chapter will show how interpreter's gaze orientation bears on the negotiation of possible *transition relevance places* and how it contributes to the smooth continuation of the projected extended multi-unit turn.

Chapter 4 starts with a brief literature overview on the organization of turn-taking in same-language interaction and the role of gaze in that process (Section 1), followed by a discussion on the organization of turn-taking in interpreter-mediated interaction (Section 2). Section 3 zooms in on the organization of extended multi-unit turns as the main focus of analysis. Section 4 analyses the role of interpreter's gaze as part of next-speaker selection by comparing its use in two sequential environments. Chapter 4 concludes with a discussion on the implications of this study for our understanding of turn-taking in interpreter-mediated interaction and for the regulatory functions of gaze.

[28] This chapter is largely based on the findings presented in Vranjes et al. (2018c).

4.1. Turn-taking and gaze in face-to-face interaction

The organization of turn-taking in human interactions has been studied extensively by researchers working in Conversation Analysis (Sacks et al., 1974) who are interested in analyzing how interactants manage to coordinate their speaking turns in a finely tuned way. Sacks et al. (1974) observed that generally 'one party talks at a time' and that this is achieved through a tightly organized, moment-by-moment coordination between interlocutors' turns (e.g. Schegloff 2000, Mondada 2007). According to Sacks et al. (1974), every turn consists of one or multiple *turn constructional units* (TCUs), i.e. potentially complete, meaningful utterances, such as sentences, clauses, phrases and words (Clayman 2013). On the basis of syntactic, prosodic, pragmatic and embodied cues, the recipient can anticipate the type of unit that is being produced and foresee its possible completion (Sacks et al. 1974, Lerner 1991, Ford and Thompson 1996, Selting 2000, Auer 2005, Mondada 2007, Stivers & Rossano 2010). The moment in talk where the transition to a next speaker becomes possible is referred to as the *transition relevance place* (TRP). Sacks et al. (1974) defined two types of techniques for turn allocation in spontaneous talk: (i) the current speaker may select someone else to speak next ('current-selects-next') or (ii) the participants can self-select to produce the next turn.

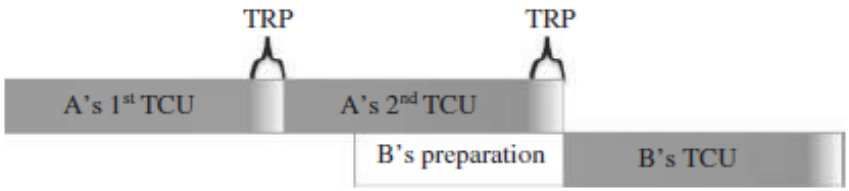


Figure 4.1: Transition Relevance Places (TRP's) (Clayman 2013)

As for *self-selection techniques*, the next speaker can project or claim incipient speakership with an appositional beginning (e.g. *well*, *so*), vocal or non-vocal resources (e.g. an audible inbreath, opening of the mouth), gaze and hand gesture (for an overview, see Hayashi 2013). For example, Mondada (2007) has shown that by pointing before the completion of the current turn, listeners progressively establish themselves as possible

next speakers and thereby publicly display their understanding of the turn-in-progress in real-time (Mondada 2007). In his study on gaze direction in dyadic interactions, Kendon (1967) found that incipient speakers tend to look away just before speaking, which may signal to their interlocutors that they are about to take the floor (for similar results, see Duncan 1972). This result was replicated in an eye-tracking study by Brône et al. (2017).

As for *current-selects-next techniques*, the most effective way for the current speaker to select the following speaker is by producing a first pair-part (e.g. a question) in combination with some form of addressing (an address term and/or gaze direction at the interlocutor) (Hayashi 2013). While address terms are not frequently used in ordinary conversations (Hayashi 2013: 170), selection through gaze is part and parcel of face-to-face interaction. Early studies by Kendon (1967) and Duncan (1972) argued that speaker gaze has a 'floor apportionment' function, in that the speakers tend to shift their gaze away at turn-beginning and gaze back toward the recipient near turn completion in order to signal that they are ready to hand over the floor to the recipient (for similar findings, see Bavelas et al. 2002). More recently, Rossano (2012) reconsidered previous work on the regulatory functions of gaze by claiming that gaze is "not organized primarily by reference to turns-at-talk" and that there is "no evidence that speakers use gaze as a turn-yielding cue"[29]. According to Rossano, gaze is "mainly organized in relation to sequences of talk and the development of courses of action or ongoing interactional projects" (2013, 319). He observed that, for instance, tellings require a more sustained gaze by the recipient toward the speaker than questions. Rossano's claims were supported by Streeck (2014).

However, aforementioned studies primarily focused on gaze conduct in conversations with two participants. Gaze direction is of particular importance for the

[29] Rossano cites several studies, such as Beattie's (1978, 1979), that stated that a speaker's looking away during early utterance production and reengagement at the completion is "occasioned purely by the need to reduce cognitive load and that they do not have any regulating function in terms of turn-taking" (Rossano 2012: 38). Note that those studies were based on two-person interactions. In this chapter I will argue that interactions with more than two persons can provide evidence for such regulatory functions of gaze.

organization of turn-taking in conversations with more than two participants, where it is can be unclear *who* will be the next speaker. Goodwin (1979) and Lerner (2003) have argued that directing gaze at a co-participant in a multi-person interaction is an explicit way of addressing an utterance, but that its success depends on the gazing practices of other participants; both the addressed and the non-addressed participant must see the gaze direction of the current speaker (see also Hayashi 2013). Building on this distinction between currently addressed versus non-addressed participants in multi-person talk, Auer (2017) argues for making an analytical distinction between gaze for *next-speaker selection* and gaze for *addressee selection*. According to Auer (2017: 12):

> addressee selection (by gaze) and next-speaker selection (by gaze) are temporally ordered. While addressees may vary during an emerging turn, the co-participant who is gazed at toward the end of a speaker's turn is the one who is selected as next speaker[30].

At the same time, Auer (2017) acknowledges that gaze alone as a cue for next-speaker selection is less effective than multimodal techniques for next-speaker selection (such as gaze in combination with an address term). Despite the above-mentioned limitation with the use of gaze for the allocation of turns, the distinction between gaze for next-speaker selection and gaze for addressee selection will prove to be particularly useful in the context of interpreter-mediated interactions, as this study will show.

4.2. Turn-taking in interpreter-mediated interaction

The organization of turn-taking in interpreter-mediated interaction is in many aspects different from turn-taking in same-language interactions (Bot 2005, Englund Dimitrova

[30] However, this does not prevent another co-participant from self-selecting, which leaves us to conclude that turn organization is not all about the regulation of speaking rights, but "about qualified opportunities for certain activities" (Schmitt 2005: 81).

1997, Li 2015, Wadensjö 1998). First, turn-taking in interpreter-mediated interaction usually follows a specific pattern: typically, the interpreter takes every second turn in order to provide a rendition of the primary speaker's utterance in the language of the other participant[31] (Bot 2005: 112; see also Englund Dimitrova 1997, Wadensjö 1998). Moreover, as observed by Merlini & Favaron (2005), "the absence or delay of the interpreter's turn would be a noticeable occurrence or, as the case may be, a noticeable non-occurrence" (p. 271). They continue by adding that "the utterance of a primary speaker "sequentially implicates" not only the utterance of the other primary speaker, but, prior to this, the translating act of the interpreter" (ibid.: 271). Although a consecutive interpreter is indeed expected to take the turn in order to provide a rendition, the moment *when* this should occur is not predetermined but managed in a locally situated way (cf. Mondada 2013). Second, although interpreters take part in the conversation, they have different needs and interests concerning turn-taking than the primary speakers: the interpreter usually benefits from shorter primary speakers' turns in order to be able to provide accurate renditions of the preceding talk (Bot 2005: 112). Third, recent studies have shown that the interpreter, by taking turns at talk, is an active participant in the exchange, who is critically engaged in the negotiation of meaning and in the coordination of the interaction (Wadensjö 1998, Bolden 2000, Davidson 2002, Merlini & Favaron 2005, Van De Mieroop & Mazeland 2009, Baraldi & Gavioli 2012, Llewellyn-Jones & Lee 2014). Research has demonstrated that the interpreters – given their understanding of both languages - can resolve overlap (Roy 2000), decide when turn transfer can occur (Bot 2005, Pasquandrea 2011) and give primary participants the opportunity to come out in their tellings (Englund Dimitrova 1997, Gavioli 2012, Vranjes et al. *forthcoming*).

Altogether, these studies show that interpreter-mediated talk is a joint accomplishment between the primary participants and the interpreter, who acts both as oral translator and coordinator during the talk. This also means that primary participants "may not retain control on extended parts of the conversation" and "must yield part of

[31] However, in practice, it may occur that the primary participant responds directly to the other primary participant without waiting for the interpreter's rendition or that the interpreter initiates a dyadic sequence with one of the participants (see e.g. Bolden 2000, Davidson 2002).

their interactional power and share responsibilities for the management of the interaction" with the interpreter (Pasquandrea 2011: 456). Thus, given its complex participation format interpreter-mediated interaction offers researchers unique opportunities for the study of collaborative and multimodal processes of turn-taking and action formation.

Previous research has demonstrated that **multimodal resources** (speech, gaze, body posture etc.) are of crucial importance for the management of turn-taking in consecutively interpreted conversations with an on-site interpreter (e.g. Bot 2005, Davitti and Pasquandrea 2017, Mason 2012, Pasquandrea 2011, Wadensjö 1998). Especially gaze conduct appears to have an important function in this process (Bot 2005, Davitti 2013, Lang 1978, Mason 2012, Pasquandrea 2011). Lang's (1978) pioneering study revealed that by not gazing at the primary speakers, the interpreter can miss out important turn management cues. Lang argued that "constant visual monitoring by the interpreter of his clients, especially at turns or possible turns, is an absolute necessity" (Lang 1978: 241). Further studies have demonstrated the importance of interpreter's gaze direction for the negotiation of transition relevance places (Pasquandrea 2011) and for the allocation of turns (Mason 2012) in interpreter-mediated encounters. However, no systematic conceptualization has been developed on the way in which interpreter's gaze is organized with reference to turn-taking and action formation (Levinson 2013) in consecutively interpreted conversations with an on-site interpreter.

This chapter will examine the relevance of gaze in the local management of turn-taking and in the development of ongoing action in interpreter-mediated interactions by drawing on insights from conversation analysis (Sacks et al. 1974) and multimodal analysis (e.g. Mondada 2007). More specifically, it will focus on the interpreter's involvement in the construction of 'larger projects' or multi-unit turns that are 'chunked' into shorter units in order to allow the interpreter to provide a rendition of the talk piecemeal. Of special interest will be those moments at the end of the interpreter's turn, when, according to Bot (2005), "transfer can take place to any of the primary speakers" (p. 112).

4.3. Extended tellings in interpreter-mediated talk

This chapter focuses on a specific form of sequence organization, namely the production of multi-unit turns or extended tellings (Houtkoop and Mazeland 1985, Ford 2004, Sacks et al. 1974, Schegloff 1982, Selting 2000). Multi-unit turns involve elaborate actions such as extended descriptions, story tellings, pieces of advice and the like (Houtkoop & Mazeland 1985, Selting 2000). In contrast to adjacency pairs, which project a transition relevance place immediately after the first-pair-part, turn-taking during the production of multi-unit turns is temporarily suspended (Selting 2000). The recipient will refrain from taking the turn until syntactically, prosodically, pragmatically and visually recognizable turn completion has been reached (Selting 2000, see also Clayman 2013, Houtkoop & Mazeland 1985). Selting (2000) distinguishes non-final TCU's and final TCU's with "operative TRP's" in the production of multi-unit turns (see Figure 4.2). According to Selting, multi-unit turns are an "interactive achievement in which speakers suspend TRPs, and recipients refrain from making use of suspended TRPs" (2000: 487). However, the recipients can - and are indeed expected to - make contributions during the production of a multi-unit turn, in the form of acknowledgment tokens, assessments or head nods, through which they confirm their recipiency (see Goodwin 1986, Houtkoop & Mazeland 1985, Stivers 2008). Extended tellings also require more sustained gaze by the recipient towards the speaker (Rossano 2012, 124). Therefore, the recipient must be able to recognize or project that a multi-unit turn is in progress, in order to be able to respond in an appropriate way[32]. As noted by Rossano:

> "For this coordination to be successful, the collaboration of the co-participant is not only important but necessary, as s/he could block the production of an extended telling by simply taking the floor at the first transition relevance place.

[32] Rossano (2012: 69) observes that the timing of listener responses "suggests that recipients understand early in a TCU whether it is projecting the beginning of a multi-unit turn, and behave as recipients, at least by not taking the floor".

> This implies that the co-participant must be able to recognize that, at least approaching the transition relevance place, the current TCU is not going to be the last one of the telling but rather there will be further ones." (Rossano 2012: 69)

The speaker employs various practices to signal a multi-unit turn and to secure recipiency of the interlocutors in order to be able to bring the multi-unit turn to its completion (Stivers 2013). Typical practices include list-initiating markers ('first of all', see Schegloff 1982) or story prefaces ('let me tell you something', see Sacks 1974); rising intonation (Duncan 1972, Selting 2000); gestural holds (Duncan 1972) and markers of pragmatic incompletion (Sack et al. 1974, Schegloff 1982) that project turn continuation.

Figure 4.2: A schematic representation of extended tellings in same-language interaction (adapted from Rossano 2012). 'Operative TRP' is distinguished from 'non-operative TRP' (between brackets).

The production of extended tellings is a more complex matter in interactions conducted with the aid of a consecutive interpreter. Extended multi-unit turns may pose a challenge for the interpreters to memorize and reproduce accurately what has been said in the other language, which could then reduce the quality of communication between the primary participants (Li 2015). Therefore, a primary participant may 'chunk' longer multi-unit turns into shorter units (see Figure 2) in order to allow the interpreter to make renditions piecemeal (Bot 2005: 117, Hansen 2016). In contrast to the production of extended multi-unit turns in same-language interaction, where only 'final' TCUs of the telling end in an 'operative' TRP (Selting 2000), during the activity of 'chunking' with an interpreter the current speaker opens up the possibility for the

interpreter to take the turn, while projecting continuation of the telling. Thus, as shown in Figure 5.2, this results (as I will demonstrate) in at least two more projectable, activity-internal TRPs: at the end of the counsellor's TCU and at the end of the interpreter's TCU. Turn transfer to the other primary participant (S2 in Figure 2) typically occurs at the end of the multi-unit turn (Bot 2005).

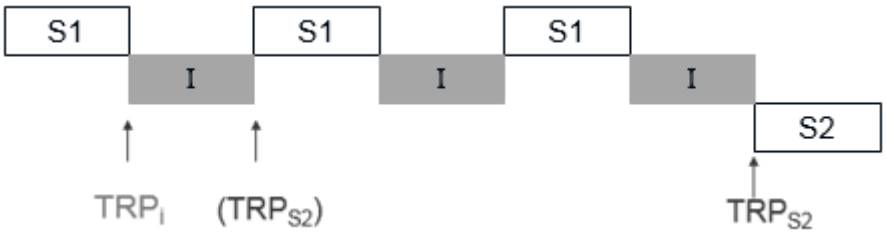


Figure 4.3: A schematic representation of 'chunking' of extended tellings in interpreter-mediated talk.

Chunking of extended multi-unit turns in interpreted encounters reduces the time that the primary participant has to wait for the interpreter's rendition, allowing him or her to respond (e.g. by producing an assessment or a head nod) and thus confirm his or her recipiency during the progression of the multi-unit turn. However, when an extended telling is delivered in shorter units, "each piece is translated as a decontextualized whole" and "participants in conversation (including the interpreter) may draw premature conclusions about the points a speaker wishes to make" (Wadensjö 1998: 234). It is thus essential to maintain a sense of a larger activity in progress. My research questions are the following:

> How is chunking of such longer multi-unit turns multimodally accomplished in interpreter-mediated talk?
> What is the role of the interpreter in this process?
> How does the interpreter display her understanding that a multi-unit turn is in progress?

I am particularly interested in the moments at the end of the interpreter's turn when, since the prior speaker whose talk is being translated lacks knowledge of the other language, (s)he has less control over the turn transfer between the interpreter and the recipient. Therefore, close collaboration and understanding between the primary speaker and the interpreter is crucial in this process (see also Merlino 2014). The importance of smooth chunking of multi-unit turns is demonstrated by Licoppe et al. (2018) in court proceedings with asylum seekers where one of the parties is interacting via video link; if chunking does not run smoothly between the asylum seeker and the (remote) interpreter, this can lead to negative inferences on the part of the asylum seeker[33]. I will now examine in detail how chunking in face-to-face interpreter-mediated dialogues is accomplished in real-time and what is the role of gaze in this process. I will focus on those moments at the end of the interpreter's turn where, according to Bot, "turn transfer can take place to any of the speakers" (2005: 112). However, Bot did not take into account the sequential context in which turn transfer takes place. By comparing the interpreter's gaze conduct in two sequential environments (first pair-parts of adjacency pairs and non-final TCUs of an extended telling in progress), I will analyze the role of gaze in the process of action projection and turn-taking in interpreter-mediated talk.

[33] Licoppe et al. (2018) observe that, if chunking does not run smoothly between the interpreter and the asylum seeker, then the latter becomes "potentially vulnerable to negative inferences and characterizations (…) because s/he can be viewed as breaching the kind of fundamental trust involved in the orderly production of a recognizable courtroom interrogation sequence". Furthermore, the 'ecology of interaction' is important in this respect: Licoppe et al. (2018) show that if the interpreter is physically distant from the speaker (via video link), then s/he might be more inclined to use more obtrusive multimodal resources for explicit management of turns.

4.4. Analysis

The following section starts with an analysis of the interpreter's gaze during the production of adjacency pairs.

4.4.1. First pair-part of an adjacency pair: gaze at the selected next speaker

In contrast to extended tellings, where turn transfer between interlocutors is temporarily suspended, first pair-parts such as questions project a responding action of a particular type (Sacks et al. 1974). By responding, the recipients display their understanding of the preceding action. In the present dataset, interpreters always orient their gaze toward the addressee and expected responder while rendering a question (see also Stivers & Rossano 2010). An example of this pattern is shown in Extract 1. Note that all figures in the extract are screenshots from the eye-tracking camera worn by the interpreter and thus represent the latter's gaze direction (indicated with a colored dot). The transcription symbols can be found in the annex. Prior to the excerpt, the student (STU) has been telling that he will be going abroad for a research stay during the summer.

Extract 1

1 **CNS** **wanneer# zou je terug hier #zijn?**
when would you be back here?
cns gaze at Stu----------> 1.3
int gaze at Cns----------gaze away->
fig #fig. 1 #fig. 2

figure 1

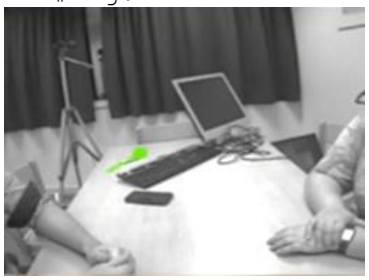


figure 2

2 **INT** **а ког#да вы вернётесь в Бельгию?**
and when are you coming back to Belgium?
int -----gaze at Stu---------->

fig #fig. 3

figure 3

3 **STUD** **#Я должен приехать на: пару недель а: в июле**
I have to come for a couple of weeks uh in July
int ----------table--gaze at STU---------->
fig #fig. 4

figure 4

In line 1, the counsellor (CNS) asks a question while gazing at the Russian student, who is the addressee and the expected responder. Already the first word (the adverb 'when') allows the interpreter (INT) to ascertain what type of action the counsellor is initiating. Even before completion of the question, the interpreter starts to shift her gaze to the table in front of her (Figure 2), thus anticipating the end of the current turn and displaying her readiness to take the floor. Such early gaze shift provides evidence for the interpreter's online processing of the speaker's utterance (on 'projection', see e.g. Auer 2015; for unaddressed participant's anticipatory gaze shifts, see Holler & Kendrick 2015). In line 2, the interpreter reproduces the question while gazing at the student (Figure 3). Gazing at a particular co-participant is treated as an 'explicit' method of addressing (see Hayashi 2013, Stivers & Rossano 2010). The interpreter thus displays her alignment with the projected action of the previous turn through her gaze orientation; by maintaining her gaze at the student while she orients to the relevance of his response (line 3). In this sequential position, it might be difficult to maintain Auer's analytical distinction between speaker-gaze for addressee selection and speaker-gaze for next-speaker selection, as the student is both the addressee and the

expected next speaker. Against this background, we will now turn to the interpreter's gaze conduct in the process of chunking of multi-unit turns.

4.4.2. Chunking of multi-unit turns

The results reported in this chapter are based on an examination of 48 cases in which a primary speaker's multi-unit turns were 'chunked' into shorter units. In the following sections, I will show that the interpreter's gaze direction at the end of the turn is guided by her understanding of the ongoing course of action.

4.4.2.1. Immediate gaze shift to the previous speaker at turn-end: Close collaboration

In the following, we focus on the interpreter's gaze conduct during stepwise production of multi-unit turns. Most multi-unit turns in our data were initiated by a question of the other primary speaker ("second-position tellings", see Mandelbaum 2013). In the following extract, the counsellor is engaged in an explanation about wedding ceremonies in Belgium. Again, the figures in the extract are screenshots from the eye-tracking camera worn by the interpreter and thus represent her perspective.

Extract 2

1 **CNS** **dus er wordt eerst een=uh (0.2) je hebt ten #eerste #al**
so there is first an uh you have first of all
cns gaze at Stu--wall----------->
int gaze at Cns----------Cns hand-
fig #fig. 1 #fig.2

figure 1


figure 2

2 **(0.3) twee ceremonies,**
two ceremonies,

cns ------gaze at Stu----->
int Cns------------------>

3 ***(part of the transcription omitted)***

4 **en (.) je moet (.) uiteraard burgerlijk ook trouwen**
and you have to obviously also have a civil wedding

cns gaze down----gaze at Stu--------------------------->
int gaze at Cns-->

5 **in het (.) #stadhuis of in het ge#meentehu#is.**
at the city hall or at the town hall.
cns --->
int --gaze away->
fig #fig. 3 #fig. 4 #fig. 5


figure 3


figure 4


figure 5

6 **INT .h здесь тоже две: є: церемонии**
.h here are also two uh ceremonies

cns -----gaze at Int----------gaze at Stu->
int ------------------------------>

7 ***(part of the transcription omitted)***

8 **но еще плюс и в з#агсе.**
but in addition also at the Records Office.

cns --gaze at Int--Stu--Int---Stu->
int ----------gaze at Stu-----gaze away->
fig #fig.6

figure 6

9 **(0.2) в загсе это обычно є: или в ратуше где-то в г#ороде.**
at the Records Office usually uh or in the town hall somewhere. in the city.
cns ------------------------gaze at Stu-----gaze at Int—Stu->
int -------gaze at Stu--------away--Stu----------------Cns--->
fig #fig. 7

figure 7

10 **(0.4)# (0.3)**

---------->
---------->
#fig. 8

figure 8

11 **CNS** **en vaak eer#st 't stadhuis en d#an de kerk,**
and often first in the city hall and then the church,
cns --wall------Stu---------->
int ----------gaze away----->
fig #fig. 9 #fig. 10

figure 9

figure 10

12 **(0.2)#**

cns ------>
int ------>
fig #fig. 11

figure 11

```
13  INT   и обычното в ратуше, а потом ц#ерковь,  (…)
          and usually in the city hall, and then the church,
    cns   --gaze at Int---Stu------------------------------>
    int   ------------------------------gaze at Cns-------->
    fig                                 #fig. 12
```

figure 12

In line 1, the counsellor (CNS) starts with a list-initiating marker ‘first of all’ (see also Schegloff 1982) accompanied by a co-speech counting gesture (Figure 1) that signals to the recipient that an extended description will follow (see also the interpreter’s gaze shift towards the counsellor’s hands in Figure 2). From lines 1 to 5, the counsellor produces his turn while looking at the student[34]. All the while, the interpreter maintains her gaze at the counsellor (Figure 1). Towards the end of his turn (line 5), which is produced with a falling intonation contour and with a retracting gesture (Figures 3-4) (see also Duncan 1972, Kendon 1967), the interpreter shifts her gaze toward a middle position (Figure 5) and produces an audible inbreath. Her gaze shift together with the audible inbreath clearly belongs to the pre-onset phase of her turn (Schegloff 1996, Mondada 2007) and projects her incipient speakership. In line 6, she starts rendering the counsellor’s utterance into Russian. Interestingly, in contrast to what we have seen in Section 4.1, she does not maintain her gaze at the student, who is the addressee of her turn, but immediately shifts her gaze back to the counsellor (Figure 7). Following the

[34] The counsellors in the present dataset mostly gaze at the student (who has no understanding of Dutch) while speaking. In this way, they display orientation towards the student as the addressee of their talk. At the same time, they display disengagement from the interpreter, who is expected to take the next turn. By not gazing at the interpreter during and at the end of their turn, the counsellors are reducing their speakership to that of a conduit or a ‘translation machine’. (see Vranjes et al. 2018 for similar observations in the context of an interpreter-mediated therapeutic encounter)

interpreter's gaze movement, the student then turns toward the counsellor who is already looking at her (Figure 7)[35].

Thus, by shifting her gaze, the interpreter makes it publicly visible that she orients towards continuation of the counsellor's turn and thereby selects him as next speaker. In this sequential environment, her gaze shift works as a kind of 'return signal', produced to display her ongoing understanding "that an extended unit is in progress and is not yet completed" (Schegloff 1982: 84). At the same time, by shifting her gaze before the end of her turn, the interpreter is in a way *blocking* the completion points of her turn from being treated as transition places by the student (see also Houtkoop & Mazeland 1985, Selting 2000). In this way, the interpreter is closely collaborating in the production of the extended telling.

During the pause in line 10, the counsellor starts lifting his hand again (Figure 8) and establishing himself as a speaker (see also Mondada 2007). The conjunction 'and' at the beginning of his TCU (line 11) makes it recognizable as an increment of the turn-so-far and marks continuity with the previous utterance. The counsellor supports the chronology of the described events with hand gestures: 'first' (pointing to the left, Figure 9) and 'then' (pointing to the right, Figure 10). All the while, he gazes at the student, who is the principal recipient of his telling. Again, the interpreter anticipates a possible transition relevance place in line 11 by shifting her gaze to a middle position (Figure 11). She starts speaking after a slight pause (line 12), during which the counsellor retracts his gesture. Although the lowering of the gesture indicates the end of the TCU, he finishes his utterance with a rising intonation contour, projecting 'more to come'. Interestingly, the student seems to anticipate turn continuation as well, as she keeps her head oriented towards the counsellor (see Figures 11-12) during the interpreter's short rendition in line 13[36]. Once again, the interpreter shifts her gaze to the counsellor before the completion of her rendition in line 13.

---

[35] In most cases, the student's gaze shift towards the counsellor follows the interpreter's gaze shift. As such, the interpreter's gaze shift appears to function as a gaze cue for the student's shift of attention towards the counsellor (on 'gaze cueing', see Frischen et al., 2007).

[36] This was not a consistent pattern in my data. In other cases, the student gazes at the interpreter when the interpreter is speaking to her.

Thus, chunking is a collaborative achievement between the speaker and interpreter. The interpreter relies on a combination of available cues (syntactic structure, intonation, pragmatics and gesture deployment) in deciding when possible completion of the current speaker's TCU has been reached and when to take the turn. It can also be observed that through her gaze shift, the interpreter projects both the incipient end of her own turn and continuation of the previous speaker's telling.

Extract 3 further illustrates the phenomenon. Here the speaker selects the interpreter as the next speaker by directing her gaze towards her[37].

Extract 3

1 **CNS** **er is in vlaanderen de laatste jaren heel wat veranderd**
there has been a lot of change in Flanders in recent years
cns gaze at wall----------Stu-->
int gaze at Cns---------->

2 **in het hoger onderwijs,**
in the higher education,
cns ---------->
int ---------->

3 **CNS** **dus bij de# hoge scholen en #de universitei[ten],**
so at colleges and universities,

4 **INT** **[mm hm,**
cns ----------gaze at Int------->
int ---------->
fig #fig. 1 #fig. 2

figure 1


figure 2

5 **INT** **.hh#**
cns -------------->
int -gaze away---->

[37] I found mutual gaze between the current speaker and the interpreter in 52% of turn transitions during chunking, and no gaze contact in 21% of the cases. In other instances, the speaker looks at the interpreter shortly before (6%) or right after the end of his turn (17%), after the interpreter has already shifted her gaze to the other interlocutor in order to start rendering the turn.

fig #fig. 3

figure 3

6 ***(part of the transcription omitted)***

7 **INT** **в Фландрии є: очень# много изменилось**
in *Flanders uh very much has changed*
cns gaze at wall---------------gaze at Int->
int gaze away----------gaze at Stu--------->
fig #fig. 4

figure 4

8 **что касается высшего образования-**
regarding higher education-
cns ------------------------------->
int ----------------------------gaze away-->

9 → **є: институтов и универси#тетов,**
uh the institutes and the universities,
cns ------------------------------->
int ---Stu------away--Stu---Int---->
fig # fig. 5

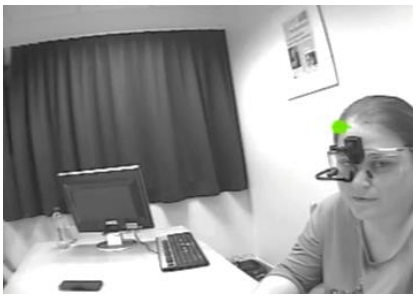


figure 5

10 **(.)**

11 **CNS** **.h uhm dus vroeger (…)**
.h uhm so before (…)
cns --wall---------------->>
int ---------------------->>

The counsellor is explaining here to the student why there are two universities in the city of Antwerp. By starting with the observation that a lot has changed in the organization of higher education in Flanders, the counsellor leads the listener to expect more information on the nature of those changes. The screenshots in the transcript represent the interpreter's gaze direction. From lines 1 to 3, the counsellor addresses her utterance to the Russian student, while the interpreter is looking at her (Figure 1). Towards the end of her utterance in line 3, she moves her gaze toward the interpreter (Figure 2), thereby inviting her to take the turn. At the same time, the pragmatic incompletion and rising intonation of the counsellor's turn indicate that her explanation will be continued ('at colleges and universities then,').

Note that the counsellors in the present dataset rarely use recipient reference terms ('you') (Lerner 2003) to select the interpreter as next speaker. In this context, the use of verbal address by the counsellor would disrupt the progress of the multi-unit turn and bring the interpreter's presence more to the foreground. Therefore, a smooth turn transfer between the counsellor and the interpreter during the process of chunking necessarily depends on the interpreter's projection of the upcoming transition space and continuous understanding of the ongoing action.

In line 4, the interpreter projects possible completion by producing an acknowledgment token. During an audible inbreath (line 5), she shifts her gaze to a middle position as a way of displaying her incipient speakership. Comparable to extract 2, the interpreter projects turn continuation by moving her gaze to the expected next speaker (the counsellor) towards the end of her turn-in-progress.

What is then the quantitative distribution of this pattern? A quantitative analysis reveals that in 90% of non-final TCU's[38] of projected multi-unit turns, the interpreter shifts her gaze to the previous speaker even before the end of her current TCU. I annotated the timing of those early gaze shifts (defined as the difference (in milliseconds) between the time that the interpreter's turn-final gaze[39] shift starts and the

[38] Following Selting (2002) a TCU was coded as non-final if the utterance has reached syntactic and prosodic completion, but further talk is projected syntactically (e.g. in case of a when-then construction), pragmatically and prosodically (e.g. rising intonation).

[39] This means that only the final gaze shift of the final TCU at turn completion was taken into account.

moment that the interpreter's turn-in-progress ends) in ELAN and found that the average timing of the interpreter's gaze shift was -645 ms (median =-464 ms; minimum value = 0 ms; maximum value = -1864 ms). If we assume that speakers need about 600 ms to plan their speech (Levinson & Torreira 2015), then such an early gaze shift by the interpreter may contribute to the shortening of the transition space between the interpreter and the primary speaker. Furthermore, through those gaze shifts the interpreter is in a way blocking the completion points of the TCU's from being treated by the recipient as normal transition-relevance-places (cf. Houtkoop and Mazeland 1985, Selting 2000) and thus project the continuation of the extended telling.

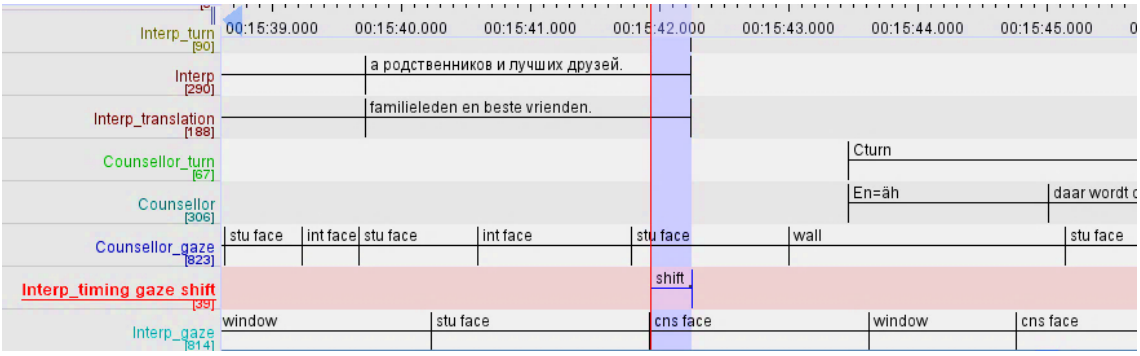


Figure 4.4: A visual representation of gaze and timing measurements during chunking

To summarize, this analysis has shown how the interpreter projects turn-transition and continuation of the developing action with her gaze, ensuring in this way a smooth continuation of the prior speaker's multi-unit turn. Lang (1978) observed in this respect that it is important for the interpreter to maintain visual contact with all primary participants, "but especially with the one who is talking" (1978: 233), or in this case, who is expected to resume the talk. The following sections will address the issue of normativity of interpreter's gaze behavior by presenting two deviant cases.

4.4.2.2. No gaze shift to the previous speaker at turn-end

4.4.2.2.1. The selected speaker does not take the turn

As shown in the previous extracts, the interpreters closely collaborate in the construction of a 'chunked' extended telling by anticipating transition relevance places and projecting

turn continuation towards the end of their rendition, as becomes evident through their gaze behavior. In the following extract, which is a continuation of Extract 3, I present a deviant case[40] in which continuation of a multi-unit turn is projected by the speaker, but not by the interpreter. This can lead to a temporary incongruity in the progression of the projected multi-unit turn. Note that the screenshots first represent the interpreter's and then (from line 6) the counsellor's perspective.

Extract 4

1 **CNS** **dus vroeger had #je (.) de universiteiten en de hogescholen,**
so before you had the universities and the colleges,
Cns gaze at wall-----gaze at Stu---------wall----------->
Int gaze at Cns--------------------away--Cns----------->
fig #fig.1

figure 1

2 **en op die hogescholen werden ook een aantal opleidingen**
and at colleges also a number of courses were
Cns ----------------------------------Stu----------->
Int --->
3 **van academisch niveau #aangebod[en.**
offered at academic level.
4 **INT** **[mh hm**
Cns ----------------------wall--------->
Int ----------------------away--------->
fig #fig. 2

figure 2

5 **INT** **[((opens mouth))**

[40] Deviant case analysis is a common method in Conversation Analysis. It is through deviant cases that participants' own orientations to a 'norm' become visible.

6 **CNS** **[bijvoorbeeld #de opleiding v#ertaler-tolk#.**
for example the training for Translator-Interpreter.

```
Cns   -------------------Stu-------Int---------Stu->
Int   -----------Cns-----------------------away--->
```

#fig. 3 #fig.4 #fig. 5

fig

figure 3 (INT gaze)

figure 4 (CNS gaze)

figure 5(CNS gaze)

7 **INT** **раньше є: существовали университеты и институты,**
before uh: there were universities and institutes

```
Cns   Int-----------------------wall----Int---------->
Int   ------gaze at Stu------------------------------>
```

8 **(part of the transcription omitted)**

9 **и институт переводчиков#.**
and institute for translators.

```
Cns   gaze at Int------------>
Int   gaze at Stu------------>l.12
Stu   gaze at Int------------>
```

fig #fig. 6

figure 6 (CNS gaze)

10 **(0.7)#**

```
Cns   ------>
Stu   gaze down->l.13
```

fig #fig. 7

figure 7 (CNS gaze)

11 **CNS** **.h #**
cns gaze away-->
fig #fig. 8

figure 8 (CNS gaze)

12 **CNS** **(0.5)#**
cns -------->
int -----gaze at Cns->
fig #fig. 9

figure 9 (CNS gaze)

13 **CNS** **en dan uh is er een wijziging geweest (…)**
and then uh there was an alteration (…)
cns ------------------------------gaze at Stu-->
int -->
stu ----------gaze at Cns---------------------->

The counsellor starts with the explanation of the situation 'before' (in line 1) while looking at the student. In line 3, the counsellor's utterance ends with a falling intonation contour, which is treated by the interpreter as turn completion; she displays her incipient speakership with the token 'mh hm' (l. 4) while opening her mouth and shifting her gaze towards a middle position (Figure 3). The interpreter's gaze shift is thus part of the pre-onset phase of her turn. However, the interpreter immediately abandons her incipient speakership as the counsellor continues with an expansion ('for example…' at line 6). During this expansion, the counsellor quickly looks at the interpreter and then back at the student (Figures 4-5). By looking at the interpreter, she projects an imminent completion point and displays her orientation towards the

interpreter as next speaker (for a similar observation, see Extract 3). Although the counsellor's utterance is syntactically and prosodically complete, her explanation about the current situation has not yet reached its pragmatic completion.

In line 7, the interpreter starts rendering the counsellor's utterance, while being engaged in mutual gaze with the student. At the end of her turn, the interpreter maintains her gaze at him, thus selecting him as the next speaker (Figure 6). There is a notable pause at line 10, during which the student moves his gaze towards the table in front of him; he appears to disengage from the conversation and signal that he is not going to take the turn. The interpreter's gaze orientation opens up the opportunity for turn-transition and thus makes recipient reaction relevant[41]. The pause turns into a place "where a response could have been produced but was not" (Couper-Kuhlen & Ono 2007: 514). The student's gaze withdrawal from the interpreter has not so much to do with a "bid for closure" (Rossano 2012: 320) in this context, but rather with his refusal to take up the turn after being selected to respond by the interpreter through gaze. The interpreter maintains her gaze at him (Figure 7), possibly in an attempt to elicit a response (Stivers & Rossano 2010). In this way, she appears to be reshaping the trajectory of the primary speaker's talk and projecting a transition relevance place[42]**.** This example nicely illustrates how gaze as a turn-allocation technique is an inviting instead of obliging technique for next-speaker selection (Auer 2017).

After registering that the interpreter is not oriented towards her (Figures 6-7)**,** the counsellor self-selects by producing an audible inbreath (Schegloff 2000) in line 11 while shifting her gaze to a middle position (Figure 8). After a slight pause (line 12), the interpreter eventually moves her gaze toward the counsellor (Figure 9). The counsellor's turn-beginning coincides with the moment when the interpreter's is oriented at her (line 13). This example nicely shows how speakers may display sensitivity to the gaze (and head) orientation of the interpreter in the process of chunking.

---

[41] As noted by Levinson (2013) the type of action that is being done by the turn "is revealed by the response of a next speaker".

[42] The example provides a nice illustration of the 'transformative' (Goodwin 2013) power of gaze in the process of action formation. It illustrates how the interpreter's gaze opens up the opportunity for a response by the student.

As illustrated in this extract, the interpreter is doing more than just translating the talk; she is involved in the progression of the multi-unit turn through her gaze.

4.4.2.2.2. The selected participant takes the turn

The following excerpt illustrates another case in which the interpreter does not immediately shift her gaze back to the previous speaker who has projected a multi-unit turn, but displays her orientation to the recipient as next speaker. However, in contrast to the previous example, the recipient (STU) takes the opportunity to take the turn and initiate a question.

Extract 5

1 **CNS** **die krijgen bijvoorbeeld hé #van studentenvoorzieningen**
they get for example eh? from the student services
Cns gaze at Stu----------gaze at laptop->
fig #fig. 1

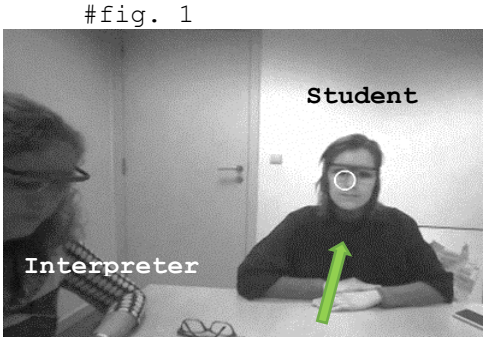


figure 1 (CNS GAZE)

2 **ik heb het #eventjes (.) op#engezet ook voor jou,**
I've just opened it (.) for you as well,
Cns ----------gaze at Stu---------->
fig #fig. 2 #fig. 3

figure 2 (STU GAZE)

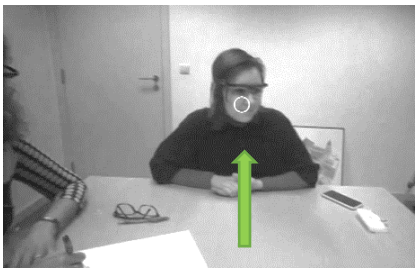


figure 3 (CNS gaze)

3 **(0.4) die krijgen een aantal mogelijkheden aangeboden**
(0.4) they are offered a number of options
Cns ---gaze at laptop----Stu--Int--Stu---------->

4 **en dat is ook toegan#kelijk voor de (.)# Erasmusstudenten.**
and those are also accessible to the Erasmus students.
Cns ----------gaze at Int---------->
Int gaze at laptop-----gaze at Cns----------gaze down->
fig #fig. 4 #fig.5

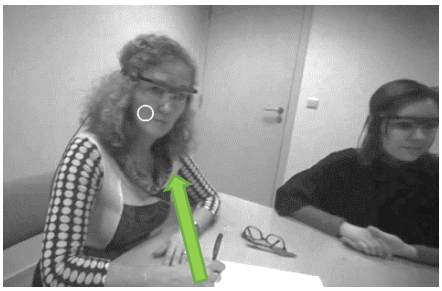


figure 4 (CNS gaze)


figure 5 (INT gaze)

5 **INT** **#mh hm:?**
Cns ------->
Int ------->
fig #fig. 6


figure 6 (CNS gaze)

6 **CNS** **maar #misschie[n #weer ef]kens ((lacht))**
but maybe again for a moment ((laughs))
7 **INT** **[Mh hm]**
Cns ----------gaze away-->
Int ----------gaze at Cns---------gaze down---------->
fig #fig. 7 #fig. 8


figure 7


figure 8

8 **INT** **oké**
okay

9 ***(part of the transcription omitted)***

10 **вот# я приготовила для вас посмотрите какие є:**
here I've prepared for you see which uh:
Cns gaze at Int----------gaze at Stu----->

fig #fig. 9-10

figure 9 (CNS gaze)

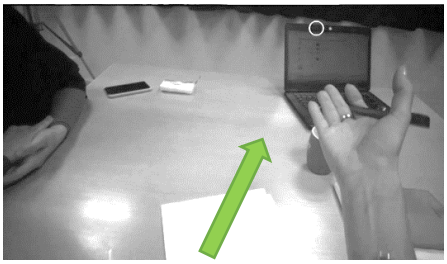


figure 10 (INT gaze)

11 **службы у нас в нашем распоряжении есть**
services we have at our disposal
Cns ---------->

12 **ч[то мы можем] предложить, (0.2) так же мы стараемся е**
what we can offer, (0.2) we also try to
13 **STU** **[mh hm]**
Cns ------gaze at Int-----gaze at Stu------------------>

14 **INT** **ознакомить иностранных студентов с этими воз#можностями.**
notify the foreign students about these opportunities.
Cns ------------------------------gaze at Int---------Stu->l.19
Int gaze at laptop------------------------------Stu------>l.18
fig #fig. 11

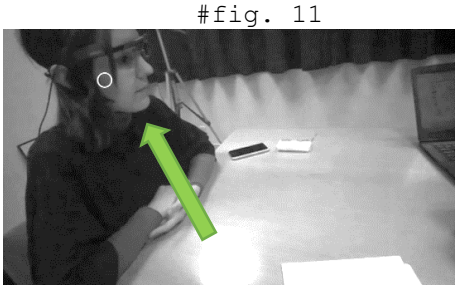


figure 11 (INT gaze)

15 **(0.4)**
16 **STU** **+ mh +**
stu +nodding+

17 **CNS** **#((opens mouth))**

fig #fig. 12

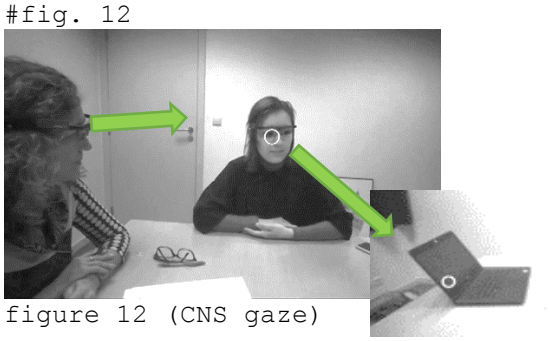


figure 12 (CNS gaze)

18 **STU** **интересно. #это только: в этом университете?**
interesting. is it just at this university?
Int ---------->

fig #fig. 13

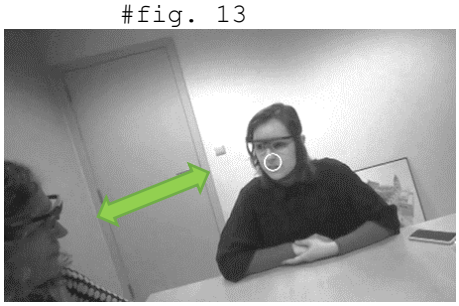


figure 13

19 **INT** **(.) dat is alleen hie:r? in een (.) uh**
(.) that is only here at a (.) uh
Cns ----------gaze away-->
Int --gaze at laptop----------gaze at Cns->

20 **universiteit van Antwe#rpen?**
university of Antwerp?
Cns ----------gaze at Int-->
Int ---------->
fig #fig. 14

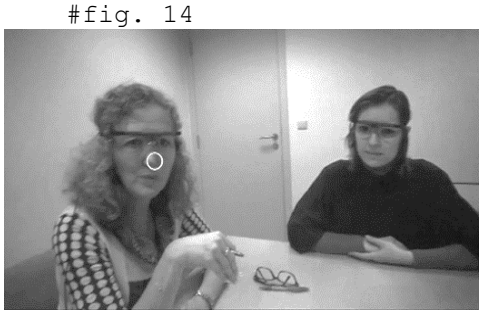


figure 14

21 **(0.3)**

22 **CNS** **uhm nee:n? (.) allé (.) ik zal heel in het kort**
uhm no? (.) well (.) I shall very briefly
Cns --gaze at Stu--gaze at laptop----------gaze at Stu->

23 **eventjes uitleggen,**
explain just for a moment
Cns ---------->

Prior to this excerpt, the student had asked how the integration of foreign students proceeds at the university. The counsellor is explaining here that the university organizes various student activities that are open for foreign students as well. At one point during her explanation (in line 2), she turns her laptop screen with the opened web page of the university student service towards the student (Figure 2). The student maintains her gaze at the laptop screen until the end of the counsellor's utterance. Near the end of her TCU (line 4), the counsellor looks at the interpreter, thereby inviting her to take the floor. This is supported through her deictic gesture (Figure 5).

Instead of taking the turn, the interpreter produces a continuer (‘mh hm?’) and gazes down at her paper to take notes (Figure 6). Since the interpreter does not react in the expected way (by taking the turn) and the continuer appears to signal her continuing recipiency, the counsellor then verbally invites the interpreter to provide a rendition (line 6) by suggesting that she could perhaps render the talk again briefly (‘but maybe again for a moment’) in combination with gaze and gesture towards her (Figure 7). Thus, before expanding on those options that are ‘also accessible for Erasmus students’, the counsellor gives the interpreter the opportunity to render the preceding utterance. The interpreter immediately displays her acceptance of the offered turn by uttering an acknowledgment (line 7) and turning her gaze to the counsellor (Figure 8)[43].

During her rendition of the preceding unit, the interpreter gazes from the student to the laptop screen (Figure 10). She explicitly orients the student’s attention to the screen as the referent of the talk through her gaze and gestures (Figure 9-10). The interpreter ends her rendition in line 14 with falling intonation contour, while looking at the student, who maintains her gaze at the laptop screen (Figure 11). Note that the interpreter’s TCU is syntactically and prosodically complete at that moment and contains no indications that more will follow. She appears to select the student as next speaker by gazing at her. After a slight pause in line 15, the student produces an acknowledgment token (‘mh’) accompanied by nodding, while maintaining her gaze at the screen (Figure 12). In contrast to extract 4, the student’s gaze at the laptop is not a sign of disengagement from the conversation, but appears to project her upcoming question. The interpreter displays her orientation towards this possibility, as she provides a slot for the student to offer a reaction to what has been said. We see that the interpreter in this example is closely monitoring the student’s behavior for any cues of incipient speakership.

In line 17, the counsellor appears to self-select by opening her mouth, but abandons taking the turn; she also seems to forestall an upcoming question by the

[43] The immediacy of the interpreter’s response (‘mh hm’, line 7) indicates that the interpreter had already understood that she may take the turn and that she interprets the counsellor’s suggestion ‘but maybe again for a moment’ as a *directive* to start interpreting. The counsellor then produces a laugh particle, probably to attenuate this turn transition.

student. It is the student who finally takes the turn by producing an assessment (‘interesting’), followed by a question (‘is it just at this university?’).

4.5. Conclusion

This chapter has focused on the multimodal process of chunking of extended tellings in interpreter-mediated dialogues with an on-site interpreter. Chunking in interpreter-mediated face-to-face encounters is necessarily accomplished through finely tuned coordination between the current speaker and the interpreter (see also Licoppe et al., forthcoming). We have seen how the counsellor chunks his or her projected multi-unit turn to open up the opportunity for the interpreter to take the turn and how the interpreter monitors the counsellor’s turn for such moments of possible completion (as evidenced, for instance, through her early gaze shifts). The question is if such places are still to be treated as TRPs (Selting 2000) or are they to be seen as mere temporary ‘suspensions’ (Hansen 2016) of the turn in progress. They do result in speaker change, but they are not meant to disrupt the overall progression of the larger activity. The presented analysis points in favor of the former view, since rules of turn allocation as proposed by Sacks et al. (1974) - viz. next-speaker selection and self-selection - become relevant at those places. Thus, the counsellor and the interpreter jointly accomplish moments of turn transfer between them. What makes ‘chunking’ possible, then, is participants own orientation to the progression of the larger activity (Selting 2000: 512).

The main focus of the analysis was on the role of the interpreter’s eye gaze for the management of turn-taking in interpreter-mediated talk. By using Auer’s (2017) analytical distinction between gaze for addressee selection and gaze for next-speaker selection in the specific context of ‘chunking’, I have demonstrated how the interpreter is involved both in the process of *turn allocation* and *action formation* during the talk. We have seen how the interpreter’s gaze at the end of her TCU may select the participant who is expected to speak next. Most importantly, the interpreter does not select just any of the primary participants to speak next, but the one with the sequentially-relevant speaking rights: within the context of an adjacency pair, it is the producer of a second-

pair part, and within the context of an extended multi-unit turn, it is the producer of the ongoing telling. Furthermore, the interpreter's gaze shifts at the end of her TCU are constitutive for the ongoing action (i.e. production of an extended multi-unit turn) in face-to-face interpreted talk. The timing of those gaze shifts is particularly interesting; in that way, the interpreter *blocks* the completion points from being treated by the addressee as TRP's and signals to the previous speaker that s(h)e may continue, thus ensuring a smooth continuation of the ongoing multi-unit turn. Furthermore, through a detailed analysis of 'deviant' examples, I have demonstrated how the interpreter can contribute to or transform the projected activity (for example, by projecting a TRP for the participant to take the turn to respond) through her gaze orientation.

Responding to recent discussions on the organization of gaze in social interaction, the present study offers support for Rossano's (2013) claim that gaze is organized with reference to "sequences of talk and the development of courses of action or ongoing interactional projects". At the same time, it offers some additional evidence for the relevance of gaze at the local level of turn-taking (as argued by Kendon 1967 and Auer 2017). Thus, Kendon's and Rossano's claim concerning the main organizational locus of gaze are not necessarily 'competing' (Streeck 2014), but are both essential for our understanding of the regulative role of gaze in social interaction.

Finally, this chapter argues for a multimodal approach to the study of turn-taking in interpreter-mediated interaction. Through fine-grained analyses of the participants' behavior during the exchange, this study has revealed some processes of micro-collaboration and coordination, that would have remained 'invisible' in a purely 'textual' approach of interpreting. Furthermore, the study provides additional evidence for a growing body of research arguing that the interpreter is not just a linguistic 'conduit', but an active participant in the interaction. In that process, interpreter's gaze is more than just one additional 'layer' (Wadensjö 2015) of the interpreter's presence within the interaction, but is be constitutive for the management of the talk and developing courses of action.

Chapter 5

Gaze and backchannel responses in interpreter-mediated interaction

In this chapter, I will focus on the production of backchannel responses, i.e. brief listener signals such as *mh hm*, *yeah* and head nods. Through quantitative and qualitative analyses, my aim is to gain insight into the (multimodal) organization of backchannel responses in interactions conducted with the aid of a consecutive interpreter[44].

Chapter 5 is structured as follows: I will first provide a literature review on backchannel responses in same-language interaction (Section 1.1) and in interpreter-mediated interaction (Section 1.2). In the subsequent analytical section, I will give an overview of the quantitative distribution of backchannel responses in my dataset (Section 3.1). I will then analyze the role of gaze in the production of backchannel responses in interpreter-mediated interaction (Section 3.2). In section 3.3, I will zoom in on the directionality of backchannel responses in interpreter-mediated interaction, with a specific focus on the 'dual feedback pattern' (section 4), and discuss its impact on the participation framework of the exchange. I will close off the chapter with a general discussion and conclusion regarding the relevance of the current analysis for existing models of interpreting and for our current understanding of multimodality in interpreter-mediated interaction.

[44] This chapter is based on the findings presented in Vranjes et al. (2018b)

5.1 Backchannel responses in face-to-face same-language interaction

5.1.1. Defining the phenomenon

Face-to-face interactions are characterized by a high degree of reciprocity between speakers and listeners (Bavelas et al. 2002). During ongoing utterances, listeners manifest their recipiency in a multimodal way: by orienting their gaze towards the current speaker, nodding or shaking their head and producing short tokens such as 'mm hm', 'yeah'. Such listener practices have been called - among others - *accompaniment signals* (Kendon 1967), *backchannel responses* (Yngve 1970, Duncan 1974), *continuers* (Schegloff 1982), *listener responses* (Dittmann and Llewellyn 1968, Bavelas et al. 2002), *acknowledgment tokens* (Jefferson 1984, Mazeland 1990, Clark & Brennan 1991, Drummond & Hopper 1993), *response tokens* (Gardner 2001), *reactive tokens* (Clancy et al. 1996) and *minimal responses* (Mazeland 1990, Barnes 2011, Norrick 2012) in the literature. Through such practices, recipients display attention, understanding, (dis)agreement, assessment and/or affiliation with the ongoing discourse, without claiming a turn-at-talk (for an extensive overview, see Gardner 2001).

From the above-mentioned studies, it becomes clear that this phenomenon covers a broad spectrum of tokens that are not always clearly defined (Drummond & Hopper 1993, Gardner 2001). Response tokens such as 'yeah' and 'mm hm' are sometimes treated in a generic way as indicators of 'attention' or 'listenership'. According to Norrick, for instance, such tokens "just demonstrate recipiency without provoking any specific response, without indicating any particular emotional involvement and without conveying anything about the uptake of the information received" (2010: 527). However, if this were the case it is unclear why backchannel responses would be needed then**,** as the display of 'attention' can also be achieved through continuous gaze orientation towards the speaker (Schegloff 1982).

A large body of research, predominantly within the conversation-analytic framework, has focused on the work accomplished by the tokens such as 'mm hm', 'yeah', 'oh' and head nods, as well as on the differences in their use. Although these

observations refer in the first place to American and English conversations, they also seem to hold for their Dutch equivalents (Mazeland 1990: 252). Discussions have revolved around the following aspects:

- PLACEMENT: short tokens such as 'mm hm' and 'yeah' are placed predominantly at the boundaries of turn-constructional units (Goodwin 1984) or places of possible completion of one speaker's talk. According to Schegloff, "it is structurally relevant at such places for parties to display their understanding of the current state of the talk" (Schegloff 1982: 81). Stronger actions such as assessments (e.g. 'Oh wow', 'That's great') have a different placement within the speaker's units of talk and are consequently treated differently by speakers than 'continuers' such as 'mm hm' (Goodwin 1986, Stivers 2008). Stivers (2008) argues that assessments in mid-telling position can be treated as "too strong" by the teller. On the other hand, nonverbal tokens, such as head nods, can be produced both at boundaries of turn-constructional units or within them (Goodwin 1984, Stivers 2008). Berger & Rae (2012: 1825) noted that visual responses can "fulfil the obligation to respond in a particular time and place (see Sacks et al., 1974) but can evade some of the constraints that impinge on vocal responses".

- DEGREE OF 'SPEAKERSHIP INCIPIENCY': Jefferson (1984) compared the use of '*mm hm'* and 'yeah' in regular conversations and found that 'mm hm' shows a low level of speakership incipiency ('passive recipiency'), whereas 'yeah' indicates recipient's readiness to become next speaker (for similar results, see Drummond & Hopper 1993; for somewhat different observations see Mazeland 1990).

- LISTENER'S STANCE: Traditionally, neutral tokens such as 'yeah' and 'mm hm' and head nods have been contrasted with more evaluative tokens like assessments (e.g. 'That's great') and change-of-state tokens (e.g. 'Oh') (Heritage 1984, Goodwin 1986, Betz & Golato 2008, Stivers 2008, Heinemann & Koivisto 2016). Bavelas et

al. (2000) refer to the former as 'generic' and to the latter as 'specific'[45] listener responses. Specific responses are usually produced later in the narrative than generic responses, because they "require more information about the story and about the narrator's perspective" and "would not be possible or credible until the listener has this information" (Bavelas et al. 2000: 945). Stivers (2008), in this respect, proposes a distinction between *affiliation* and *alignment* in her study of response tokens in storytelling contexts. According to Stivers, head nods in mid-telling environments are used to display endorsement of the teller's stance (affiliation). Other mid-telling tokens such as 'mm hm' and 'yeah' merely acknowledge the information provided and support the progress of the telling, without claiming support of the teller's stance (alignment). Furthermore, the way the response tokens are clustered together also reveals something about the recipient's stance towards the talk (Gardner 2001). For instance, if the listener always uses the same token ('mm hm'), it is likely that (s)he does not orient to the utterance as new: this may even be an indication of listener's incipient disinterest (Schegloff 1982; Jefferson 1984, Gardner 2001).

For the sake of consistency and clarity, I will use the term 'backchannel response[46]' in this chapter to refer to listener's actions within current speaker's ongoing talk through which the listener displays *how* (s)he has understood/analyzed the prior talk and how (s)he is projecting further actions in the talk (Gardner 2001). For instance, by producing a continuer such as 'mm hm', the listener treats the preceding talk as part of a larger, incomplete structure, while at the same time indicating no problems of understanding with it (Gardner 2001; see also Schegloff 1982). Backchannel responses

[45] According to Bavelas et al. (2000: 243), *generic* backchannel responses "are not specifically connected to what the narrator is saying, in the sense that the same generic response would be appropriate to a wide variety of narratives", whereas *specific* backchannel responses "are tightly connected to what the narrator is saying at the moment. They take on a form specific to the narrative content of the moment and are not generically appropriate to all narratives"

[46] Despite the fact that the term 'back-channel' has been criticized within CA framework as attempting to "analyze recipient's behavior in isolation from speaker's talk" (Goodwin 1986: 207), it will be used in the present analysis for ease of reference.

include the various types of tokens, as surveyed by Gardner (2001:2-3) and presented in Table 5.1.

Table 5.1: Functional types of backchannel responses.

GENERIC	**Continuers**	function to hand the floor[47] back to the immediately prior speaker (e.g. *Mm hm, Uh huh*);
	Acknowledgments[48]	claim agreement or understanding of the prior turn (e.g. *Mm, Yeah*)
SPECIFIC	**Newsmarkers**, and other newsmarker-like objects:	mark the prior speaker's turn as newsworthy in some way (e.g. *Really?*, the **change-of-state token** *Oh*, the "idea-connector" *Right*)
	Change-of-activity tokens	mark a transition to a new activity or a new topic in the talk (e.g. *Okay*, *Alright*);
	Assessments	evaluate the talk of the prior speakers (e.g. *Great, How intriguing*);
	Brief questions	for clarification or other types of repair, which seek to clarify mishearings or misunderstandings (e.g. *Who?*, *Which book do you mean?*, or the very generalized *Huh?*);
	Collaborative completions:	whereby one speaker finishes a prior speaker's utterance (e.g. A: *So he's moved into …* B: *commercial interests*);
GENERIC/ SPECIFIC	Many **nonverbal vocalizations** and **kinesic actions**	(e.g. head nods and head shakes, sighs, laughter).

[47] Gardner's (2001) definition sounds somewhat contradictory here, because the listener does not take the floor by producing continuers, but rather signals to the current or immediately prior speaker that (s)he may continue.

[48] Interestingly, Clark & Brennan (1991) use the term 'acknowledgments' as an umbrella term for both 'continuers' and 'assessments'. This illustrates the lack of consensus on the terminology used to describe such hearer practices.

Backchannel responses are considered to be central for the success of communication as a **joint activity** (Goodwin 1984, 1986; Clark & Brennan 1991, Bavelas et al. 2000, Gardner 2001, Mondada 2011, Tolins & Fox Tree 2014). Speakers and listeners do not construe their actions in isolation from each other, but in concord. As noted by Heath (1992: 121) "the very production and understanding of the activity relies upon the way in which the potential recipient is participating in the moment by moment accomplishment of the activity". Studies have shown that speakers are sensitive to the presence or absence of listener's backchannel responses and that these shape the development of the speaker's talk (Goodwin 1981, 1986, Stivers 2008, Mondada 2011). Experimental research confirmed these observations by showing that speakers are less able to tell their stories well in front of unresponsive listeners as they are in front of active listeners (Bavelas et al. 2000). Furthermore, Goodwin (1981) demonstrated how the speaker, who experiences a lack of listener's visual attention, produces phrasal breaks, such as restarts and pauses, to draw the listener's gaze. Recipient's actions are thus important for the overall progressivity of the speaker's turn. On the other hand, a recipient's backchannel responses are closely linked to the speaker's actions; for instance, studies have demonstrated how speakers use their gaze to solicit a backchannel response from the listener (see Goodwin & Goodwin 1986, Bavelas et al. 2002, Stivers & Rossano 2010), showing that speaking and listening are interrelated activities in face-to-face interaction.

5.1.2. Backchannel responses and grounding

According to Clark & Brennan (1991) backchannel responses and continued attention are the most basic forms of positive evidence of understanding and are inherent in the collaborative process of grounding (Clark & Schaefer, 1989; Clark & Brennan, 1991) or the establishment of intersubjectivity in communication (Deppermann, 2015). Grounding is the process through which participants collaborate with each other moment by moment to establish the mutual belief that what has been said has been understood, or that it has been made part of their common ground "to a criterion sufficient for current

purposes" (Clark & Schaefer, 1989; Clark & Brennan, 1991). This is accomplished through a three-step procedure: when an action-to-be-understood is presented by the speaker (i), the recipient should display how (s)he understands it (ii), and finally the speaker should display if (s)he accepts[49] what (s)he takes up as the recipient's understanding of his/her turn (iii) (Deppermann, 2015: 66). This basic process, which is essential for conversational cooperation, is "an emergent, observable, shared product of the interaction" (Deppermann 2015: 65-66). Grounding does not only proceed on a turn-by-turn basis, but also within unfolding phrases (see Goodwin, 1986; Goodwin & Goodwin, 1987; Clark, 1996). This is where backchannel responses perform an important role. Being closely linked to the immediate context in which they are produced, backchannel responses reveal much about the recipient's analysis of the talk in progress and the development of intersubjectivity in interaction. For example, a change-of-state token such as '*Oh*' marks that its producer has undergone some sort of transformation in his "current state of knowledge, information, orientation or awareness" (Heritage, 1984: 299). The seemingly simple tokens such as 'mh hm' and 'yeah' are used by the recipient when there is opportunity to initiate repair, and thus indicate no problems of understanding with the turn so far (Schegloff 1982, Clark & Brennan, 1991; Gardner, 2001). Backchannel responses are thus a highly efficient way to establish common ground in interaction.

5.1.3. Visual backchannel responses and multimodality

Face-to-face interactions are characterized by their *visibility* (Bavelas et al. 2000), which means that interlocutors can "produce and receive at once and simultaneously" (Clark 1996: 9). Thus, apart from tokens such as 'mm hm' and 'oh', recipients' employ a wide range of visual practices such as head nods, head shakes and facial expressions to display their attention, understanding, (dis)agreement and affiliation (Kendon 1967, Goffman 1981, Dittmann & Llewellyn 1968, Goodwin 1981, Clark & Brennan 1991, Bavelas et

[49] The third step, the acceptance phase, can be realized by the speaker's relevant next turn (Davidson, 2002).

al. 2000, Bavelas et al. 2002). In recent years, there has been a growing interest in the function and employment of such visual backchannel responses (Bavelas et al. 2000, 2002; Bertrand et al. 2007, Stivers 2008, Whitehead 2011, Mondada 2011, Ford & Stickle 2012, Peräkylä & Ruusuvuori 2012, Muntigl et al. 2014, Hömke et al. 2017). Yet, there is much more to be learned about their sequential organization in actual talk.

Visual backchannel responses are often lumped together in a single category 'nonverbal behavior' (Tolins & Fox Tree 2014) or 'kinesic actions' (Gardner 2001, see Table 4.1. above) without differentiating between them in more detail. Head nods, for instance, are usually put on a pair with 'continuers', such as *mm hm* (Schegloff 1982, Goodwin 1986, Clark & Brennan 1991, Bavelas et al. 2000). However, Stivers (2008) has argued that head nods perform different functions (i.e. affiliation with the speaker's stance) than vocal continuers in story-telling environments, which illustrates that more research is needed on the design and different interactional/sequential uses of such visual practices within actual talk.

Another important aspect in the production of visible backchannel responses is the aspect of *simultaneity.* Visible backchannel responses are often co-produced with verbal tokens such as 'yeah' and 'mm hm'. According to the "integrated message model" developed by Bavelas et al. (2002), such co-occurring verbal and nonverbal backchannel responses work together to convey meaning. Allwood & Cerrato (2003), for instance, report that gestural backchannel responses (e.g. head nods) mostly co-occur with short vocal or verbal responses, together with which they perform a related or complementary function. Yet, within Conversation Analysis literature, the multimodal manifestation of backchannel responses has received only scant attention (Mondada 2011). In her widely cited study on head nods in the context of story-tellings, Stivers (2008) made it explicit that she focused only on head nods "occurring on their own (i.e., not in overlap with a spoken item)". Also, other modalities such as gaze, facial expressions were not taken into account. It is however worthwhile examining whether and how different resources interact and how this bears on their communicative import in the interaction, as I will illustrate in this chapter.

5.1.4. Backchannel responses within a multiperson participation framework

In dyadic interactions, the recipient's production of backchannel responses is related to the behavior of the current speaker (see Bavelas et al. 2002, Stivers 2008, Stivers & Rossano 2010). Within a multiperson[50] participation framework, however, the display of recipiency and collaboration is a more complex matter. In triadic interactions, which are of special interest in this study, the speaker may be addressing only one of the participants, while the third remains a momentarily 'unaddressed participant'[51] (Goffman 1981, Bolden 2013) or a 'side-participant' (Clark 1996). This results in more competition for visual attention between interlocutors. Gaze is then of particular importance to disambiguate who is being addressed at a particular moment during the talk and who is being selected to speak next (Rossano et al. 2009).

Triadic interactions also create the possibility for two participants to be involved in the production of a telling, acting as 'consociates' (Lerner 1992) or 'co-members of a party of tellers' (Bolden 2013) to a third person. This is especially the case when two interlocutors have shared experience or knowledge of some subject matter. Gaze is then employed to acknowledge the state of affairs between participants (Lund 2007). As famously shown by Goodwin (1981), speakers orient to listener's knowledge states during the production of their talk. For instance, while moving his/her gaze from an unknowing to a knowing recipient, the speaker may display uncertainty about some aspect of the telling that (s)he may share with the knowing recipient and ask for verification (Goodwin 1981:153). Knowledge states (or epistemics) have also been shown to play an important role in conversational turn-taking (Lerner 2003) and in the organization of participation in repair sequences[52] (Bolden 2013). In the context of

[50] The term 'multiperson' instead of 'multiparty' is used here, as participants in an interaction may act collectively as a single unit or a multiperson "collectivity" (see Bolden, 2013).

[51] The distinction between "addressed" and "unaddressed" is mainly based on the speaker's visual orientation (see also Goffman 1981).

[52] i.e. "sequences of talk dedicated to resolving problems of hearing, speaking or understanding" (Bolden 2013).

other-initiated repair in multiperson interaction, Bolden (2013) found that repair initiators use their gaze to select the "owner of the experience" to resolve repair, and not the source speaker (or 'animator').

The question then is whether listeners in triadic, face-to-face interactions also display their sensitivity to the differences in their co-participants' participation status and knowledge states during the production of backchannel responses. This will be discussed further in the analytical part of this chapter (section 5.3).

5.2. Backchannel responses in interpreter-mediated interaction

In reviewing the literature on this phenomenon, it is surprising to find that there are only few previous studies on backchannel responses in interpreter-mediated talk. Interpreter-mediated interaction creates a particularly complex situation for the production of backchannel responses and the processes of grounding for two main reasons: (1) the interpreter's participation status within the conversation and (2) the primary participants' lack of direct communicative contact or understanding. In the following, I will discuss each of these aspects separately.

5.2.1. The interpreter's participation status

The interpreter's involvement within the interaction is obviously not the same as that of the primary participants. The interpreter typically provides renditions of one speaker's talk in every second turn, in the language of the other participant. In this way, the primary speaker's and the interpreter's voice are merged into one in the interpreter's rendition. This necessarily raises the question of 'speakership' in such exchanges. The interpreter's involvement has often been described in terms of Goffman's (1981) notion of 'footing', according to which a person can take up different stances to the words being expressed - as the 'animator' of the words (who produces the words), 'author' (who selects the words and sentiments that are being expressed) and 'principal' '(i.e. who is committed to the expressed stance and beliefs). According to this view, the interpreter

acts as 'animator' or 'author' on behalf of another interlocutor (see also Wadensjö, 1998, Pöchhacker, 2012). In other words, when the interpreter is rendering primary participant's assessment such as 'I don't like the weather today', she is the spokeswoman for that assessment, for which the prior primary participants is the principal (see also Clark & Carlson 1982). This is particularly evident when the interpreter uses the first-person pronoun ("I-form") to render the prior speaker's turn, which is in line with most codes of conduct for community interpreters (Bot 2005, Van De Mieroop 2012).

Goodwin & Goodwin (2004) pointed to the limitations connected to Goffman's model, stating that "the methods offered for investigating participation take the form of a typology, a set of static categories" and that "no resources are offered for investigating how participation might be organized through dynamic, interactively organized practices" (225). Thus, although Goffman's work allows us to rethink the 'speaker' as a laminated entity, the question is how those categories are grounded in actual interaction. How do participants themselves orient to these differences? In her study of conversational repair in multiparty interactions, Bolden (2013) investigated the epistemic underpinnings of Goffman's notions in actual talk. She found that interlocutors address their repair initiation not just to the producer of the trouble source (or the 'animator), but also to the qualified participant with the 'epistemics of experience' (or the 'principal' in Goffman's terms), who has primary epistemic rights and responsibility over the trouble source to provide repair. As Bolden puts it, "when somebody talks about experiences or life circumstances of a co-present participant, that participant retains privileged access to what is being said" (2013: 320). Knowledge appears to be a "moral domain" with clear implications for the management of social relationships (Stivers et al. 2011).

In terms of the socially distributed rights to knowledge or **epistemics**, when the interpreter is rendering the thoughts, life circumstances and experiences of a primary participant, the latter remains the 'owner' of those thoughts and circumstances based on his/her 'epistemics of experience' (Bolden 2013). Thus, the interpreter has a different

status and responsibilities with respect to knowledge[53] than the primary participants, who retain epistemic primacy or authority over the content of what is being said[54].

Previous research has shown that primary participants can display various orientations towards the interpreter, treating him/her as a mere 'animator' of the preceding utterance or as a real participant with primary epistemic rights in the interaction (Bot 2005, Mason 2012, Pasquandrea 2011). Interpreter-mediated interactions are thus a particularly interesting setting to study how recipients orient to differences in participation status of their co-participants in ongoing conversation. As observed by Wadensjö:

> "the interpreter-mediated conversation offers unique opportunities for researchers to trace how participation status is marked and confirmed in and by interaction. Through detailed analysis it should be possible to uncover how interlocutors understand the interaction order as interaction proceeds, (…) and how participants perceive the distribution of responsibility for the substance and the progression of current talk" (1998:94).

Interpreters do not only act as speakers, but also as **recipients** of the ongoing talk. Although they are not considered as the 'main' recipients, they are the *'first'* recipients of the primary speakers' talk (Gavioli 2012). In fact, interpreters are arguably the most attentive listeners in the conversation. They focus on the meaning, but also on the context and the overall conversational goals of the encounter. For example, they judge if the patient's turn is appropriate (Hsieh 2012, Bolden 2000) and in line with overall conversational goals, before rendering it to the other participant.

---

[53] The interpreter does have some epistemic authority arising from her knowledge of both languages in the interaction, which allows her to speak on behalf of the primary speakers. This is what Heritage (2013) and Bolden (2013) refer to as 'epistemics of expertise'.

[54] This may also explain why people who interact with the aid of an interpreter are usually recommended to gaze at each other, instead of at the interpreter.

In line with Goffman's decomposition of the 'speaker' into the roles of an 'animator', 'author' and 'principal', Wadensjö (1998) devised a 'reception format' to describe the different ways in which recipients can relate to the ongoing talk. She distinguished between a 'reporter' (who memorizes for repetition), a 'recapitulator' (who recapitulates and gives an authorized voice to a prior speaker), and a 'responder' (who acts as the final addressee, by e.g. producing backchannels and orienting gaze to the current speaker) (ibid. 92). These categories were further elaborated and expanded by Merlini & Favaron (2005). Still, the interpreter's role as recipient-responder of talk remains a delicate subject (Baraldi & Gavioli 2012; Gavioli 2012).

The production of **backchannel responses by interpreters** during the talk has received scant attention in the literature. Linell, Wadensjö & Jönsson (1992) reported an overall low incidence of (verbal or vocal) backchannel responses by the interpreters in a large corpus of interpreter-mediated dialogues in legal and medical contexts (however, see section 3.1. for different results). On the basis of these findings, they concluded that backchannel responses such as *mm hm* and *yeah* are "dysfunctional from the point-of-view of the interpreter's task", whose main focus is to understand and memorize what is being said (Linell, Wadensjö & Jönsson 1992: 133). Their study only took verbal backchannel responses into account, without acknowledging the complex interactional work performed by such tokens. In addition, the interpreter's production of backchannel responses can be treated as in discord to the principle of 'neutrality' that forms one of the building blocks of the interpreter's professionality (Englund Dimitrova 1997). One could argue that backchannel responses are an appropriate action only for 'true' participants (see also Gardner 2001: 10) and thus not for the interpreter. However, the interpreter's provision of backchannel responses is a highly context-sensitive matter, which depends on the type of action being performed and overall conversational goals of the interpreted event (Gavioli 2012). In the current dataset, the establishment of mutual rapport and reciprocity ('interpersonal goals', see Clark 1996) is of greater importance than for example in legal settings.

Backchannel responses are increasingly seen as an important aspect of the interpreter's coordinating activity (Merlini & Favaron 2005, Gavioli 2012). Through

backchannel responses, the interpreter has the opportunity to coordinate the talk, "to show that she still has the listener role, i.e. that she does not claim her turn, and that the speaker can therefore go on speaking" (Englund Dimitrova 1997: 161), while putting the other primary participant 'on hold' (Gavioli 2012). Also, through their production of verbal backchannel responses (such as 'okay') towards the end of the current speaker's turn, the interpreters may signal a shift from the activity of listening to the activity of speaking. Gavioli sees advantage in the use of minimal responses as they make "understanding possible and 'public' to different participants who speak different languages" (2012: 222).

In the following, I will discuss how interpreting affects the primary participants' orientations towards each other and their opportunities for full participation in this type of encounters.

5.2.2. Lack of direct communicative contact between primary participants

One of the main challenges in conversing with the aid of an interpreter is "based largely in the difficulty in establishing reciprocity of understanding between the primary interlocutors in the discourse" (Davidson 2002: 1274). As they have little or no knowledge of each other's language, primary participants usually have to wait for the interpreter's rendition of the preceding talk in order to understand and react to what the other participant has said. Even when primary participants produce backchannel responses, interpreters rarely translate (or repeat) them in the other language (Englund Dimitrova 1997, Wadensjö 1998). Englund Dimitrova (1997) noted that it might be challenging for the interpreters to reproduce such backchannel responses, because the link between the recipient's response token and the part of the utterance that it is responding to would be lost. Moreover, given their relative transparency in different languages, 'close-renditions' of such minimal tokens can be experienced as redundant by the primary participants:

> "It is as if the relative 'transparency' (Müller 1993) of this communicative activity reduces the relevance of translating (…) Providing 'close renditions' of

> back-channeling tokens, the interpreter can demonstrate, as it were, her image as a 'close-texts-producer' (…) this may play down the uniqueness of the interpreter's position in interaction, since interpreting **odd words of this kind** points at the possibility of interpretation being superfluous." (Wadensjö 1998: 122, my emphasis)

Merlini & Favaron (2005) reported different observations in their study of speech pathology sessions, where the interpreters displayed tendency to convey backchannel responses such as 'mm hm'. According to Merlini & Favaron, the interpreter's repetition of the speech therapist's backchannel responses reinforces the empathic communication model adopted by the therapist. Interpreter's (re)production of backchannels is thus a context-sensitive matter, that should be in line with overall conversational goals of the interpreted event.

According to Davidson's (2002) well-established collaborative model of interpreting, interpreter-mediated interaction takes the shape of two overlapping dyads where two sets of common ground are being co-constructed; between the interpreter and each of the primary participants (A and B in 5.1 below). The primary participants are fully dependent on the interpreter whose task is not only to create shared understanding (Davidson 2002), but also to establish contact and rapport between them (Linell, Wadensjö & Jönsson, 1992)[55]. Consequently, during a dyadic exchange in one particular language, the participant who has no understanding of the language is temporarily excluded from the participation framework (see Figure 5.1 below). Although Davidson's model provides for a better understanding of the interpreter's role in the organization of grounding in interpreter-mediated interaction, it rules out other micro-phenomena in the interaction as 'deviations' from the model. Furthermore, there is no place for nonverbal practices such as head nods in this model, as it focuses only on the linguistic achievement of reciprocity and understanding.

[55] In cases where the primary participants can understand their co-participant's utterance, this can lead to direct grounding between them.

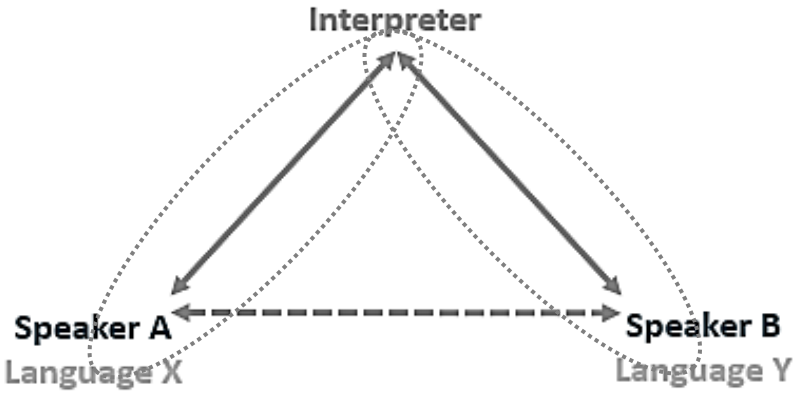


Figure 5.1: Interpreter-mediated interaction as two overlapping participation frameworks.

As any form of face-to-face communication, interpreter-mediated talk is characterized by its **visibility and simultaneity** (Clark 1996) and requires collaboration on every level of communication. It has been shown that an interpreter-mediated talk is coordinated through the collective efforts of all participants and that participants' visual behavior and orientation towards each other is crucial for the moment-by-moment negotiation of the (dyadic/triadic) participation framework (see Davitti & Pasquandrea, 2017; Pasquandrea, 2011). Yet, no systematic research has been done on the multimodal organization of backchannel responses and their impact on the organization of participation in this type of encounters.

Altogether, previous sections have shown that interpreter-mediated interaction differs in some important aspects from same-language (triadic) interaction, which can have an impact on the production of backchannel responses in this type of exchange. This raises an important basic question: *how* do listeners and speakers collaborate and adapt their backchanneling behavior to the specificities of this type of exchange? Through a combination of quantitative and qualitative (conversation-analytical) methods, the following analyses aim to provide answers to the following questions:

- What is the distribution of backchannel responses in the current dataset? Are there any differences in the production of backchannel responses between the participants in an interpreter-mediated interaction? (Section 5.3.1)

- What is the role of speaker gaze in the production of backchannel responses in interpreter-mediated talk? (Section 5.3.2)
- Given the complex participation framework of an interpreted talk, what is the *directionality* of backchannel responses in such an exchange? (Section 5.3.3)

I will link these findings to the complex participation framework of interpreted encounters and to the differences in the participation status of the participants within the exchange.

Table 5.2: A summary of the previous studies on backchannel responses in interpreter-mediated interaction.

	LINELL et al. (1992)	**ENGLUND DIMITROVA (1997)**	**GAVIOLI (2012)**
Definition	Feedback	Feedback	Minimal responses
	OK, 'mhm', 'mm', 'yeah', 'I see', repetitions, …	verbal ('ah', 'I see', 'mmm') and nonverbal (head nods & facial expressions)	'mhm', 'yes', 'no', short (or partial) repetitions or short turn completions
Data / setting	40 audiotaped medical and legal interactions	2 videotaped medical interactions	150 audiotaped medical interactions
Approach	(unclear)	descriptive, qualitative	qualitative
Nonverbal resources	no	yes	no
Primary speakers' backchannels	yes	yes	no
Placement	turn-initial, turn-medial, turn-final	turn-initial, turn-medial	turn-final; turn-medial
Function	Markers of acceptance and/or understanding; coordination devices	Display of understanding, reactions and attitude + coordination	Mechanism of 'back-reference' and projecting translation (turn-final); negotiating understanding, coordination

5.3. Analysis

5.3.1. Backchannel responses in the dataset

In the following, I will focus on verbal and nonverbal backchannel responses in turn-medial position, i.e. stand-alone backchannel responses preceded and followed by the talk of the main speaker. The analysis does not include backchannel responses that occur at places where the main speaker's turn has recognizably reached a point of possible completion or those that are produced by the listener to claim the conversational floor (so-called 'turn-initial cue phrases' Hjalmarsson 2010, Gravano et al. 2013)[56]. In coding head nods, the *amplitude* (shallow vs. deep), *velocity* (slow vs. fast) and *number* (single, double or multiple) of produced head nods were taken into account (see Stivers 2008, Muntigl et al. 2012). For instance, continuous nods that were produced without change in velocity were counted as *one* backchannel response. Cases where continuous nods suddenly changed in amplitude (from shallow to very deep) - which typically corresponds with increased interest and understanding - were counted as two backchannel responses. Furthermore, clusters of tokens (such as 'head nods + *mm hm*') were coded as one backchannel response if the temporal distance between them was less than 250 ms. I have made no attempt to subdivide nonverbal backchannel responses according to their formal and functional characteristics, as this would exceed the scope of the present study. I coded the first 15 minutes of each conversation, which resulted in a dataset of 1442 instances of backchannel responses in turn-medial positon.

Figure 5.2 below summarizes the distribution of turn-medial backchannel responses of students, interpreters and counsellors in my dataset during speakers' turns. In total, I made an additional distinction between interpreter's speaking turns in Russian and between interpreter's speaking turns in Dutch. Consequently, primary participants are listeners in 3 different constellations: (1) during the interpreter's rendition in their

[56] Such tokens often display a double orientation towards the talk: on the one hand, responding to the previous turn and on the other, projecting speaker transition. Gavioli (2012) observed that interpreters often produce tokens such as 'okay' or 'mm hm' to claim the turn.

own language, (2) during the interpreter's rendition in the foreign language and (3) during the primary participant's turn. Thus, primary participants listen both to the 'original' (principal's) utterances in the other language, as well to the interpreter's renditions (in both languages). Although this might seem as an obvious observation, it has been largely overlooked in previous studies.

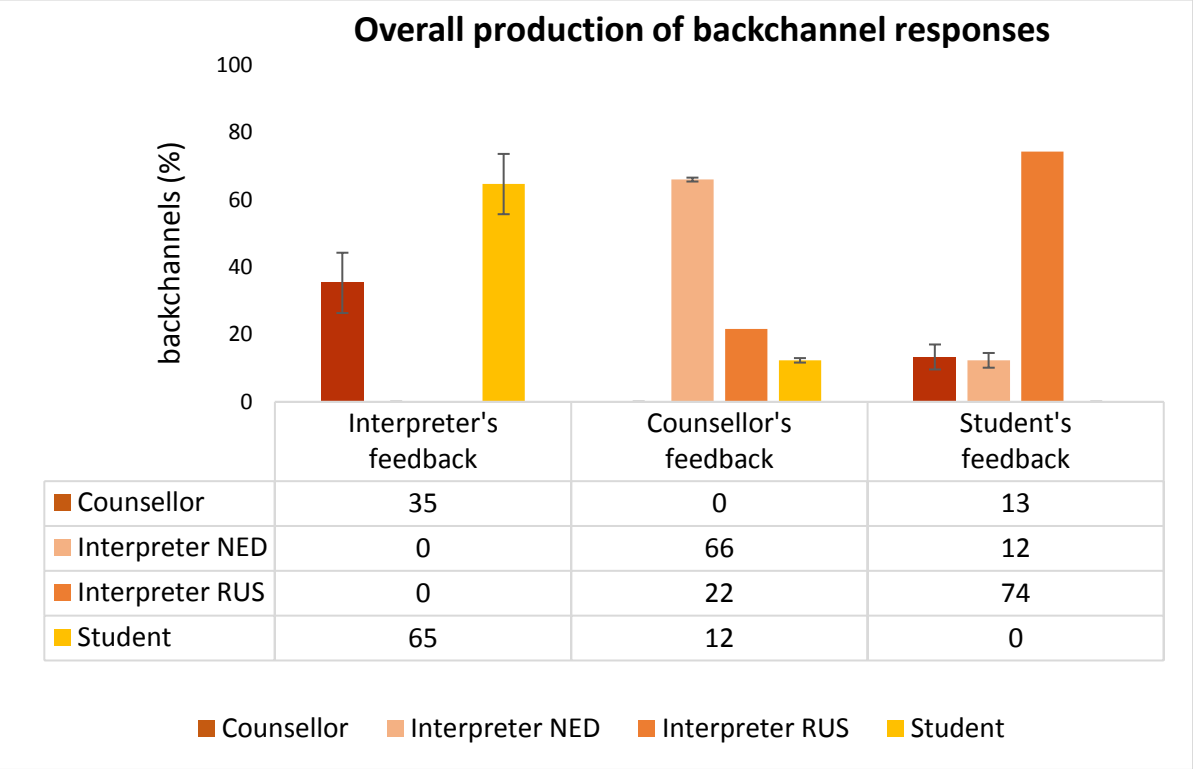


Figure 5.2: Overall production of turn-medial backchannel responses (in percentage) in the dataset: Interpreters (n=247), Counsellors (n=528) and Students (n=667). Errors bars show the standard deviations.

As one would expect, both students and counsellors produce backchannel responses most frequently during interpreter's renditions in their own language (74% and 66% respectively). When listening to Russian, counsellors tend to provide more backchannel responses during the interpreter's renditions of their own utterances (22%) than during student's utterances (12%). Although in both cases the utterances are produced in an unfamiliar language, in the former case, the interpreter is rendering the counsellor's preceding turn in Russian, which implies that the counsellor has knowledge of the content of the turn (but not of the exact order nor of the words used). Such

backchannel responses are often a *display of endorsement* of the interpreter's turn. As an example, consider the counsellor's backchannel responses in the following extract (the figures are taken from the counsellor's scene camera).

Extract 1

```
1   INT     у вас же там #какой-то есть коор+[динат]ор, *[да]
            you already have some   coordinator there, isn't it
2   CNS                                      [((nose exhale))]
3   STUD                                                [mh hm]
    cns                                   +repeated nods------->
    stud                                             *double nod-->
    fig                    #fig.1-2
```

figure 1

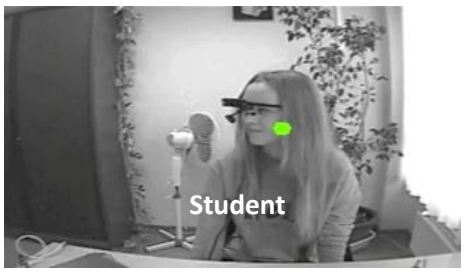


figure 2

```
4   INT     который всех* (-) как иностранных студ+ентов курируют
            who supervises all    those foreign students
    cns     ------------------------------------+
    stud    ------------*
```

In this example, the interpreter is rendering the counsellor's question in Russian to the student[57]. After hearing the word 'coordinator' (which is similar in Russian and in Dutch), the counsellor starts producing a series of nods, while quickly shifting his gaze from interpreter (Figure 1) to the student (Figure 2). Through these head nods, the counsellor appears to be asserting confirmation with what the interpreter is saying to the student (as knowledgeable participant) and displaying his authority as the 'principal' (Goffman 1981) towards the ongoing talk.

Students, on the other hand, tend to produce slightly more backchannel responses (predominantly head nods) during counsellor's turns (13%) than during the interpreter's renditions in Dutch (12%). There are two possible explanations for this: first,

[57] Note how the counsellor is left outside of the participation framework (as 'side-participant') created through mutual orientation between the interpreter and the student.

counsellors in the present dataset tend to address students while speaking[58] (see Chapter 3), which may mobilize a nonverbal backchannel response from the student from time to time (Bavelas et al. 2002). Second, students may be less inclined to display themselves as 'principals' during the interpreter's turn. As shown in Chapter 3, students orient towards the interpreter as 'real' interlocutor, with a certain authority ('epistemic of expertise', Bolden 2013) within the exchange and not as a mere linguistic conduit. Overall, these results illustrate how a study of the distribution of backchannel responses in face-to-face talk can reveal much about interlocutors' orientations toward each other and their own positions within the exchange.

The interpreters in my dataset produce significantly less backchannel responses (n=247) than the primary participants (counsellors: n=528; students: n=667), which is in line with their distinct interactional role within the exchange (see also Linell, Wadensjö & Jönsson 1992). They also tend to produce more backchannel responses during student's turns (65%)[59], than during counsellor's turns (35%). We conducted a Paired Samples T-test in order to test whether the average amount of backchannel responses produced by interpreters was different during students' turns and counsellors' turns. The results of this analysis indicated that on average, interpreters produced more backchannel responses during students' turns (M = 14.625, SE = 10.730) in comparison to counsellors' turns (M = 8.125, SE= 8.374), t (7) = -2.402, p = 0.047, r = 0.705. These findings suggest that interpreters display orientation towards the 'weaker' participant in the interaction. This may in part be attributed to the fact that, unlike counsellors, students usually address interpreters while speaking (see Chapter 3). Furthermore, an analysis of the sequential environments in which interpreter's backchannel responses are produced shows that interpreters rarely (only in 7% of the cases) provide backchannels during first-pair parts (such as questions, see Sacks et al. 1974, Stivers & Enfield 2010) of the primary participants. During questions the interpreter appears to be less manifested as a

[58] Students, on the other hand, tend to gaze at the interpreter while speaking (see Chapter 3).

[59] It is important to take into account that students had less turns (n=203) than counsellors (n=310) in the present dataset.

recipient than in other sequential environments (such as extended tellings, see Chapter 6).

Modality of backchannel responses

Given that previous research on backchannel responses in interpreter-mediated interaction has primarily focused on *verbal* responses, we will now look at the proportion of *nonverbal* backchannel responses in the dataset. I made a formal distinction between (a) nonverbal backchannels (such as head nods, head shakes, smiles), (b) verbal backchannels (such as '*mm hm'* and '*ja*') and (c) combined backchannels (for example, a nod accompanied by *'yeah'*).

Table 5.3: Modality of backchannel responses.

	BACKCHANNEL RESPONSES	INSTANCES	%
Interpreters	Nonverbal	176	71%
	Verbal	24	10%
	Combined	47	19%
Counsellors	Nonverbal	351	66%
	Verbal	36	7%
	Combined	141	27%
Students	Nonverbal	543	81%
	Verbal	35	5%
	Combined	89	14%

As shown in Table 5.3, the majority of backchannel responses in the dataset are situated on the nonverbal level. For instance, more than two thirds (71%) of interpreter's backchannel responses do not contain a verbal component. Such findings demonstrate the importance of including the visual dimension of interaction in the study of interpreter-mediated encounters.

A further analysis of the type of verbal backchannel responses shows that interpreters in the majority of cases (78%) provide continuers (e.g. *mm hm*) and acknowledgments (e.g. *ja*). I found only 4 cases (6%) of newsmarkers (e.g. *Ah*), which

were then usually produced after resolving some problem of understanding or hearing on the part of the interpreter (on 'repair solution', see Schegloff 1977). The interpreters in the present dataset did not produce assessments (e.g. *Oh wow, Great*) during primary speakers' talk.

5.3.2. Backchannel responses and speaker's gaze

Now that we have established the importance of a multimodal approach in the study of backchannel responses, I will focus on the relationship between speaker's gaze and listener's backchannel responses in interpreter-mediated talk. More specifically, the question is whether listeners in the current dataset produce more backchannel responses during the so-called 'gaze window' (Bavelas et al. 2002) or mutual gaze with the current speaker (see also Kendon 1967). Gaze direction is crucial for the display of recipiency in face-to-face interaction. Interlocutors employ their gaze to display attention and (dis)engagement (Goodwin 1981) and their participation status within the given interaction (see also Rossano 2013). By orienting their gaze to the speaker, listeners show that they are acting as hearers to the speaker's utterance (Goodwin 1980). Bavelas et al. (2002) showed in an experimental study that speaker's gaze is an important cue for eliciting backchannel responses in dyadic face-to-face interactions (Bavelas et al. 2002). They observed that the listener "tended to respond when the speaker looked at her, and the speaker tended to look away soon after the listener responded", creating a "gaze window to coordinate their actions" (Bavelas et al. 2002: 576-7). Thus, according to Bavelas et al., timing of backchannel response is a "collaborative process", linked to speaker's gaze. However, Bavelas et al. (2002) did not differentiate in their analysis between verbal and nonverbal backchannel responses. Also, their dataset consisted of interactions between two participants, whereas in the present study I am focusing on interactions with three participants. In the following, I will focus on the interpreter's and the counsellor's backchannel responses in relation to current speaker's gaze.

Table 5.4 shows that two thirds (66%) of interpreters' nonverbal backchannel responses are produced after mutual gaze with the current speaker (student or counsellor), whereas verbal responses are only in 25% of the cases produced during

mutual gaze. A Chi-square test shows that nonverbal backchannel responses are significantly more likely to occur after mutual gaze than verbal responses ($\chi^2(1) = 14.85$, $p < 0.001$). This may be attributed to the fact that nonverbal backchannels such as head nods are meant to be seen by the speaker. Yet, the counsellors' production of nonverbal backchannel responses does not seem to follow the same distribution, as less than half of them (48%) are produced with mutual gaze. This difference in distribution between counsellors' and interpreters' nonverbal backchannel responses is statistically significant ($\chi^2(1) = 12.32$, $p < 0.001$), viz. interpreters are more likely to produce nonverbal backchannel responses after mutual gaze with the speaker than counsellors. I have already noted in a previous section that, whereas interpreters orient at one speaker at the time while listening (see Chapter 3), counsellors are in a more 'complex' position as listeners, dividing their visual orientation between the interpreter and the other primary participant (to whom the words attest as 'principal', Goffman 1981). That may also explain why their nonverbal backchannel responses are less dependent on mutual gaze with the current speaker than the interpreters' nonverbal responses.

Table 5.4: Production of backchannel responses in relation to mutual gaze.

	MODALITY	**MUTUAL GAZE**	**NO MUTUAL GAZE**	TOTAL
Interpreters	Nonverbal	116 (66%)	60 (34%)	176
	Verbal	6 (25%)	18 (75%)	24
	Combined	15 (32%)	32 (68%)	47
Counsellors*	Nonverbal	87 (48%)	96 (52%)	183
	Verbal	15 (43%)	20 (57%)	35
	Combined	67 (54,5%)	56 (45,5%)	123

*the focus is on counsellors' backchannel responses during interpreter's renditions in their own language.

Previous sections suggest nonverbal backchannel responses are more strongly linked to mutual gaze than verbal backchannel responses. Moreover, it appears that Bavelas' et al. (2002) claims should be treated with caution in triadic interactions with an interpreter. Bavelas' et al. (2002) study has been criticized recently by Rossano (2013), who commented that "other communicative behaviors to solicit a response were

simply not taken into account, and, as such, the specific actions performed through talk ignored" (2013: 317). If speaker gaze were to have the "strongest and most consistent relationship to a backchannel response" (Bavelas et al. 2002: 569), then it may be able to mobilize a backchannel response independently from other linguistic factors, such as syntax and semantics. Interpreter-mediated talk provides an ideal environment to test whether this is the case, given the fact that there is always one participant during the talk who has no understanding of the language. If speaker gaze were enough to mobilize a response from the listener, then we would expect that counsellors will display a tendency to provide a backchannel response every time that student gazes in his/her direction. I chose to focus on counsellors here, because students mostly gaze at interpreters while speaking, which makes their gaze toward the counsellor a 'marked' event. The dataset contained 65 instances of counsellor's backchannel responses during student's turn.

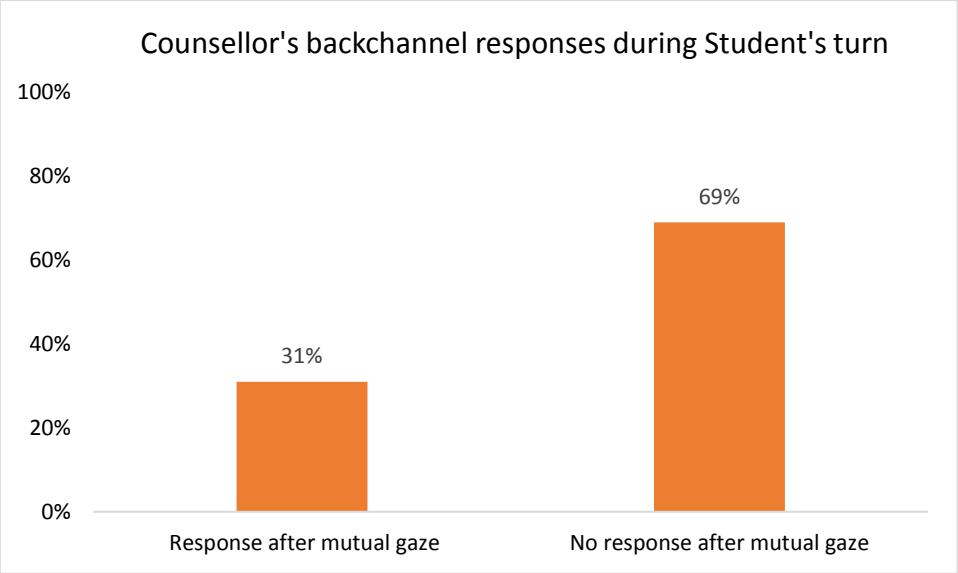


Figure 5.3: Counsellor's production of backchannel responses during Student's turn in Russian in relation to mutual gaze (n=65).

As we can see from Figure 5.3, mutual gaze is in one third of the cases (n=21, 31% of the data) followed by a backchannel response. These distributions indicate that gaze alone is not, on its own, enough to mobilize a backchannel response, as suggested previously by Bavelas et al. (2002). We can assume that gaze functions together with other factors, such as syntax, intonation, pauses and action performed (Stivers &

Rossano 2010) to mobilize a backchannel response from the recipient. Nevertheless, we will need a larger dataset from different conversational settings to investigate this question further.

In sum, previous sections have shown that most backchannel responses in the present dataset are produced through embodied resources (e.g. head nods, facial expressions). Furthermore, the analysis of the relationship between speaker's gaze and backchannel responses in this setting suggests that interpreters are more likely to produce nonverbal backchannel during mutual gaze with the current speaker than primary participants (counsellors). These differences can be, in part, attributed to the different participation frameworks they are engaged in as listeners. The findings also indicate that gaze is indeed important, but it is not, on its own, sufficient to mobilize a backchannel response. The visual dimension of interaction may change our approach to the study of backchannel responses and reveal some new aspects in the organization of interpreter-mediated interaction that would remain 'invisible' in a purely 'linguistic' or audio-based studies.

5.3.3. Directionality of backchannel responses

In the present dataset, interpreters always gaze at the current speaker while producing backchannel responses. They are thus oriented at one interlocutor at the time while listening and memorizing what is being said. Primary participants (the counsellor and the student), on the other hand, display various patterns in their visual orientation during the production of backchannel responses. In the following sections, I will discuss the impact of primary participants' gaze orientation during the production of backchannel responses for the conversational dynamics of an interpreter-mediated talk. For the sake of consistency and clarity, I will focus on the counsellor's backchanneling behavior.

5.3.3.1. Interpreter-directed backchannel responses

When the interpreter is speaking, the counsellors in most cases gaze at the interpreter while producing backchannel responses (Figure 5.4 below).

An example of this pattern is shown below (Extract 2). Prior to this excerpt, the counsellor has asked the student what he is planning to do after he graduates. The following extract provides the interpreter's rendition of the student's answer. The green dots in the figures represent the gaze direction from the perspective of the counsellor (figure 1 and 2) and from the perspective of the interpreter (figure 3)[60].

Extract 2

1 **INT** **ik zou een meer dynamische plaats willen zoe+ken #zoals+**
I would like to look for a more dynamic place such as uh

Cns gaze to Int----------->l.7
cns **+double nod+**
fig #fig 1

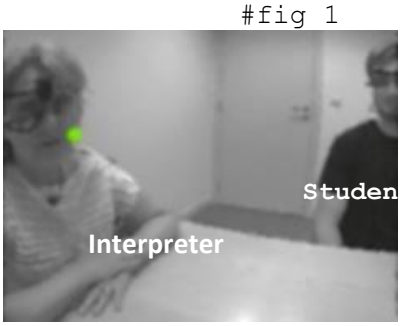


figure 1 (CNS gaze)

2 **INT** **euh verenigd koninkrijk of australië**
United Kingdom or Australia

3 **CNS** **+mm hm#**
cns **+nod-->**
fig #fig 2-3

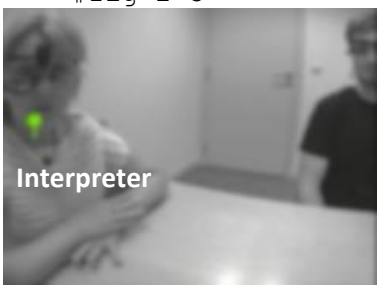


figure 2 (CNS gaze)

figure 3 (INT gaze)

[60] The images are somewhat blurred for the reason of anonymization.

```
4   INT   .h+ ik wil nog een beetje euh r+eizen.+
          .h I  want to travel a bit uh more.
    cns   ->+                              + nods +

5   CNS   ja.
          yeah.
6         (0.3)

7   CNS   ja dat kan ik begrijpen
          yes I can understand that
    Cns   gaze to Stu--------------->
```

The interpreter renders the student's talk into Dutch in the first person ('I would like to look for'). Note that she is oriented towards the counsellor while rendering the student's preceding turn. As soon as the interpreter has finished formulating the turn-constructional unit (TCU) 'I would like to look for a more dynamic place', the counsellor produces a nod to indicate her understanding of the prior utterance, while gazing at the interpreter (figure 1). Then, at the end of the TCU in line 3, the counsellor produces a vocal continuer ('mm hm'), accompanied by a double nod, while keeping her gaze focused on the interpreter. By orienting to the interpreter in such a way, the counsellor foregrounds the interpreter's speakership. At the same time, the student - who has no understanding of Dutch - is left outside of the participation framework created by the counsellor's and interpreter's mutual orientation; he is momentarily excluded and relegated to the status of a 'bystander' (Goffman, 1981). Only at line 7, when the counsellor takes the turn, does she orient her gaze at him. This excerpt fits nicely into Davidson's (2002) model of interpreted discourse, according to which interpreter'-mediated interaction consists of two separate (monolingual) dialogues, between the interpreter and each of the primary participants. This implies that only one dyad is 'active' at a time. However, as we will show, the picture is not always that simple. This will be discussed in more detail in the following sections.

5.3.3.2. Primary-directed backchannel responses

Counsellors may also gaze at the student while producing backchannel responses during (or at the end of) the interpreter's turn. In the following extract, the interpreter is rendering the student's turn into Dutch.

Extract 3

1 **INT** **#maar meestal is het in de wee:k+#**
but mostly it is during the week
Cns gaze at Stu----------gaze at wall--->
cns **+double nod--->**
fig #fig 1 #fig 2

figure 1 (CNS perspective)


figure 2

2 **INT** **en rond vijf uur 's a:vonds+**
and around five in the evening
Cns ----------gaze at Stu-->
cns **----------+**

3 **INT** **en om van [plaatsnaam] te+ ko:men**
and to come from [place name]
Cns ----------gaze at wall-->
cns **repeated nods----------+**

4 **CNS** **+Ja.#**
Yeah.
Cns --gaze at Stu-->1.6
cns **+repeated nods & sympathetic facial expression-->**
fig #fig 3-4

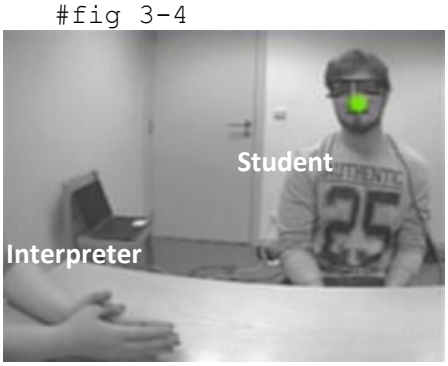


figure 3

figure 4 (INT perspective)

5 **INT** **is dat wel ja+.**
is that actually yeah
cns **---------->+**

6 **INT** **(.)# da's een beetje onhaalbaar**
that's a bit unfeasible
Cns -gaze at wall---------->

fig #fig 5

figure 5

7 **(.)**

8 **CNS** **#Ja, dat is inderdaad nie=niet makkelijk**
yeah, that is indeed no=not easy

cns gaze at Stu----------->

fig #fig 6

figure 6

The student, who lives in a small town outside of Antwerp, remarks that most student activities in Antwerp are organized during the weekdays, which makes it difficult (‘unfeasible’) for him to attend them. While listening to the interpreter’s rendition of the student’s turn, the counsellor gazes away from the interpreter; in line 1 she starts producing head nods while moving her gaze to a middle-distance look (Heath 1986), as shown in Figure 2. This allows her to monitor the actions of both participants, without having to engage in mutual gaze with either of them; towards the end the interpreter’s TCU in line 2, she moves her gaze back to the student while indicating her understanding with head nods. At the end of line 3 (‘and to come from [place name]’), that clearly projects the student’s stance on this matter, she produces the acknowledgment token ‘yeah’ while gazing at the student and making a sympathetic facial expression. The immediacy of her acknowledgment token indicates that she has achieved understanding of the student’s situation, which is followed by a display of affiliation with the student’s stance (see also Pomerantz & Heritage 2013). She

maintains her gaze at the student well into line 8, where she conveys her understanding and endorsement of the student's stance in more substantive terms ('Yeah, that is indeed no=not easy'). Thus, the counsellor displays disengagement from the interpreter and orientation towards the student, who is the primary source of the telling. Such gaze behavior goes against Goodwin's 'rule' that the listener should be gazing at the speaker when the speaker is gazing at the listener (Goodwin 1981). However, the counsellor's gaze orientation can be attributed to the fact that the interpreter does not have the same conversational status as other participants in the dialogue; although she is producing the utterance (as an 'animator' or 'author', Goffman 1981), she does not relate to its content in the same way as the primary speaker. Therefore, the recipient's gaze aversion from the interpreter is not treated as problematic. However, by not gazing at the interpreter while listening to her, the counsellor is reducing her speakership (see also Licoppe & Veyrier 2017).

Primary directed feedback is typically found at the end of the interpreter's turn, when the counsellor turns his/her attention back to the student before initiating a following turn (see Figure 6). Also, the counsellor may provide immediate feedback to the student without waiting for the interpreter's rendition. This is mostly the case when the other participant produces a short turn (e.g. a negating or confirming response).

5.3.4. Dual feedback

5.3.4.1. The phenomenon

In my dataset, I found that counsellors regularly shift their gaze from the interpreter to the student while producing backchannel responses (see Figure 5.4. for an overview). In total, my dataset contained 168 instances of this pattern.

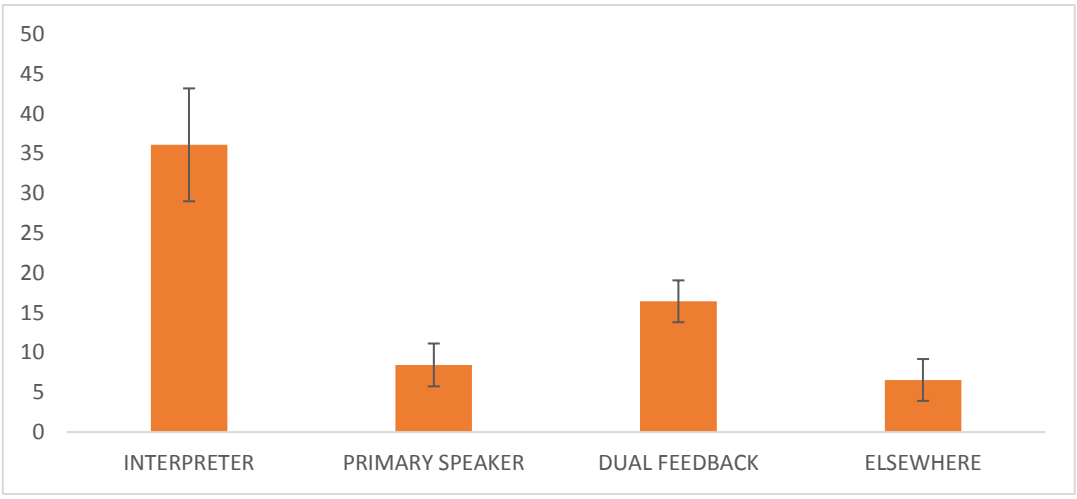


Figure 5.4: Average distribution of counsellors' gaze direction during their production of turn-medial backchannel responses. Error bars show the standard deviation.

We will refer to this to this pattern as *dual feedback* (Vranjes et al. 2018b). Such rapid gaze movements are often produced without a (noticeable) head reorientation towards the other participant and may be difficult to capture with a video camera. An example of this pattern is provided below:

Extract 4

```
1   INT   da's het eerste jaar dat wij eu:h
          it's the first  year that we uh
    Cns   gaze to Int---------------------->

2         eu:h examenperiode in januari#[hebben ],
          uh  have exams    in January
3   CNS                                  [+ah ja+]
                                          I see
    Cns   --------------------------gaze to Stu-gaze to Int->l.5
    cns                                  + nod +
    fig                                #fig 1-2
```

figure 1

figure 2

4 **#+(0.6)**

cns **+double nod-->**

fig #fig 3

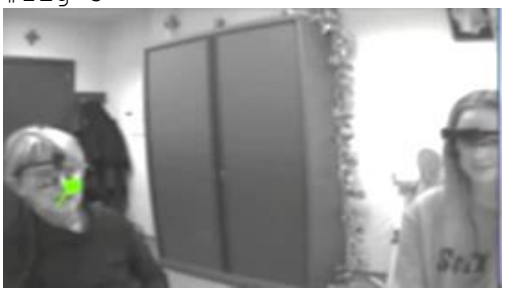


figure 3

5 **INT** **euh vroe+ger hadden wij alle examens voor eenendertig december.**
uh we used to have all the exams before the thirty-first of December

cns --------+

6 **CNS** **(0.9) voor ah ja (.) oké**

Prior to the extract, the counsellor has asked the student if she normally takes exams in January and in June at her university in Russia. While listening to the interpreter's rendition of the student's turn, the counsellor responds to the information that 'this is the first year that we uh have exams in January' with the newsmarker 'ah ja', while shifting his gaze towards the student (Figure 2-3). The counsellor orients towards the 'principal' (Goffman, 1981), whose utterance is being rendered in Dutch by the interpreter. This phenomenon thus illustrates "the simultaneous use of semiotic resources by participants" (Goodwin, 2002: 1490) in interaction.

Drawing on insights from Conversation Analysis and from research on multimodality in interaction, the aim of the following analyses is to investigate such turn-medial backchannel responses that are produced in close correlation with gaze shifts. Can we find some regularities in the production of this phenomenon? Which consequences does it have for the organization of participation in interpreter-mediated talk? After presenting a quantitative analysis on the relationship between gaze shifts and backchannel responses in turn-medial position (section 5.3.4.2), I will focus on the sequential positioning of this pattern in ongoing interpreted talk (section 5.3.4.3). I will argue that recipient's gaze shift towards the other interlocutor during the production of backhannel can be seen as a recipient's heightened display of understanding and

affiliation towards the 'principal'. Also, it will be shown that this pattern plays an important role in maintaining a triadic participation framework in these interactions.

5.3.4.2. Gaze and backchannel responses in turn-medial position

A first step in the analysis of the *dual feedback* pattern was to chart the relationship between the counsellor's gaze and backchannel responses in turn-medial position during interpreter's renditions. As mentioned earlier, turn-final backchannel responses were not taken into account, where it is unclear whether counsellor's gaze is part of next-speaker selection, or whether it is a cue for incipient speakership (Kendon 1967)[61]. Interestingly, 89% of all counsellor's gaze shifts in turn-medial position towards the student (or the 'principal') are coincident with a backchannel response. Thus, when the counsellor shifts his/her gaze towards the 'principal' during the interpreter's rendition, it is most likely to be coincident with some form of backchannel response, be it nonverbal or multimodal. The results indicate that those turn-medial gaze shifts coproduced with a backchannel response together form a "**composite signal**" (Bavelas et al., 2002; Clark, 1996) that contributes to the establishment of reciprocity between the primary participants. Further analysis revealed that 95% of backchannel responses that co-occur with a gaze shift are produced in combination with head nods[62] (and, in a few cases, head shakes or head tilts). As such, this pattern comprises a strong multimodal signal. A possible explanation for this could be that such backchannel responses accompanied by a gaze shift are meant to be seen (and potentially responded to) by the other primary participant. The visibility of the head nod in combination with a gaze shift makes it recognizable as one single action and is thus conducive to making the gaze shift

[61] However, in a few cases, it is clear that the listener did not anticipate the end of the current turn and in those instances, turn-final backchannel responses were in fact a display of continued recipiency (see also Tolins & Fox Tree 2014 concerning a posteriori categorizations).

[62] A head nod is considered to be a "rhythmical vertical head motion consisting of at least one down-up trajectory" (Stivers 2008: 37).

noticeable to the other interlocutor. Also, head nods can be extended in time, which makes it easier to coordinate them with gaze shifts towards the other participant.

The following step in the analysis was to test if some types of backchannel responses were more often produced with a gaze shift than others. This might provide an indication of the work accomplished through this pattern. I annotated all instances of backchannels in the corpus according to the functional types described in Gardner (2001) (see Appendix D). Head nods, head shakes and other embodied signals occurring on their own were annotated as one category 'nonverbal'[63]. In coding backchannel responses, I also took other (prosodic and visual) aspects into account that were relevant in interpreting the function of the backchannel response. For instance, 'ja?' with rising intonation contour was usually used as a continuer, whereas 'ja.' with a falling intonation contour usually functioned as an acknowledgment. In order to test the reliability of the annotations, an inter-coder agreement was performed between two coders (Geert Brône and myself), who independently annotated a random sample (n=34) of backchannel responses from the dataset. Details about the coding can be found in Appendix D. The inter-reliability coding showed a Cohen kappa score K=0.80 (84,8% agreement), which indicates very strong reliability (Cohen 1960).

As shown in Table 5.7 below, newsmarkers and newsmarker-like objects such as *Oh!* and *Ah ja*, which mark the prior speaker's turn as in some sense newsworthy to the recipient (see Gardner 2001), are mostly produced with a gaze shift towards the primary speaker. Acknowledgments and continuers occur in less than half of the cases with a gaze shift. Thus, when displaying change in their state of knowledge, counsellors tend to orient towards the 'principal' or the party with the 'epistemics of experience' (Bolden 2013). This will be discussed further in section (5.3.4.3).

Note that *dual feedback* is produced overwhelmingly by the counsellors, and much less frequently by the students during these encounters. This can be partly attributed to the differences in social roles between the counsellor and the student; the

[63] Given their great variability in form and function, and the fact that they are still relatively understudied, I did not differentiate between them in more detail.

counsellor being the 'representative' of the faculty and the foreign student being a 'newcomer'. In contrast to the students, the counsellors might have been driven by specific 'interpersonal goals' (Clark 1996), namely to establish and maintain contact with the student during the talk, which is also evident from their gaze behavior (see Chapter 3). Also, I will discuss this further in section 5.3.4.5, where I will argue that *dual feedback* is a manifestation of increased interpersonal engagement between primary participants.

Table 5.7: Distribution of gaze shifts to the 'principal' in relation to the backchannel function.

FUNCTIONAL TYPE	**NO GAZE SHIFT**	**GAZE SHIFT**	TOTAL
Newsmarker and newsmarker-like objects (*e.g. Ah ja, Ah, Oh*)	16 (27%)	43 (73%)	59
Assessments (*e.g. leuk*)	2 (40%)	3 (60%)	5
Acknowledgments (*e.g. ja*)	44 (61%)	28 (39%)	72
Continuers (*e.g. mh hm*[64])	39 (59%)	27 (41%)	66

Overall, we can conclude from this analysis that (a) counsellor's gaze shifts towards the student during interpreter's renditions are strongly related to the production of backchannel responses, (b) they comprise a 'composite signal' and (c) newsmarkers and newsmarker-like objects are in most cases produced with a gaze shift towards the student in these interactions. In the following sections, I will argue that this pattern objectifies the double conversational ground between the primary interlocutors and the interpreter, and plays an important role in maintaining a triadic participation framework in these encounters.

[64] Note that in the present dataset, recipients produced 'continuers' (see Schegloff, 1982) such as *mh hm,* almost exclusively with head nods, forming a 'multimodal Gestalt' (Mondada 2014).

5.3.4.3. Dual feedback and triadic participation framework

If we take the counsellor's perspective in the dataset, as described in the previous section, the phenomenon of interest is organized as follows:

(1) The counsellor listens to the interpreter's rendition of the student's preceding turn.
(2) The counsellor shifts his/her gaze from the current speaker (interpreter) towards the student while producing a backchannel response. The student may align by producing a head nod towards the counsellor.
(3) The counsellor returns his/her gaze towards the interpreter who is still speaking. Alternatively, (s)he may maintain his gaze at the other student for a while (see section 5.3.4.5. for an example).

In order to address the question *when* the dual feedback pattern is produced and how it is organized in ongoing interaction, it is necessary to take a closer look at the details of its production in its sequential environment. I analyzed 40 randomly selected occurrences of this phenomenon in my dataset, following a CA-inspired analytical procedure as outlined in Hoey & Kendrick (2017). I analyzed each case individually by taking the following four aspects into account:

- Position — Where does this phenomenon occur in the course of interaction? Which action occurred before (e.g. an assessment, an informing)?
- Composition — How is it realized? Which combinations of resources (verbal/nonverbal) verbal are used (e.g. a head nod in combination with a facial expression)? How do these resources contribute to what is being done?
- Action — What is accomplished through talk?

- Epistemics — What are the epistemic domains of the participants? What do they know? How much do they know relative to the other? (cf. Heritage 2012)

I found that backchannel responses accompanied by a gaze shift occurred in specific sequential environments in the dataset, namely as

- receipt of **newsworthy information** in the context of a (question-elicited) informing or counterinforming (30%),
- display of **affiliation or agreement** with the principal's stance in the context of an assessment (15%),
- display of **confirmation** of something that is within the epistemic domain of the recipient (e.g. after a request for confirmation, display of uncertainty or a statement) (12,5%),
- less clear cases, where the listener seems to merely display **uptake** and alignment with the ongoing talk, usually in the context of question-answer sequences (43,5%).

I then compared dual feedback to 40 randomly selected 'regular' backchannel responses (oriented to the interpreter) and found that 'regular' backchannel responses occur less in the same contexts: receipt of newsworthy information (2,5%), display of confirmation (2,5%) and none as display of affiliation or agreement. Although display of uptake in the context of question-answer sequences is the most frequent environment for *dual feedback* (43,5%), it is even more frequent with regular backchannel responses (80%). Therefore, in the following, I will focus on the two sequential environments in which *dual feedback* most frequently occurs: receipt of newsworthy information and display of affiliation.

5.3.4.4. Receipting newsworthy information

In the current dataset, *dual feedback* is frequently produced to register a change in the counsellor's state of knowledge towards the student (or 'principal'). This is usually the case in the context of **(elicited) informings**, which are typically constructed to deliver some information for a non-knowing recipient (see Thompson et al. 2011). I illustrate this in Extract 5. Prior to what is presented in this extract, the counsellor has asked whether secondary schools in Russia offer their students any help with the choice of a study domain for higher education, for example by helping them gather information or the like.

Extract 5

1 **INT** **dus het laatste ja+ar van 't school is het *tiende j*aa+r#,**
so the last year of school is the tenth year
Cns ---------gaze to Int--------------------------------------->
cns **+expansive nod----------------------+nod->l.3**
stu ***nod-----***
fig #fig 1

figure 1 (CNS gaze)

2→ **#(0.3)**
Cns -gaze to Stu-->
fig #fig 2

figure 2

3→ **INT** **euh+ w#orden er verschillende cursussen georganis+eerd#,**
uh various courses are organized
Cns -->gaze to Int--Stu-->
cns **->--+** **+surprised face->**

fig #fig 3

figure 3

#fig 4

figure 4

4 **(0.3)+**

Cns ---->
cns ->---+
5 **INT** **+.h euh #*waar euh* +de capaciteiten**
uh where uhm the capacities
Cns -----gaze to Int-------------------->l.10
cns +nodding-----------+
stu *double nod*
fig #fig 5

figure 5

6 **van de l+eerlingen worden getest,+**
of the pupils are tested
cns +nod---------------------+
7 **CNS** **mm [hm,**
8 **INT** **[en op +basis daarvan+ wordt**
and on the basis of which
cns +nod----------+
9 **een a+anbeveling+ +gedaa[#n**
a recommendation is made
10→ **CNS** **[ah *ja**
I see
Cns -------------------------gaze to Stu->
cns +nod-------+ +nod--------->l.12
stu *nod-->

fig #fig 6

figure 6

11 **(0.3)#***

Cns -----gaze to Int--->

stu **----->***

#fig 7

figure 7

12 **INT** **v+oor wel+ke (.)# richting (0.2) #euh**
which study domain uh

Cns ----------Stu->

cns **-+ +nodding----------->**

fig #fig 8 #fig 9

figure 8 figure 9

13 **zou (je) (.) beter+[(.)*kie+zen*].**
(you) should optimally choose.

14 **CNS** **[ah ja oké]**
I see okay

Cns ----------Int-->

cns **----------+head tilt to the side+**

stu ***nodding***

15 **CNS** **ja. (0.2) en is ook jouw keuze bepaald.**
yes. and has your choice also been determined.

Instead of starting with a type-fitted response (a confirmation or a negation) to the counsellor's question, the interpreter begins with an informing action ('so the last year of school is the tenth year') in line 1, introducing a contrast with the situation in Belgium (where the last year of secondary school is usually the sixth year). The counsellor starts producing an expansive nod with an upward movement of the head, which "foreshadows" his receipt of this informing as news (Whitehead 2011). At the end of the turn-constructional unit (TCU), he displays uptake by producing another head nod and shifting his gaze to the student (figure 2) during a 0.3-second pause (line 2). Thus, the gaze shift and the onset of the second nod are nicely timed to the point at which the interpreter's TCU reaches its possible completion. Interestingly, the student aligns by producing a head nod towards the counsellor. She does not look at the interpreter, who is rendering her preceding turn into Dutch, but maintains her head and gaze at the counsellor for the most part of the interpreter's rendition. As argued by Merlino & Mondada (2014), by gazing at the counsellor during the interpreter's turn, the student displays herself as the 'principal' of the current talk. The counsellor immediately shifts his gaze back to the interpreter (figure 3), who is about to initiate the following TCU, thereby displaying his continuing attention and showing that he has no intention to take the turn.

The interpreter continues with the informing and at the end of the following TCU 'uh various courses are organized' (line 3), the counsellor again shifts his gaze towards the student (figure 4) while displaying surprise to the information provided with his facial expression (lifted eyebrows and lowered chin) (see also Ruusuvuori & Peräkylä 2009 on showing stance through facial expressions). This action marks his receipt of the content of the TCU as newsworthy in some way and is then followed by a series of nods.

3→ **INT** **euh+ worden er verschillende cursussen georganis+eerd#,**
uh *various courses are* *organized*
Cns ->gaze to Int----------Stu-->
cns ->--+ **+surprised face->**
fig #fig 4


figure 4

The interpreter appears to follow the counsellor's gaze movement as she shifts her gaze to the student as well. Again, the student acknowledges the counsellor's display of surprise by producing a head nod towards him, although she does not understand Dutch. The counsellor then moves his gaze back to the interpreter (figure 5), who continues with the specification of the content of those courses. During the interpreter's utterance, the counsellor starts nodding and at the end of the TCU (line 6), he produces a continuer ('mm hm'). Unlike in previous cases, the counsellor maintains his gaze at the interpreter, as he is merely indicating to her that the utterance has been received and that she may continue (Tolins & Fox Tree 2014). Also, the student does not display alignment here with a head nod.

Following the information that 'a recommendation' (line 9) is made on the basis of those tests, the counsellor starts producing a nod with an upward head movement followed by the newsmarker 'ah ja' (*I see*), while shifting his gaze towards the student (figure 6). The counsellor thus addresses successively both interlocutors with his gaze while producing the listener response. Through this composite signal, he displays a change in his knowledge state and at the same time acknowledges the other interlocutor's epistemic authority (Bolden 2013, Heritage 2012) over the content of the talk. Again, the student aligns with this expression of change in counsellor's knowledge state by responding with a head nod towards the counsellor. There appears to be some form of **preliminary (or semi-)grounding** at that stage: although the student cannot

check on what the interpreter says at the moment, she acknowledges the counsellor's change in information state[65].

In line 12, then, the interpreter continues with further specification 'which study domain uh (you) should optimally choose'. As she utters 'study domain' (*richting*), she starts moving her head and gaze towards the student (figure 8), while lifting her hand and gesturing in the student's direction. The counsellor immediately follows by directing his gaze at the student (figure 9) and maintaining it at her. In contrast to the previous cases, this gaze shift was brought about by the interpreter's gaze and head movement towards the student (for 'gaze cueing', see Frischen et al. 2007). Also, the gaze shift occurs towards the end of the interpreter's turn, which is marked by slowing down of the pace of the interpreter's rendition. As it is unclear how this gaze shift is motivated, I did not count it as an instance of *dual feedback* in the analysis. At the end of the interpreter's turn, the counsellor produces 'ah ja' followed by 'oké' (okay), which is used to pre-figure changes in topic or activity (Gardner 2001: 55).

Another example is provided in Extract 6. The counsellor had asked the student what type of program she is following as part of her exchange in Belgium (not presented in the transcript). Here, the interpreter starts her rendition with the **counterinforming** (Heritage 1984) 'it is not an exchange' (line 1), that explicitly rejects the counsellor's presupposition. The counsellor immediately responds to this TCU with the newsmarker 'ah oké', indicating a revision in her information state as a result of this counterinforming, while moving her gaze to the student. The student responds to the counsellor's action with head nods while maintaining her gaze at her.

Extract 6

```
1   INT    .h het is niet uitwisseling,#
           .h it is not an exchange
    Cns                      gaze to Int--->
```

[65] This brings us to another issue in interpreter-mediated dialogue, namely **the issue of trust**. The student takes it for granted that the interpreter is rendering everything accurately. Although being basically ignored in interpreting literature, the issue of trust seems to be very much at stake here (see Vranjes et al. 2018).

```
    fig                                  #fig.1
```

figure 1

```
2           #(0.2)
    cns     gaze to Stu--->
    fig      #fig.2
```

figure 2

```
3           e[+uh:m ik kwam+] naar hier om volledig e:h
            uhm I came here to complete euh
4   CNS      [ A:h oké      ]
    Cns     ----------------------gaze to Int-------------->
    cns       +  head nod  +
```

Previous sections have suggested that, during the interpreter's ongoing turn, the counsellor displays his just-changed state of knowledge to the 'principal' and not to the interpreter. Gaze is in those instances "wholly interactive and a function of a communicative intent" (Lund 2007:299). By displaying a change in the epistemic state directly to the student, the counsellor reduces the epistemic asymmetry (Bolden 2013) between himself and the student already *before* he can take a turn-at-talk. The student aligns with this change in epistemic position by producing a head nod towards the counsellor. With *dual feedback*, the counsellor also acknowledges the epistemic authority of the student over what is being said (as the 'principal'); although the interpreter has knowledge of the other language and of the content of the student's talk, she and the student have different states of knowledge in relation to what is being expressed.

5.3.4.5. Displaying agreement or affiliation

The following example shows how dual feedback can perform a 'social-affiliational' function during the interaction. We have found that counsellors produce dual feedback in the environments where student's stance toward a situation is expressed (as being funny, sad, awful or exciting). *Stance taking* is another social activity in which intersubjective understanding between interlocutors is publicly displayed. The most typical and recognizable environment of stance taking are assessments, viz. "utterances that are positively or negatively balanced through use of specific lexically assessing terms" (Thompson et al 2015:139). Assessments typically make relevant an affiliative uptake on the part of the responder that mirrors the speaker's conveyed stance (see Stivers 2008). A first example is provided in Extract 7. Prior to the extract, the student was telling that he will be abroad during the summer, which was then positively assessed by the counsellor in the following turn (not represented in the transcript). Here, the student recounts that it is actually not the best timing to go abroad. In the following excerpt, we provide the interpreter's rendition of the student's turn.

Extract 7

1 **INT** **.h het is gewoon zonde van de timing +#(.)omdat in d+e zomer**
.h it is just pity for the timing because in the summer
Cns **----------**gaze to Int**----------**>1.3
cns **+ head nod +**
fig #fig.1

figure 1

2 **euh zijn er heel veel interessante activiteiten**
uh there are a great many interesting activities
3 **in (plaatsnaam),+# (.) waar ik euh (0.2) nu verblijf**
in (place name) where I uh am currently staying
Cns ----------gaze to Stu----------->1.5

cns→ +sympathetic facial expression & slow nods->

fig #fig.2

figure 2

4 **(0.5)+**

cns ---->+

5 **INT** **[e:n] euh:: ja dat zal ik wel moeten missen.**
and uh yeah I shall miss out on that.
CNS **[J#a]**
Yeah
Cns ---gaze to Int----------->
fig #fig.3

figure 3

6 **CNS** **ja (-) ah dat is inderdaad dan wel (-) wel jammer.**
yeah oh that is indeed really really pity then.
Cns -gaze to Stu----------->

At the end of the TCU 'pity about the timing' (line 1), the counsellor produces a single head nod, while maintaining her gaze at the interpreter (figure 1). Note that at that point, the counsellor does not yet have full access to the reason as to *why* it could be a bad timing for the student to go abroad. In order to be able to affiliate with the expressed stance, the recipient has to understand why such stance is taken in the first place. Only after the interpreter has provided the reason for it in lines 1-2 ('because in the summer there are great many interesting activities') does the counsellor gain full understanding of the student's stance. She immediately displays her increased understanding and

affiliation (see also Stivers 2008) with the expressed stance by shifting her gaze towards the student (figure 2), who is looking at her, and making a sympathetic facial expression accompanied by a series of slow, repeated nods[66]. The counsellor's affiliative backchannel response is directed to the 'principal', whose views are being expressed and who retains primary epistemic rights over what is being said, and not merely to the interpreter who produces the utterance. The counsellor and the student stay engaged in mutual gaze for 2 seconds. As noted by Kendon "extended mutual gazes appear to be indicative of an intensifying of the direct relations between the participants" (1967:48). In other words, there is a moment of **convergence in stance** (see also Haddington 2006[67]) between the primary participants during the interpreter's rendition. This is important for strengthening the interpersonal relationship between primary participants and for the establishment of a triadic participation framework. After the student withdraws his gaze, the counsellor turns her attention to the interpreter again (figure 3). At the end of the interpreter's turn, the counsellor verbalises her understanding and affiliation with 'that is indeed really pity' (line 6). This example thus shows how "[t]he informational and social-affiliational functions of common ground are closely interlinked" (Enfield 2008: 229) in social interactions. That is, only by establishing a certain level of common ground with our interlocutor, are we able to affiliate in a socially relevant way with that person.

Another example is shown in Extract 8. The counsellor has asked the student, who had a cold the week before, if she is feeling better now (not represented in the transcript). After telling the counsellor that she has recovered from her cold, the student starts assessing the currently unstable weather in Belgium.

[66] It has been shown that the more affiliative, supportive tokens "are more varied in intonation, in lexical selection and also in length" (Müller 1996, in Gardner 2001: 20).

[67] Whereas Haddington (2006) demonstrated the importance of mutual gaze between the speaker and recipient during 'second assessments' (see Thompson et al. 2011), I show how a display of convergent stance can also be achieved during a speaker's ongoing turn in multiperson interactions.

Extract 8

```
1   INT    dat het dan wel (0.2) euh +moeilijk is om in te schatten,
           that it is then        uh difficult to estimate

    Cns    >>-gaze to Stu-----------gaze to Int----------------->l.3
    cns    head nods----------------+

2   INT    omdat het zo# koud is 's morgens  (.)
           because it is so cold in the mornings
                        #fig.1
```

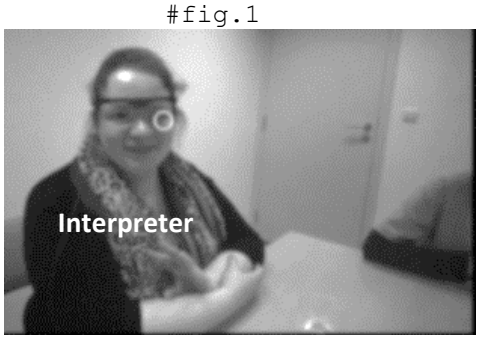


figure 1

```
3          warm 's [middags],
           warm in the afternoon
4   CNS→            [+ Ja:+#]
                      Yeah
    Cns    ----gaze to Stu----->
    cns              +sympathetic facial expression & tilted nod+
    stu    >>gaze to Int-----gaze to Cns-->l.7
    fig                    #fig.2-3
```

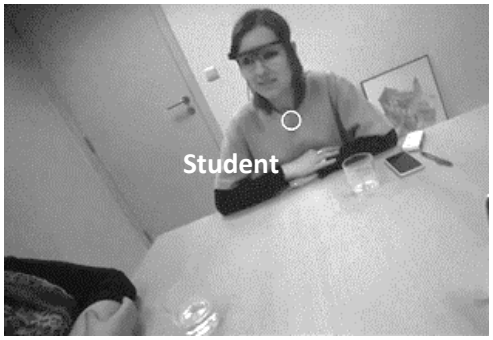


figure 2

figure 3 [STU perspective]

```
5   INT    *en [vooral als je] dan +ook nog* heel veel+ wind,
            and  especially if you then also a lot of wind
6   CNS      [  +   Ja:  +  ]
                  Yeah
    Cns    ----------------------------------------gaze to Int->

    cns          +   nod  +         +----head nods----+

    stu    *repeated nods-----------------*

7   INT    en als je dan fietst,
           and if you ride a bike then,
    Cns    -------------------------------->>
```

Line 1 starts with the assessment that it is 'difficult' to estimate how the weather will be given the fluctuating temperatures in Belgium ('so cold in the mornings' and 'warm in the afternoon'). Towards the end of the TCU at line 3, the counsellor produces a prolonged acknowledgment token ('Ja:'), accompanied by a tilted nod and a sympathetic facial expression[68] (see figure 3), while shifting her gaze towards the student. Thus, she orients towards the previous utterance as an assessment, that makes expression of convergent stance and engagement with the student ('principal') relevant. Through her tilted nod and face she expresses this 'uncertainty' in weather conditions in an embodied way, while displaying agreement with the student's stance. This is followed by a second acknowledgment 'Ja:' in overlap with interpreter's speech (in line 6). Her prolonged acknowledgment tokens 'Ja: (.) Ja:' do not only signal her strong agreement and affiliation with student's stance, but also her own epistemic access (Heritage 2013) to the situation. Thus, the counsellor empathises with the student from a 'knowing' position, being Belgian herself. In response to counsellor's first acknowledgment 'Ja:' and head tilt directed to her, the student shifts her gaze to the counsellor (line 4) and starts producing a series of repeated nods (line 6). Similar to the previous extract, convergent stance is not established with the interpreter, who is the 'animator' of the primary speaker's assessment, but with the 'principal'. In this way, counsellor momentarily establishes interpersonal bond with the student during the interpreter's turn.

[68] viz. raized eyebrows and corners of the mouth pointed downwards.

5. Conclusion

This chapter makes a contribution to the almost non-existing multimodal accounts of backchannel responses in interpreter-mediated interactions and, more broadly, to multimodality research that takes into account participants' simultaneous employment of various semiotic resources (Goodwin 2002). The first part of the chapter provided a quantitative analysis of the occurrences of backchannel responses in the dataset, showing that most backchannel responses are situated at the nonverbal level. Second part of the study focused on the role of gaze as a backchannel response mobilizing cue, suggesting a relationship between interpreter's nonverbal backchannel responses and mutual gaze with the current speaker. The main part of the study focused on the directionality of backchannel responses in interpreter-mediated talk and the *dual feedback* pattern. I demonstrated the crucial role of recipient's gaze direction in the production of backchannel responses and in the establishment of a triadic participation framework in interpreter-mediated interactions.

Extending Davidson's (2002) analysis that focused on the achievement of reciprocity at the verbal level in interpreter-mediated talk, the current study has shown how participants who have no access to each other's language can nevertheless achieve reciprocity of understanding through gaze and backchannel responses. *Dual feedback* thus appears to further the accomplishment of common ground and reciprocity between the primary participants and, as such, objectifies the **double conversational ground** between the primary interlocutors and the interpreter. Also, it exemplifies the relevance of backchannel responses in interpreter-mediated talk not just for the achievement of common ground or intersubjectivity in interaction, but also on an interpersonal level. As argued by Enfield "our efforts to maintain and build common ground have significant consequences for the type of relationship we succeed in ongoingly maintaining, that is, whether we are socially close or distant" (Enfield 2006: 230). *Dual feedback* appears to be an *active* communicative signal, produced at specific moments during the interpreter's rendition, through which the recipient tries to engage with the 'principal' of the current talk.

Furthermore, through *dual feedback*, recipients in an interpreter-mediated encounter display their **orientation to the participation status and speaking rights** of their interlocutors. While producing *dual feedback*, the recipient addresses both participants, who are momentarily acting as 'consociates' (Bolden 2013) and who both have the epistemic access to the content of the talk, but not the same epistemic authority over what is being said. Although the 'principal' relies on the interpreter to express her experience or thoughts in the other language, (s)he nevertheless retains the epistemic authority as 'experiencer' of those thoughts and experiences.

The analysis has shown that *dual feedback* occurs strikingly often with **newsmarkers** (Gardner 2001), pointing to the fact that the recipient wants to indicate directly to the 'principal' that the information is newsworthy in some way and that a change in the knowledge state has occurred. Furthermore, we have demonstrated how the recipient displays **affiliation** with the 'principal', whose stance is being expressed, rather than with the interpreter, who is merely producing it in the recipient's language. So even when producing backchannel responses, the recipient orients to the "socially distributed rights to knowledge" (Bolden 2013) in the ongoing interaction. As Wadensjö (1998: 93) puts it: "To interact means to continuously evaluate others' and one's own relation to a focused discourse". Thus, *dual feedback* shows that participants in an interpreter-mediated talk are cognizant of their own and others' position in the triad and that this feeds into a continuous process of 'interaction management' in this type of interaction.

I have also suggested that the recipient's redirection of gaze towards the 'principal' provides for a subtle **shift in the participation framework** during its production, which helps to maintain a triadic interaction pattern. As such, *dual feedback* appears to be a manifestation of increased social engagement between the primary participants and it throws light on new ways how temporarily "excluded" participants in this type of interaction are actually involved, even when they do not understand the language. Thus, although they are dependent on the interpreter's renditions of the talk to achieve mutual understanding, the construction of reciprocity and the management of

the interaction on the level of the triad is to a large extent in the hands of the primary participants.

Finally, the present chapter has illustrated the importance of **a multimodal approach** in gaining a better understanding of the intersubjective relations among the interlocutors in an interpreter-mediated talk. Moreover, a detailed study of synchronous gaze behavior can help us gain insights into the specific multimodal character of an interpreter-mediated encounter, which can further contribute to our knowledge of the subtleness in the multimodal exchange. Future research will be needed to investigate the *dual feedback* pattern in other constellations in face-to-face (interpreted or same-language) interactions, which will undoubtedly broaden our understanding of the phenomenon.

Chapter 6

This chapter presents a final closing case that illustrates some of the phenomena presented in the previous chapter within the context of a naturally occurring, interpreted therapeutic encounter that was recorded in a mental healthcare institution. Without aiming for the same level of analytical depth, this case study builds an argument in favor of the ecological validity of some of the previous findings. In addition, since there has been very little research on interpreting in the context of mental healthcare that takes the visual dimension of such encounters into account, the aim of this study is to contribute to that line of research and to open up some new questions for future investigations.

After presenting the data and providing a brief literature overview on interpreting in a psychotherapeutic context, this study focuses on the interpreter's and the therapist's production of backchannel responses. It starts with a quantitative overview of backchannel responses during the session. In the subsequent section, I discuss the interpreter's production of backchannel responses and how they relate to mutual gaze with the patient. Finally, I discuss the therapist's production of *dual feedback* during the session[69].

6.1. Presenting the case

The therapeutic session described in this chapter was recorded at a mental health institution in the Netherlands. We obtained the permission to record the last session of the patient's treatment. The patient was a Russian-speaking refugee of Armenian

[69]This chapter is largely based on Vranjes et al. (2018a) and Vranjes et al. (*forthcoming*).

descent with a very limited knowledge of Dutch. As he had been in therapy in that institution for several months, he was accustomed to speaking in the presence of an interpreter. The session was conducted by the therapist who had almost no understanding of Russian. A professional interpreter was present to relay the utterances of the primary parties from one language into the other. The interpreter had twenty years of experience in mental healthcare and was positively evaluated by the therapist. The conversation lasted about 47 minutes, of which the first 32 minutes were analyzed in detail. All three participants knew each other prior to the recording session. The study was approved by the KU Leuven ethics committee, the institution's ethics committee and all participants agreed to be recorded by signing a written informed consent (in Dutch and in Russian), which ensured their anonymity and stated how the data were going to be used and presented.

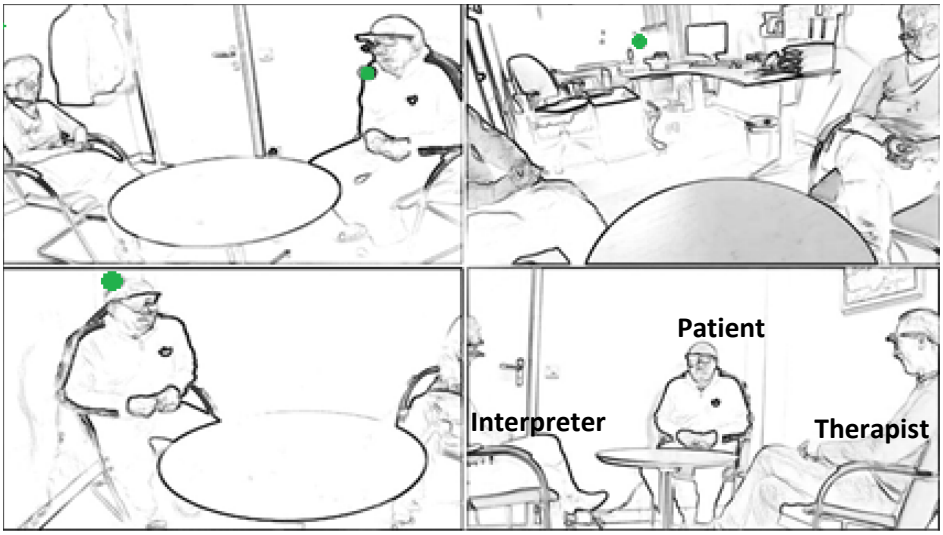


Figure 6.1: Three dynamic internalized perspectives generated by the eye-trackers and the profile shot. The dots indicate the gaze direction of each of the participants.

The recording took place in the therapist's consultation room. As preferred by the therapist, the participants were sitting in a triangular formation around a small table, with the patient in the middle, the interpreter on the right and the therapist on the left side of the patient (Figure 6.1).

In general, the recording procedure and the method were the same as described in Chapter 2. Each participant wore eye-tracking glasses. The patient was wearing the

Pupil Pro Monocular Glasses, whereas the therapist and the interpreter were wearing the *Arrington Gig-E60.* Also in this case we positioned an additional camera (*Sony HDR-FX1000E*) on the other side of the room to record the whole interaction from a side-view. The recording equipment was set up prior to the conversation, so that the eye trackers could be calibrated immediately upon the arrival of the therapist, interpreter and the patient (in that particular order). The participants were left alone in the consultation room during the recording. At the end of the session, a short interview with the participants was conducted, in which they were asked about their background and impressions about the encounter[70].

6.2. The setting: Interpreting in psychotherapy

Psychotherapy can be defined as a 'talking cure' that takes place within a privileged relationship between therapist and patient (Bot 2005: 5). During the encounter, the therapist is oriented towards the patient in a non-intrusive way in order to provide space for the patient to open out with his/her telling (Gardner 2001: 34). As previous studies have shown, therapists employ backchannel responses (such as *mm hm, yeah, right* and head nods) to delicately manage collaboration with the patient (Gerhardt & Beyerle 1997, Gardner 2001, Muntigl et al. 2012, Muntigl & Horvath 2014). The differential use of these backchannel responses reflects various levels of the therapist's involvement in the patient's telling (Gerhardt & Beyerle 1997, Muntigl 2012). Psychotherapy is thus particularly interesting for the study of backchannel responses, as these have been shown

[70] In the immediate post-interview with the participants, the interpreter and the therapist stated that they were 'fairly aware' of the fact that they were wearing the eye-tracker during the encounter; the patient on the other hand declared that he had 'almost no awareness' of the eye tracker during the conversation. Again, the recording equipment may have influenced their gaze behavior but we do not know to what extent. As it was a naturally occurring therapeutic encounter in a psychotherapeutic institution, each of the participants was there with a set purpose: the therapist was oriented toward the realization of specific goals in that therapeutic session, the interpreter was focused on her task of relaying the talk and the patient might also have had his own agenda for his last session.

to play an important role in establishing and maintaining an affective therapist-patient relationship (Gerhardt & Beyerle 1997, Muntigl et al. 2012, 2014).

The question is how this relationship is maintained when the therapist and the patient do not share a common language and communicate with the aid of an interpreter. Although professional interpreters in mental healthcare are expected to adopt a neutral and objective role, previous studies have shown that they do not only act as mere conveyors of the patient's and the therapist's words, but also influence the therapeutic experience (Bot 2005). Bot (2005: 79) argues that "the interpreter [is] a person who, through his presence, also helps to form a therapeutic reality. The essence then is the interaction between the therapist, patient and interpreter".

Apart from the studies by Wadensjö (2001) and Bot (2005), very little empirical research has been done on interpreter-mediated therapeutic encounters from a multimodal perspective, taking both verbal and visual semiotic resources into account. This is partly due to the sensitive nature of psychotherapy, which makes it an extremely difficult task to get permission to video record the sessions. Thus, to date, nobody has investigated how interpreters and therapists manage their recipiency and collaborate with the patient's telling moment-by-moment during the therapeutic session.

In therapeutic sessions, it may be challenging for the interpreter to determine what is the most appropriate way to respond to 'empathic moments' (Heritage 2011) in the patients' telling. As noted by Wadensjö (2001: 83), interpreters do not receive the same training as therapists:

> "While therapists are trained to listen attentively and respond, interpreters are trained mainly to mobilize another mode of listening: to listen attentively, render what they have heard in a new version and avoid direct response."

Although interpreters are not supposed to show their own attitude and reactions to the primary speaker's utterances, they are the ones with first and primary access to the patient's telling. Therefore, "there must be some way in which they negotiate their 'recipiency'" (Gavioli 2012: 201). The therapist, on the other hand, is unable to directly

contribute and respond to the emergent story of the patient and has to wait for the interpreter's rendition.

Drawing on the insights from Conversation Analysis (Gardner 2001, Stivers 2008, Peräkylä 2013), the following sections present an analysis of the interpreter's and the therapist's backchannel responses during the patient's extended turns (cf. Goodwin 1986). I will first provide an overview of the participants' use of backchannel responses. After that, I will examine the interplay between gaze and backchannel responses in this session. Finally, at a more general level, I will discuss the implications of these findings for the interpreting practice and for the study of interpreter-mediated psychotherapeutic interaction.

6.3. Distribution of backchannel responses

In the following, I discuss the participants' use of backchannel responses during the encounter (see Table 6.1. below). The patient displayed a low level of engagement during the conversation, which was manifested in his overall low production of backchannel responses. Therefore, the further analysis will focus on the interpreter's and the therapist's use of backchannel responses.

Table 6.1: Production of backchannel responses during the session.

LISTENER	**SPEAKER**				
	Therapist	Interpreter NL	Interpreter RU	Patient	Total
Interpreter	1	/	/	75	76
Therapist	/	136	4	19	159
Patient	3	4	6	/	13

During the session, the interpreter provided backchannel responses almost exclusively during the patient's turns (for similar observations, see Chapter 5, section 3.1). She displays active involvement in the patient's extended turns, which may require more listener-support and participation. Interestingly, the interpreter produced almost no backchannel responses during the therapist's turns. This can be accounted for the fact

that the therapist is more oriented towards the patient (66% of his gaze fixations) than towards the interpreter (14% of his gaze fixations) while speaking, and thus appeared to display no need for listener support from the interpreter during his turn (see Appendix E for an overview of the gaze distributions). It may also be attributed to the content of the therapist's turns, which were very often questions, formulations and interpretations of the patient's prior turn (see also Peräkylä 2013). Thus, the interpreter's production of backchannel responses is indicative of the structural asymmetry of this triadic interaction.

The therapist in this session displayed a very high level of participation and engagement during the interpreter's turns. Overall, the therapist produced 159 backchannel responses during the talk. Although most of them were produced during the rendition of the patient's turns in Dutch, there were also a few cases when the therapist produced a backchannel response during the patient's turns in Russian. Those backchannel responses were either a sign of recognition or understanding of the patient's utterance, or had a mere aligning function with respect to the ongoing structure of the interaction (for example, nodding to the patient's displayed intention to hold the floor). Although these cases are in themselves worth exploring further, they will not be discussed in the present dissertation.

6.4. Mutual gaze and interpreter's backchannel responses

6.4.1. Gaze as response inviting cue

In the following, we focus on the interpreter's backchannel responses during the patient's telling. A first case is shown in Extract 1. The therapist had just asked the patient how he has experienced his stay at the mental health institution. After the interpreter's rendition and a long silence, the patient starts off by expressing his stance with 'it was very bad' (line 4), while looking at the desk in front of him. There is a slight pause, but the interpreter remains silent. Instead of producing a continuer or taking the turn to render his utterance, she waits for more information. The absence of a listener

response here is thus illustrative of the interpreter's coordinative role; she judges if the patient's telling is sufficient to start rendering.

Extract 1

1	**THER:**		**en en hoe eu:h hoe heeft u het hier=eu::h gehad?**
			and and how euh how was it here euh for you?
2	**INT:**		**как (.) вам сдесь жилось всё это время?**
			how (.) was it for you to live here all this time?
	Pat		--gaze to Int----------gaze away--> l.8
	Int		gaze to Pat----------> l.8
3			(1.6)
4	**PAT:**	→	**сначало очень (0.8) плохо был,**
			in the beginning very (0.8) bad it was,
5			**(0.7)**
6			**не то что.**
			not because.
7			**(.)**
8		→	**у меня #ситуация был пло+хо#.**
			my situation was bad.
	Pat		----------gaze to Int->
	Int		---------->
	int:		**+double nod -- >>**
	fig		#fig. 1 #fig.2

figure 1

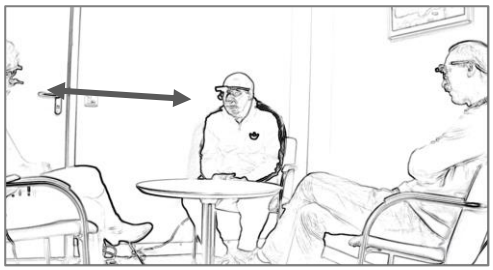


figure 2

The patient continues in line 6, but breaks off his talk. He then self-repairs (Schegloff et al. 1977) in line 8 by explicating that his own situation ('my situation') was bad, and not the circumstances at the institution, as he explains later on. While producing the word 'bad' in line 8, he shifts his gaze towards the interpreter, who immediately responds with a head nod. Thus, by orienting his gaze at the interpreter, the patient selects her as the addressee and responder of his telling (cf. Stivers & Rossano

2010). Immediately after her response, he moves his gaze away and continues his talk. This example shows how the establishment of mutual gaze can elicit the interpreter's head nod and thus functions as a response inviting cue (see Heath 1992, Stivers & Rossano 2010).

In excerpt 2, we find a similar pattern. Prior to this extract, the patient was telling that he was in a state of shock when he came to the clinic. Here, the patient is asked if this feeling has disappeared. After a sigh and a long pause, the patient starts responding, while gazing at the desk in front of him. The interpreter maintains her gaze at him throughout his turn. Notice that the interpreter produces a head nod only after the patient directs his gaze at her in line 10.

Extract 2

1 **INT** **вы сказали пришли сюда, вы были в шоке;**
you said that you came here, you were in shock

Int -----gaze to Pat----------->l.19

Pat gaze to Int----------->

2 **а сейчас остается такое ощущение или ушло состояние шока** –
and now does that feeling remain or has the state of shock disappeared

Pat ----------gaze away---------->l.10

3 **(0.2)**

4 **PAT** **((sighs))**

5 **(1.4)#**

fig #fig. 1

figure 1

6 **почти ушле= ушел (.) но**
It has almost disappear=disappeared (.) but

7 **(0.7)**

8 **я говорил, что**
I have told that

9 **(0.5)**

10 → **в наше время, я, когда я сплю, во сне +увижу+.**
at night, I, when I sleep, in my dream +I see+.
Pat ----------gaze to Int-->l.13
int **+double nod+**

11 **(0.5)#**

fig #fig. 2

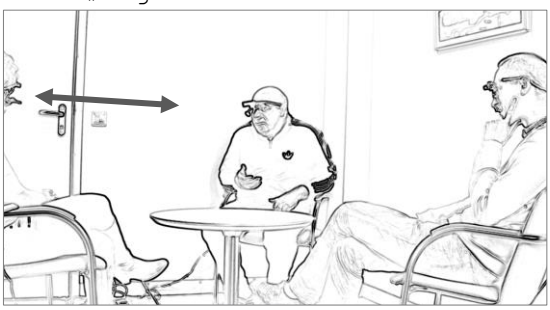


figure 2

12 **INT** **mm [hm**

13 **PAT** **[что у меня (.) со мной приходил,**
[*that what my* (.) *happened to me,*

Pat -----gaze away---------->l.19

14 **(0.6)**

15 **я не забываю это**
I will never forget that

16 **(0.5)**

17 **до конца моей жизни:= и то**
until the end of my life and that

18 **(0.8)**

19 **такой было что (.) не забуду.**
was like that (.) *I won't forget.*

The patient is telling that the initial state of shock has almost disappeared, but in lines 8-19 he repeats (*I have told*) that the bad memories keep coming back in his dreams. By repeating this and becoming more granular (cf. Stivers 2008) in his telling (*at night, when I sleep, in my dreams*), he conveys his stance towards these nightmares as something that still troubles him and that will not go away. It is only when the patient directs his gaze at the interpreter (line 10 'I see') that she immediately displays affiliation with head nods. Interestingly, the patient maintains his gaze at her and after a short pause (line 11), she aligns by producing a continuer ('mm hm', line 12). The continuer 'mm

hm' is used here to encourage the patient to continue speaking (see Drummond & Hopper 1993, Gardner 2001) and it shows how the interpreter delicately coordinates the interaction. Their mutual gaze ends in line 13 and during the remainder of the patient's turn, the interpreter does not offer any form of backchannel response. There are several occasions during the patient's turn that provide direct access to his stance, but the interpreter appears to withhold her response. She carefully monitors the patient's speech and visual behavior and waits for him to develop his stance further (Muntigl et al. 2012). In sum, these two examples reveal a direct relationship between the patient's gaze and the interpreter's backchannel responses and illustrate the subtlety of the interpreter's coordinative role in such an encounter.

6.4.2. Sustained mutual gaze and increased affiliative strength

The average duration of the patient's gaze towards the interpreter during his multi-unit turns was 4 seconds. In the following example, the patient retains his gaze at the interpreter for 14 seconds. This is strikingly long, as mutual gazes tend to be rather short (Kendon 1967: 28). As I will show, the patient's sustained gaze at the interpreter increases the recurrence (and intensity) of interpreter's nods.

Prior to this excerpt, the patient has told that he takes a lot of medication to regulate his blood pressure, as the traumatic experiences from his past keep coming back in his dreams. Here, the patient is telling that when his blood pressure is measured during the night, it is always high. He cannot forget the past. In contrast to extracts 1 and 2, the interpreter and the patient are engaged in mutual gaze throughout the whole excerpt.

Extract 3

1	**PAT**	**всегда, когда мне проверяют**
		always, when (my blood pressure) is checked
	Pat	--gaze to Int----------> l.13
	Int	gaze to Pat----------> l.13
2		**(0.2)**
3		**утром.**
		in the morning

4 (.)

5 **INT** → **+mm**
int +repeated shallow nods --->

6 **(0.7)**

7 **PAT** **нормальное+ давление.**
normal blood pressure
int ----------+

8 **PAT** **#а когда вече+ром.**
but when in the evening.
int → +slow repeated nods --->

fig #fig. 1

figure 1

9 (.)

10 **PAT** → **ночью проверяют, всегда есть да+вление+**
at night they check, it is always high
int ----------+ +repeated nods->

11 (.)

12 **INT** **(ja.)**
yeah
int ---->

13 **PAT** **Никогда не забуду.+**
I will never forget.
int --------------------+

The patient starts off by telling “when (blood pressure) is checked”, after which there is a short pause. Then he continues in line 3 by specifying when it is checked (‘in the morning’), which provides extra detail on the situation and projects the contrast with the information that is introduced in line 8 (‘but when in the evening’). By becoming more detailed in his report, the patient “heightens accessibility” (cf. Stivers 2008: 44) of his situation to his recipient. The interpreter responds with the minimal response ‘*Mm*’ (line 4), augmented with shallow nods that continue into line 7. According to Gardner

(2001: 31), '*Mm*' can be seen as a "a non-intrusive, reserved response to a delicate topic". The interpreter's nods continue during the long pause (line 6) well into line 8. By nodding, she is responding to the patient's sustained gaze at her and is displaying her affiliation with the patient's situation.

In lines 8-11, the patient introduces a contrast ('but when in the evening'). He self-repairs by adding that 'at night', when his blood pressure is checked, it is always high. All the while the patient and the interpreter are engaged in mutual gaze. In lines 8-11, the interpreter briefly closes her eyes and starts producing a series of slow nods. Closing of the eyes while nodding adds to the affiliative character of her response (see also Kendon 1967). She appears to be closely affiliating with the patient's story. In line 10, at the point when the interpreter has been given access to the situation through the patient's detailed description, the interpreter starts nodding again. At the possible completion of the patient's turn, the interpreter produces the acknowledgment token 'yeah', augmented with head nods, which display her heightened level of understanding of the preceding telling and willingness to take the turn (Gardner 2001, Drummond & Hopper 1993). She continues nodding as the patient adds in line 13 "I will never forget". Thus, it appears that the patient's sustained gaze at the interpreter increases the strength of the interpreter's affiliation with the patient's telling.

In sum, this section has shown that, although the interpreter seems to adopt a reserved attitude, she still has an important role as a recipient of the patient's talk. Mutual gaze with the patient draws her to produce affiliative head nods in response to his telling. Also, the duration of her head nods seems to be linked to the sustained mutual gaze with the patient and displays her heightened level of affiliation.

6.5. Therapist's dual feedback

We now turn to the therapist's production of backchannel responses during the interpreter's rendition of the patient's preceding turn. In this session, we find many examples of the pattern previously discussed as 'dual feedback' (Chapter 5). In the

following example, which is the continuation of extract 1, the therapist listens to the interpreter who starts with the rendition of the patient's talk into Dutch. Note that the figures are from the eye-tracking camera worn by the therapist and represent his perspective.

Extract 4

1	**INT**	→	**#het was heel s+lecht.#**
			it was very bad.
	Ther		gaze to Int----------gaze to Pat->
	ther		**+repeated nods --->**

#fig. 1 #fig. 2

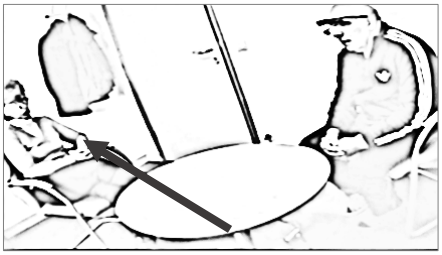


figure 1 (THER perspective)

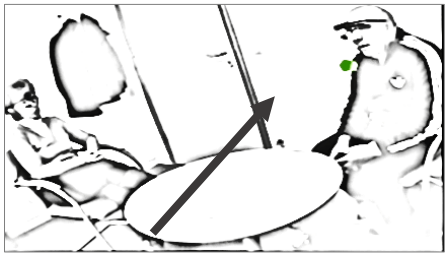


figure 2

2	**INT**		**eu:h+ en ik bedoel eu:h**
			euh and I mean euh
	Ther		------------------gaze at Int->
	ther		**----+**
3	**INT**	→	**dat mijn situ+atie heel slecht was.**
			that my situation was very bad.
	Ther		------------------------gaze at Pat-->
	ther		**+double nod -------- >>**

In line 1, the interpreter produces the assessment that 'it was very bad'. The therapist immediately reacts to this expression of stance by nodding and shifting his gaze to the patient. He then moves his gaze back to interpreter who continues with her rendition, thereby displaying his continuing attention. In line 3, the interpreter explicates that 'my situation was bad' to which the therapist immediately reacts with a nodding accompanied by a gaze shift towards the patient. Thus, by shifting his gaze to the patient while nodding, the therapist displays understanding and endorsement of the expressed stance while at the same time indicating the 'principal' or person who he is affiliating with. Finally, in contrast to the interpreter, the therapist's head nods are synchronized

with the renderings of the patient's stance in line 1 ('it was very bad') and line 3 ('my situation was very bad') (see also Muntigl et al. 2012).

Another example is provided in extract 5. In this example, the therapist's dual feedback is an immediate reaction to the patient's disagreement. Prior to this excerpt, the therapist had asked the patient if he has learned to cope with his nightmares during his stay at the clinic, since nightmares are very difficult to control and to reduce. In the following extract, we provide the interpreter's rendition of the patient's turn (not represented in the transcript).

Extract 5

1 **INT** → **dat kl+opt.**
that's true.
Ther gaze at Int-->
ther **+large repeated nods --->**

#fig. 1

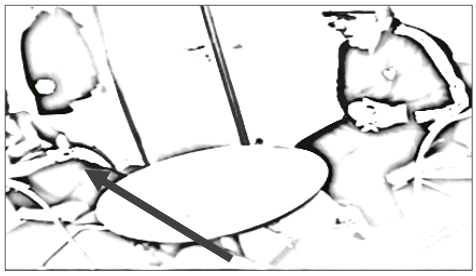


figure 1

2 **INT** **eu:h# daar kan ik ook niks mee:**
euh I cannot do anything with it
Ther→ gaze to Pat------------------>
Pat --gaze to Ther---------------->
ther **----------------------------> 1.4**

#fig. 2

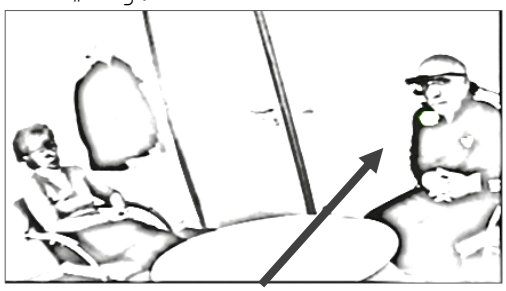


figure 2

3 **(0.2)**+
Ther ---->

```
      Pat       ---gaze away-->
      ther      -----+
4     INT   →   en daarom [wil ik] nu eigenlijk liever
                and that's why I would now actually rather
5     PAT               [(да)]
                         yeah
      Ther      ----------------------gaze to Int------>
6                 over iets anders heb[ben (.)]
                  talk about something else
7     THER                          [ +ja +  ]
                                      yeah
      Ther      --------------------gaze to Pat->
      ther                          +nod+
8     INT       ik wil daar niet meer over heb[ben.
                I don't want to talk about it anymore.
9     THER                                     [+nee precies+]
                                                 no exactly
      Ther      gaze to Int-------------------gaze to Pat--->>
      ther                                     +   nod     +
10    THER      nee precies, maar ik (0.2) het lijkt mij ook niet
                verstandig en ook niet nodig
                no exactly, but I  (0.2) it does not seem sensible nor
                necessary to me either
```

In line 1, the therapist's question is ignored, but the interpreter renders the patient's agreement ('that's true') with the therapist's suggestion that nightmares are difficult to control. The therapist starts a series of nods during this agreeing move of the patient (lines 1-3), while shifting his gaze to the patient (line 2). The therapist and the patient briefly establish mutual contact during the therapist's head nods (figure 2), which appear to reinforce the established positive alliance between them (cf Muntigl et al. 2012:14). However, in lines 4-10 a disagreeing move is introduced ('I would rather talk about something else, I don't want to talk about it anymore'), that puts the affiliation between the therapist and the patient under pressure. In response to this rejection of the proposed agenda, the therapist immediately produces a head nod augmented with the acknowledgment token 'yeah' (line 9), while swiftly shifting his gaze towards the patient. The immediacy of his response indicates no trouble with agreeing with the evaluative position expressed in the interpreter's rendition (cf. Stivers et al. 2011). The

therapist is thus maximally affiliating by displaying direct endorsement of the patient's stance. As noted by Muntigl et al., the therapist is seeking to avoid disagreements, as these "may place stress on the therapist-client relationship" and may therefore hinder important therapeutic work (2014: 332). In line 11, the therapist increases his endorsement of the patient's stance by shifting his gaze towards the patient and providing the assessment 'no exactly' in combination with a nod. With his head nods, the therapist is actively working towards maintaining affiliation and alignment with the client "even in the face of outright disagreement" (Muntigl et al. 2012: 10).

6.6. Conclusion

This chapter has presented an analysis of the interpreter's and the therapist's backchannel responses in a therapeutic session. The differences in their overall production of backchannel responses may reflect differences in the interpreter's and the therapist's roles and involvement in the therapeutic process. Although the interpreter in this session appears to be the addressee of the patient's talk, she produces strikingly less backchannel responses than the therapist. This appears in line with her position within the exchange and the notion of interpreter's 'neutrality'. However, through a careful placement of her backchannel responses during the patient's production of multi-unit turns, the interpreter has been found to discreetly coordinate the talk (see Gavioli 2012) and affiliate with the patient.

The main part of the analysis focused on the role of gaze in the production of interpreter's and therapist's backchannel responses. Although both the interpreter and the therapist provide backchannel responses to the patient's telling, they do so in slightly different ways. The *interpreter's backchannel responses* are strongly linked to the patient's gaze, that functioned as a response inviting cue. Moreover, mutual gaze with the patient seems to intensify the interpreter's backchannel responses and her display of affiliation (see also Kendon 1967). The interpreter is thus not present in a detached way, but appears to engage in a co-operative relationship with the patient. According to

Merlini & Favaron, strict adherence to an impersonal style of interpreting "is not always the best way to serve one's clients, especially when their intention is to engage in a friendly and co-operative dialogue" (2005: 132). The display of cooperativeness may be especially relevant in the context of therapeutic talks with traumatized asylum seekers. However, the danger of being the responder is that the interpreter may end up doing therapeutic work, by reacting to and eliciting more elaborate responses from the patient. This means that interpreters in psychotherapeutic settings are faced with the challenge of constantly maintaining the delicate balance between professional neutrality and cooperativeness. This study has shown that head nods and visual orientation are subtle ways to display affiliation and patient orientation. As a responder, the interpreter maintains and forges the current social relationship at a level of intensity or intimacy that is related to the goals and purposes of the given interactional context (cf. Gavioli 2012, Merlini & Favaron 2005).

The therapist, on the other hand, is dependent on the interpreter's renditions and involvement in the interaction. His listener responses seem to perform a less coordinative function as the interpreter's do. The therapist is reacting to a product of the interpreter's cognitive effort, whereas the interpreter is responding to the patient's talk-in-production, with all its pauses, hesitations and self-repairs. This becomes clear from the differences in the positioning and structural characteristics of their backchannel responses. As for the therapist's head nods, the findings are consistent with what was reported by Muntigl et al. (2012, 2014). His head nods are target-specific: they are sequentially positioned and contiguous to the teller's expressions of stance (cf. Muntigl et al. 2014). Finally, the study has demonstrated how the therapist's gaze shifts as part of *dual feedback* are a manifestation of a "triadic affective interaction" (Baraldi & Gavioli 2008).

There is obviously much more to be learned about the interactional dynamics of interpreter-mediated therapeutic encounters. Due to its limited scope, this study is of an exploratory nature. Indeed, more data need to be collected in order to gain a broader picture of the interpreter's listener role in therapeutic sessions and its impact on the therapeutic relationship. This chapter has illustrated the value of a multimodal approach to the analysis of backchannel responses in the context of mental healthcare interpreting.

It has shown how a multimodal analysis may help us gain a better understanding of the intersubjective relations among the interlocutors and the interpreter's social role within such exchanges. Moreover, the observations in this study provide further argument for the importance of including gaze not only in research on interpreter-mediated interaction but also in interpreter training.

Conclusion

7.1. Summary of the findings

This dissertation set out to explore the role of gaze direction in interpreter-mediated interaction. Whereas the role of gaze in face-to-face (monolingual) interactions has been the object of study for decades, it is only recently starting to attract interest in the investigation of interpreter-mediated interactions. Recent studies have illustrated the importance of gaze direction for the display of participation and for the management of ongoing interpreter-mediated talk. Nevertheless, there was a lack of systematic research on the role of gaze direction in relation to other (visual and verbal) resources in this type of conversational settings. The aim of this dissertation was to make a first contribution in this direction by focusing on the role of gaze direction as part of two specific communicative practices: the management of *turn-taking* (Chapter 4) and the production of *backchannel responses* (Chapters 5-6). In order to get detailed information on the gaze behavior of all participants, I made use of an innovative approach, viz. mobile eye-tracking, that allows highly detailed analysis of participants' gaze during ongoing interaction.

As a starting point for the analyses on those two phenomena, I first mapped all participants' gaze distributions in the recorded sessions (**Chapter 3**). The results presented some striking differences in the distribution of gaze orientation of the primary participants that are linked to their differences in position and professional status: whereas the counsellors (as more 'powerful' participants within the exchange) gaze more at the student while speaking, the students directed their gaze more at the interpreter. Furthermore, I have shown how the differences between 'speaker' and 'hearer' gaze behavior, as put forward by Kendon (1967) and Goodwin (1981), appear

less clear-cut in triadic interactions in which one of the participants is the interpreter. Thus, the findings in this chapter refine some of the previous claims with reference to the organization of gaze in both interpreter-mediated and same-language interactions.

Chapter 4 presented a CA-inspired study on the role of gaze in the management of turn-taking. It demonstrates how the interpreter can contribute to a specific activity ('chunking' of a multi-unit turn) through her gaze. The interpreter's immediate gaze shift towards the end of her ongoing TCU during 'chunking' appears to signal two things: (1) that the end of her TCU is not a transition relevance place (TRP) for the addressee, and (2) that the gazed-at participant is selected to continue with the multi-unit turn. The study shows how gaze direction creates certain expectations with respect to the development or progression of the ongoing action that are not expressed on a verbal level. As for the interpreter's role within such face-to-face encounters, these findings corroborate the view of the interpreter as a *co-participant* in the conversation, who not only co-manages the interaction on a verbal level (through her provision of consecutive renditions), but also visually, though her gaze orientation.

The main focus of the study was on the role of gaze in the production of backchannel responses (such as *mm hm*, *yeah*, head nods) in interpreter-mediated exchanges, presented in **Chapter 5**. Although interpreter-mediated interactions create a particularly interesting (and complex) setting for the study of backchannel responses, these have received surprisingly little attention in the literature so far. It was assumed that interpreters and primary participants generally produce little backchannel responses in such exchanges, due to their asymmetric access to each other's language and the interpreter's distinct participatory role within the exchange. The aim of this study was (1) to map the *overall production* of backchannel responses in the current dataset, (2) to explore the role of *mutual gaze* in the production of listeners' backchannel response (cf. Bavelas et al. 2002) and (3) to examine the sequential organization and the directionality of backchannel responses in the dataset. Results have revealed differences in the interpreters' and primary participants' production of backchannel responses, and shown that backchannel responses are for the most part realized visually (though head nods, head shakes etc.). Furthermore, inspired by Bavelas' et al. (2002) widely-cited study, I explored the role of mutual gaze as a backchannel eliciting cue in the current dataset. In

contrast to Bavelas et al. (2002), I made a distinction between verbal and visual backchannel responses. The results indicate a strong correlation between the occurrence of mutual gaze and interpreters' production of *visual* backchannel responses (e.g. head nods). Overall, these findings show that further research is needed to refine some of the previous claims made on the regulatory functions of gaze in face-to-face interaction.

The main part of the analysis was devoted to a recurrent pattern in the production of backchannel responses in the dataset, that I refer to as *dual feedback*, and how it affects the participation framework of the exchange. Interpreter-mediated interactions are usually conceived as two overlapping participation frameworks in which two separate sets of common ground are constructed, from which one participant is excluded. However, the management of the participation framework in interpreter-mediated interaction is much more complex than previously assumed. The analysis has shown how primary participants, who have no access to each other's language, can (and do) achieve reciprocity of understanding through gaze and backchannel responses. Moreover, it has demonstrated how recipients in interpreter-mediated interactions display sensitivity to the differences in participation status and knowledge states of the interpreter and the other primary participant (or the 'principal'). This was most evident during the production of newsmarkers and affiliative backchannel responses, that are combined with gaze shifts towards the 'principal'. The analysis also suggests that such backchannel responses co-occurring with gaze shifts form 'composite signals' during the interaction. In contrast to previous claims that "video data show more, but not necessarily novel features of interactions, at least in terms of feedback signals" (Gavioli 2012: 205), the current analysis has shed light on new ways how temporarily excluded participant is actually involved in the exchange, even without understanding of the language. Overall, gaze is more than an additional layer in the analysis, but is closely intertwined with the production of backchannel responses in face-to-face interaction.

Chapter 6 presented a final closing case that has demonstrated how some of the phenomena presented in the previous chapter come together 'in the wild', more specifically, in the context of an interpreted therapeutic encounter that was also recorded with mobile eye-tracking glasses. It has shown how interpreter's backchannel responses and gaze play an important role in the emerging story of the patient and how therapist's

'dual feedback' contributes to a "triadic affective interaction" (Baraldi & Gavioli 2007) in this context.

7.2. Implications for research and practice

This dissertation contributes to research on gaze in interpreter-mediated interactions in the following four ways. First, it shows how a *multimodal approach* to the study of interpreter-mediated interactions is necessary in order to gain a thorough understanding of the interactional dynamics of such exchanges (see also Davitti 2012). Gaze does not function as an isolated communicative resource, but it contributes to situated practices in the interpreted event. A multimodal approach, in turn, implies having access to video data that allow for such fine-grained analyses of gaze and other modalities of all participants in the ongoing interaction.

Second, the current study contributes to our understanding of the interactional dynamics of interpreter-mediated interactions, by highlighting the interpreter's ***visibility*** within the exchange. Interpreters are more than simple 'voices' (Merlini & Favaron 2005) who form a direct communicative link between the primary participants; they are visible actors who contribute to the interaction through their physical presence and embodied behavior. As illustrated throughout this dissertation, interpreter's visual (gaze) conduct is an important aspect of their mediating role within such exchanges. Therefore, "it is time that the interpreters stop feeling guilty about their visibility" (see Bot & Verrept 2013: 122). Raising awareness about interpreter's visibility and the impact of the interpreter's visible behavior on the interaction will not only lead to new questions, but it will also contribute to interpreter training.

Third, this work also has addressed the primary participants' involvement in the interaction, which has received less attention in the literature. Previous studies have stressed the importance of the interpreter in the management of the interaction, and in the establishment of mutual understanding and reciprocity between the primary participants. Nevertheless, although they depend on the interpreter for the establishment of mutual understanding and for the coordination of turn-taking, the interpersonal

management on the level of the *triad* - as shown for example in the analysis of 'dual feedback' - is to a large extent in the hands of the primary participants. This understanding relieves some of the burden from the interpreter and stresses the responsibility of all participants for the accomplishment of the interaction as *joint activity*. Interpreter-mediated interaction is then, in the true sense of the word, a 'communicative *pas de trois*' (Wadensjö 1998).

Fourth, this dissertation introduces mobile eye-tracking as a novel approach to collecting detailed information on gaze in interpreter-mediated interactions. Moreover, it demonstrates how mobile eye-tracking is not confined to experimental settings, but that it can also be used for the study of natural interpreter-mediated interactions.

7.3. Limitations and suggestions for future research

This work has illustrated the importance of gaze in the production of backchannel responses and in turn-taking in interpreter-mediated interactions. Studying these phenomena in depth has opened up a number of further questions that could not be addressed in this dissertation. In particular the management of turn-taking in interpreter-mediated interaction deserves further investigation. One of the questions that emerged from the study of turn-taking is the issue of projection. Projection of points of possible completion plays a central role in the organization of turn-taking in conversation (Sacks et al. 1974). Participants in conversation must project the end of the current speaker's turn in order to prepare their turns in advance and avoid or minimize overlap. This issue appears even more pressing for a consecutive interpreter. Although interpreters are often not selected as next speaker visually (through gaze), they are nonetheless expected to take the turn at some point. In my data, I found some evidence of interpreter's anticipatory gaze shifts before taking turn-at-talk. It would be interesting to examine the systematicity of such behavior, which could reveal how interpreters prepare the interactive future and what (visual/verbal) information they rely on to project turn completions.

Since this dissertation has focused on the role of gaze in one specific interactional setting, the question arises to what extent these findings are generalizable to other settings and contexts. This issue was partly addressed in the case study presented in Chapter 6. Nevertheless, the following questions remain: how does the use of gaze differs according to the conversational setting, the seating arrangement of the participants and the physical presence of the interpreter? The findings presented in this dissertation can be used for further explorations with larger datasets, or for experimental research, that can isolate some of the variables identified in this work and examine them more systematically.

Finally, this study has helped to provide an answer to some basic research questions concerning the multimodal organization of face-to-face interpreter-mediated interactions and the role of gaze in that process. The presented findings gain additional relevance "in the light of the widespread use of digital communication technologies, which have brought about new modes of interpreting (e.g. video-mediated interpreting)" (Davitti 2015). As described by Pöchhacker (2018), the development of new technologies, and most notably the high-quality streaming of audio and video data, has paved the way for novel forms of distance or remote interpreting, in which the interpreter can provide renditions from a remote location (Braun & Taylor 2011). Although remote interpreting is seen as a (cost-) efficient way of providing interpreting services, its implementation is still subject of debate among both interpreting scholars and practitioners (Licoppe & Veyrier 2017). Furthermore, it raises a number of important questions: How does the physical separation of the interpreter affect the turn-taking and feedback behavior? Which (semiotic/multimodal) resources do the interpreters and users of interpreting services resort to in *remote* interpreting settings? Therefore, the insights gained from systematic, multimodal studies on interpreting in face-to-face settings will provide a baseline against which such novel forms of interpreting can be assessed.

Appendix

A. Transcription conventions

Data are represented in two-line transcripts. The first line represents the original utterance in Dutch or Russian; the second line gives an English translation, that captures as closely as possible, the meaning of the original utterance.

Temporal and sequential aspects

[	A left bracket indicates the onset of overlapping speech
]	A right bracket indicated the end of overlapping speech
=	An equals sign indicates contiguous speech
(.)	A dot in parenthesis indicates a "micropause" (shorter than 0.2 seconds)
(0.5)	Silences are indicated as pauses in tenths of a second

Aspects of speech delivery

:::	Colons indicate lengthening of the immediately prior sound
hh	The letter 'h' indicates audible aspirations (the number of h's the represented the length of the outbreath)
.hh	A period preceding the letter 'h' indicates audible inhalations (the number of h's the represented the length of the inbreath)
.	A period indicates a falling intonation contour
,	A comma indicates rising intonation contour
?	A question mark indicates a rise stronger than the comma
-	A hyphen indicates flat intonation contour

Other aspects

(maybe)	words within single parentheses indicate likely hearing of that word
((coughs))	information in double parentheses provides additional details

Multimodal aspects

Conventions for the multimodal transcription are inspired by the conventions developed by Lorenza Mondada: https://franzoesistik.philhist.unibas.ch/fileadmin/user_upload/franzoesistik/mondada_multimodal_conventions.pdf

+	symbol + specifies the exact moment at which the **counsellor's** gestures begin and end
*	symbol * specifies the exact moment at which the **student's** gestures begin and end
#	symbol # specifies the exact moment at which the image (fig.) refers. This is done by inserting the symbol both on the line of the talk and on the line dedicated to the image (fig. in the margins)
*-->	the action described continues across subsequent lines
-->*	until the same symbol is reached.
gaze at A-->	gaze direction is transcribed parallel to speech. Gaze at A continues until the following indication (e.g. gaze away)
-->>	The action described continues after the extract's end.

B. Specifications on the technical equipment

The research reported in this dissertation was conducted with three types of mobile eye-trackers, according to their availability at the time of the planned recordings. The new generation Tobii eye-trackers were acquired only in the later stage of the data collection and were thus used for the recordings of the third dataset.

Mobile eye-trackers

- **Arrington Gig-E603** eye-tracking glasses
  - 30 frames per second (fps)
  - 320x240 pixels

- **Pupil Pro Binocular Glasses**
  - 30 frames per second (fps)
  - 1920x1080 pixels

- **Tobii**
  - 25 frames per second
  - 1920x1080 pixels

Additional camera

- Sony HDR-FX1000E: 25 frames per second, 720x576 pixels

C. General information about the interpreters

Interpreter 1

- **Years of experience:** 9 years
- **Interpreter training (yes/no)?** yes
- **In which settings usually employed as interpreter?** all social sectors (medical, psychosocial, legal, business, etc.)

Interpreter 2

- **Years of experience:** > 20 years
- **Interpreter training (yes/no)?** yes
- **In which settings usually employed as interpreter?** all social sectors (medical, social, legal, business, etc.)

Interpreter 3

- **Years of experience**: > 10 years of experience
- **Interpreter training (yes/no)?** yes
- **In which settings usually employed as interpreter?** legal

Interpreter 4 (Chapter 6)

- **Years of experience:** > 20 years
- **Interpreter training (yes/no)?** yes
- **In which settings usually employed as interpreter?** psychotherapy (and other?)

D. Coding scheme for backchannels

Based on Gardner (2001)

TYPE	CODE	DESCRIPTION	PROTOTYPICAL EXAMPLES
Continuer	**CON**	hands floor back immediately to the previous speaker	***mm hm, uh huh***
Acknowledgment	**ACK**	claim agreement or understanding of the prior turn, typically with a falling intonation contour. Greater speakership incipiency than 'mm hm' or 'uh huh'. They can, though, also be treated by participants as continuers, and when they are, they carry a rising terminal intonation contour.	***mm, yeah***
Newsmarker & newsmarker-like objects	**NEW**	mark the prior speaker's turn as newsworthy in some way, that its producer has undergone a change of state. Expressing 'newness' or 'surprise' something contrary to the expectations	***Really?*** *;* *Change-of-state token* ***Oh!*** *;* *the 'idea-connector'* ***Right*** *(=recognizes the connection between two or more ideas)*
Change-of-activity tokens	**ACT**	mark a transition to a new activity or a new topic in the talk (**forward-looking)**	***Okay, alright***
Assessments	**ASS**	evaluate the talk of the previous speakers	***Good, Great, How intriguing, what a load of rubbish***
Brief questions or other types of repair	**QUE/REP**	seek to clarify misunderstandings or mishearings	***Who? Which book do you mean? Huh?***
Collaborative completions	**COLL**	one speaker finishes a prior speaker's utterance	
Non-verbal vocalizations & kinesic actions	**NON**		***laughter, nods, head shakes***

E. Distribution of participants' gaze orientations in the therapeutic session (Chapter 6)

1. GAZE WHILE SPEAKING

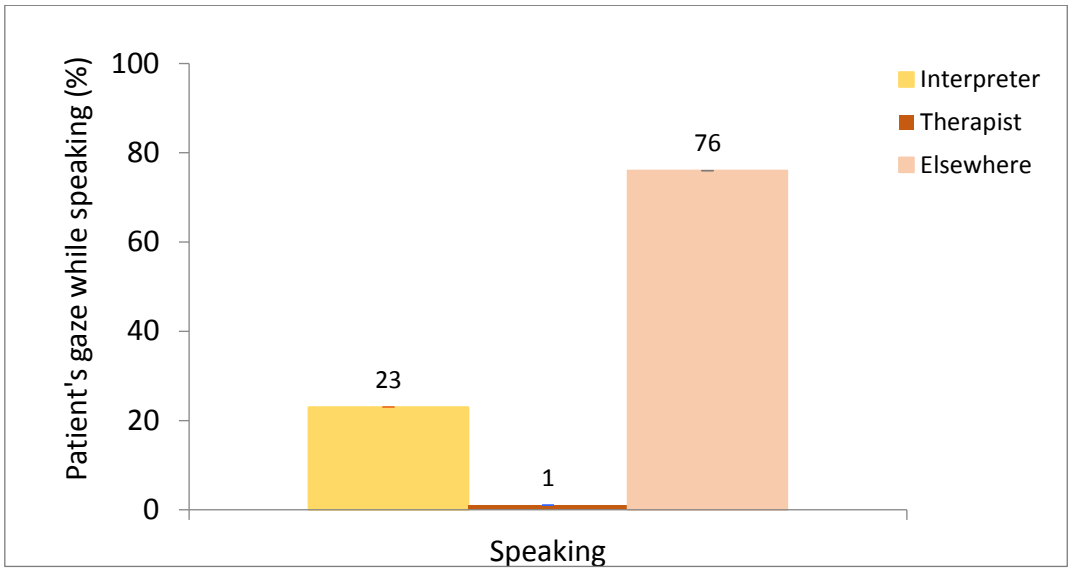


Figure 6.1. Patient's gaze direction while speaking.

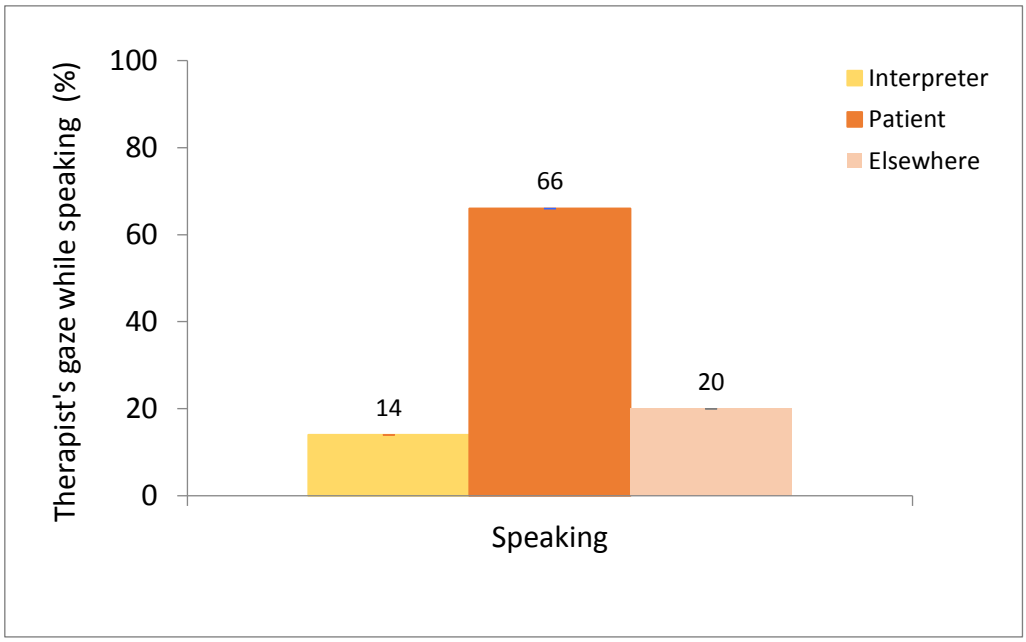


Figure 6.2. Therapist's gaze direction while speaking.

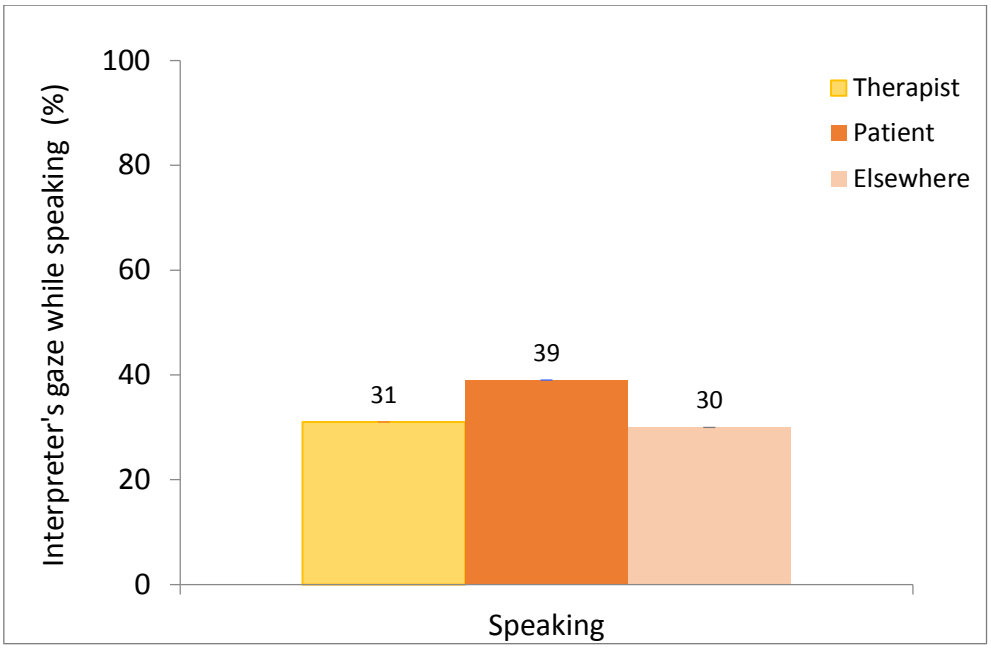


Figure 6.3. Interpreter's gaze while speaking.

2. GAZE WHILE LISTENING

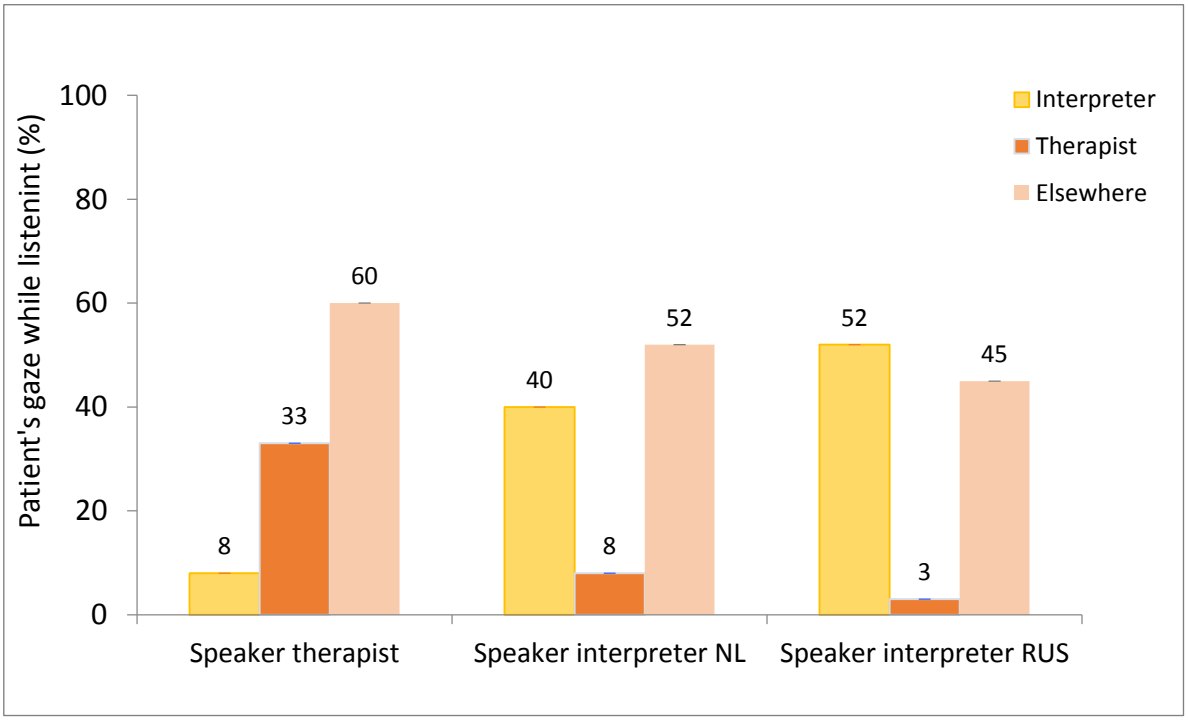


Figure 6.4. Patient's gaze while listening.

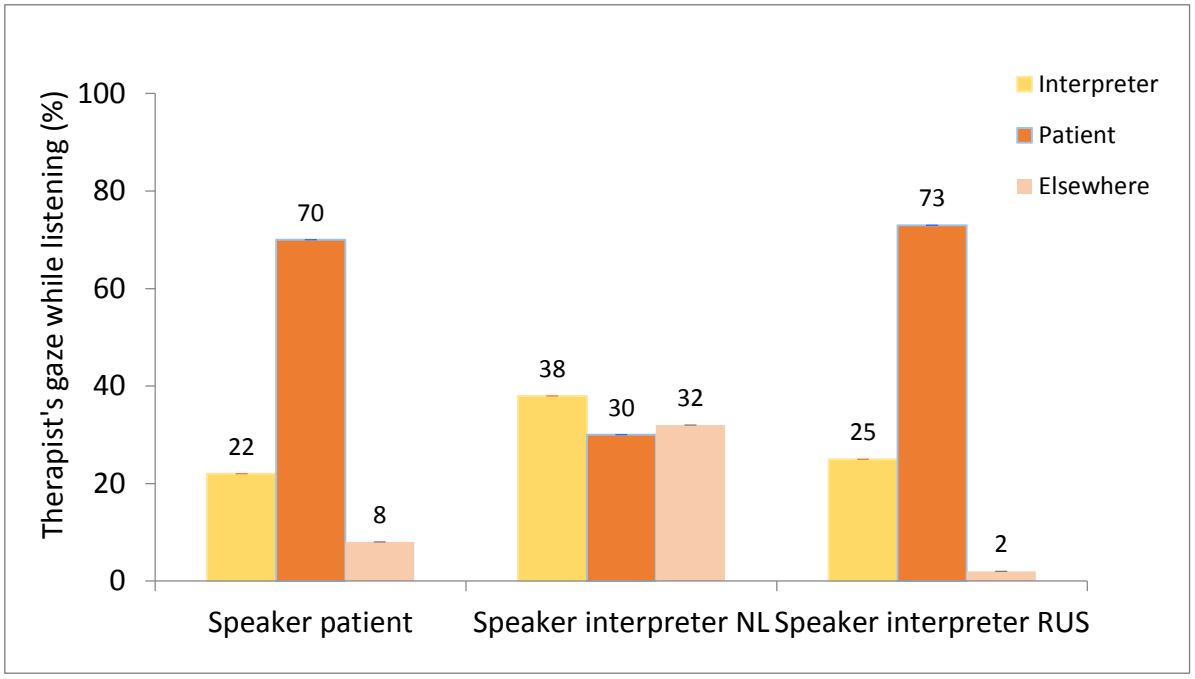


Figure 6.5. Therapist's gaze while listening.

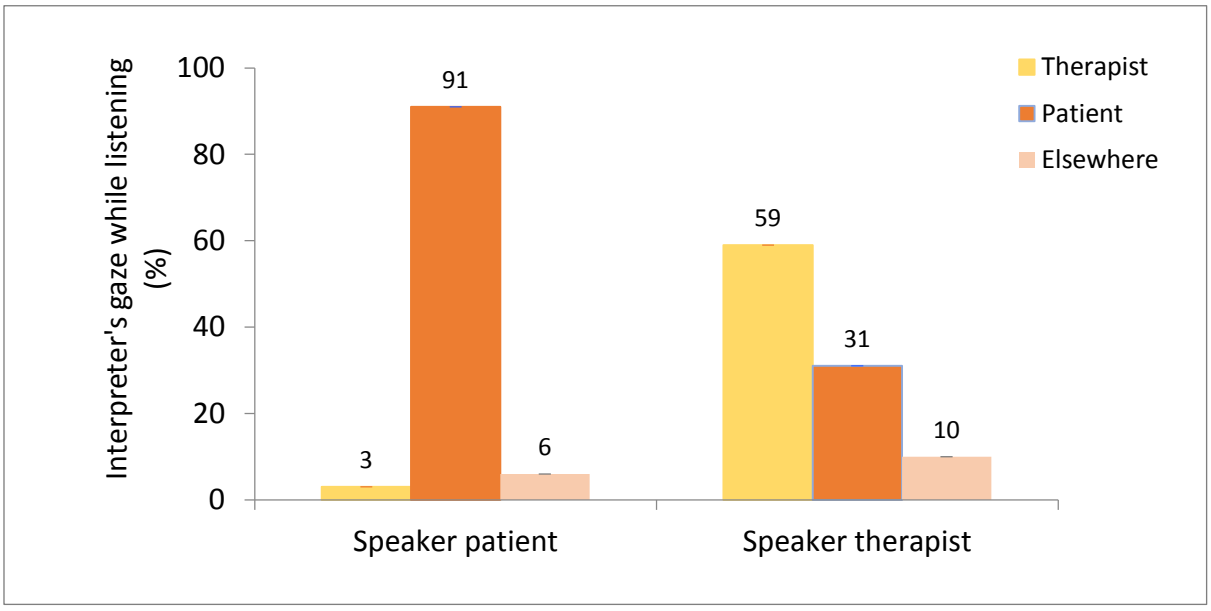


Figure 6.6. Interpreter's gaze while listening.

References

Allwood, J. & Cerrato, L. (2003). A study of gestural feedback expressions. In Paggio, P., et al. (Eds.), *Proceedings of the First Nordic Symposium on Multimodal Communication*, pp. 7-22.

Angelelli, C. (2000). Interpretation as a Communicative Event: A Look through Hymes' Lenses. *Translators' Journal*, 45 (4): pp. 580-592.

Angelelli, C. V. (2004). *Revisiting the Interpreter's Role: A Study of Conference, Court, and Medical Interpreters in Canada, Mexico, and the United States.* Amsterdam: John Benjamins.

Argyle, M. & Cook, M. (1976). *Gaze and Mutual Gaze*. Cambridge: Cambridge University Press.

Argyle, M. & Cook, M. (1994). Gaze and Mutual Gaze. *British Journal of Psychiatry*, 165, pp. 848-850.

Auer, P. (2005). Projection in interaction and projection in grammar. *Text*, 25 (1): pp. 7-36.

Auer, P. (2017). Gaze, addressee selection and turn-taking in three-party interaction. *Eye-tracking in interaction. Studies on the role of eye gaze in dialogue.* John Benjamins.

Baraldi, C. & Gavioli, L. (2007). *Dialogue Interpreting as Intercultural Mediation An analysis in healthcare multicultural settings***.** In Grein, M. & Weigand, E. J. (Eds), *Dialogue and Culture.* Amsterdam & Philadelphia: Benjamins Publishing Company, pp. 155-175.

Baraldi, C. & Gavioli, L. (2012). Understanding coordination in interpreter-mediate interaction. In Baraldi, C. & Gavioli, L. (Eds.), *Coordinating Participation in Dialogue Interpreting*, pp. 1-21.

Branes, S.E. (2011). Claiming mutual stance: On the use of that's right by a person with

aphasia. *Research on Language and Social Interaction*, 44 (4): pp. 359-3 84.

Bavelas, J., Coates, L. & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, 79 (6): pp. 941-952.

Bavelas, J., Coates, L. & Johnson, T. (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication*, pp. 566-580.

Beattie, G. W. (1978). Floor apportionment and gaze in conversational dyads. *British Journal of Social and Clinical Psychology*, 17: pp. 7-16.

Beattie, G. W. (1979). Planning units in spontaneous speech: Some evidence from hesitations in speech and speaker gaze direction in conversation. *Linguistics*, 17: pp. 61-78.

Bertrand R., Ferré, G., Blache, P., Espesser, R. & Rauzy, S. (2007). Backchannels revisited from a multimodal perspective. *Proceedings of Auditory-Visual Speech Processing 2007 (AVSP2007) 2007*, Kasteel Groenendaal, Hilvarenbeek, The Netherlands.

Betz, E. & Golato, A. (2008). Remembering relevant information and withholding relevant next actions: The German token *achja*. *Research on Language and Social Interaction*, 41 (1): pp. 58-98.

Bolden, G B. (2000). Toward understanding practices of medical interpreting: Interpreters' involvement in history taking. *Discourse Studies*, 2 (4): pp. 387-419.

Bolden, G.B. (2013). Unpacking 'self': Repair and epistemics in conversation. *Social Psychology Quarterly*, 76: pp. 314-342.

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

Bot, H. & Verrept, H. (2013). Role issues in the Low Countries. Interpreting in mental healthcare in the Netherlands and Belgium. In Schäffner, C. & Kredens, K. & Fowler, Y. *Interpreting in a Changing Landscape: Selected papers from Critical Link 6*, pp.117-131.

Brône, G. & Oben, B. (2015). InSight Interaction. A multimodal and multifocal dialogue corpus. *Language Resources and Evaluation*, 49: pp. 195-214.

Brône, G., Oben, B., Jehoul, A., Vranjes, J. & Feyaerts, K. (2017). Eye gaze and viewpoint in multimodal interaction management. *Cognitive Linguistics*, 28 (3): pp. 449-484.

Ciolek, T.M. & Kendon, A. (1980). Environment and the spatial arrangement of conversational encounters. *Sociological Inquiry*, 50: pp. 237-271.

Clancy, P.M., Thompson, S.A., Suzuki, R. & Tao, H. (1996). The conversational use of reactive tokens in English, Japanese, and Mandarin, *Journal of Pragmatics*, 26 (3): pp. 355-387.

Clark, H. & Brennan, S. (1991). Grounding in communication. In Resnick, L.B., Levine, J.M. & Teasley, S.D. (Eds.), *Perspectives on socially shared cognition.* Washington: APA Books.

Clark, H. & Schaefer, E. F., (1989). Contributing to discourse. *Cognitive Science*, 13: pp. 259-294.

Clark, H. (1996). *Using Language*. Cambridge University press: Cambridge.

Clayman, S. E. (2013). Turn-Constructional Units and the Transition-Relevance-Place. In Sidnell, J. & Stivers, T. (Eds.), *The Handbook of Conversation Analysis.* Wiley-Blackwell, pp. 150-166.

Conklin, K., Pellicer-Sanchez, A. & Carrol, G. (2018). *Eye-tracking*. A guide for Applied Linguistics Research. Cambridge: Cambridge University Press. Couper-Kuhlen, E. & Tsuyoshi, O. (2007). 'Incrementing' in conversation. A comparison of practices in English, German and Japanese. *Pragmatics*, 17 (4): pp. 513-552.

Davidson, B. (2002). A model for the construction of conversational common ground in interpreted discourse. *Journal of Pragmatics*, 34: pp. 1273-1300.

Davitti, E., & Pasquandrea, S. (2017). Embodied participation: What multimodal analysis can tell us about interpreter-mediated encounters in pedagogical settings. *Journal of Pragmatics*, 107: pp. 105-128.

Davitti, E. (2012). *Dialogue Interpreting as Intercultural Mediation Integrating Talk and Gaze in the Analysis of Mediated Parent-Teacher Meetings*. PhD Thesis, The University of Manchester.

Davitti, E. (2013). Dialogue Interpreting as Intercultural Mediation: Interpreter's use of

upgrading moves in parent-teacher meetings. *Interpreting,* 15 (2): pp. 168-199.

Davitti, E. (2015). Gaze. In Pöchhacker, F., Grbic, N., Mead, P. & Setton, R. (Eds.), *Routledge Encyclopedia of Interpreting Studies.* Oxford & New York: Routledge. Pp. 168-169.

Davitti, E. (2018). Methodological explorations of interpreter-mediated interaction: novel insights from multimodal analysis. SAGE Publications *Qualitative Research. Special Issue: Multimodality: Methodological Explorations*.

De Beugher, S. (2016). *Computer vision techniques for automatic analysis of mobile eye-tracking data*. Unpublished PhD dissertation, KU Leuven.

De Pedro, R., Perez, I. & Wilson, C. (Eds.) (2009). *Interpreting and Translating in Public Service Settings: Policy, Practice, Pedagogy*. Manchester: St. Jerome Publishing. 7: pp. 198.

Deppermann, A. & Schmitt, R. (2007). Koordination: Zur Begründung eines neuen Forschungsgegenstandes (Coordination: Foundations of a new field of studies). In Schmitt, R. (Ed.), *Koordination. Analysen zur multimodalen Interaktion (Coordination. Analyses of multimodal interaction)*. Tübingen: Narr Verlag, pp.15-54.

Deppermann, A. (2013). Turn-design at turn-beginnings: Multimodal resources to deal with tasks of turn-construction in German. *Journal of Pragmatics*, 46 (1): pp. 91-121.

Deppermann, A. (2015). When recipient design fails: Egocentric turn-design of instructions in driving school lessons leasing to breakdowns of intersubjectivity. *Gesprächsforschung - Online Zeitschrift zur verbalen Interaktion*, 16: pp. 63-101.

Dittmann, A. & Llewellyn, L. (1968). Relationship between vocalizations and head nods as listener responses. *Journal of Personality and Social Psychology*, 9 (2): pp. 79-84.

Drew, P. & Heritage, J. (1992). *Talk at Work: Interaction in Institutional Settings*. Cambridge: Cambridge University Press.

Drummond, K. & Hopper, R. (1993). Backchannels revisited: acknowledgment tokens

and speakership incipiency. *Research on Language and Social Interaction,* 26 (2): pp. 157-177.

Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology,* 23 (2): pp. 283-292.

Duncan, S. (1974). On the structure of speaker-auditor interaction during speaking turns. *Language in Society,* 2: pp. 161-180.

Enfield, N.J. (2006). Social consequences of common ground. In Enfield, N.J. & Levinson, S.C. (Eds.), *Roots of human sociality: Culture, cognition, and interaction*. London: Berg, pp. 399-430

Enfield, N.J. (2009). *The anatomy of meaning: Speech, gesture, and composite utterances.* Cambridge: Cambridge University Press.

Englund Dimitrova, B. (1997). Degree of Interpreter Responsibility in the Interaction Process in Community Interpreting. In Carr, S.E. (Ed.), *The Critical Link: Interpreters in the Community.* Amsterdam/Philadelphia: Benjamins, pp. 147-164.

Ford, C. E. & Stickle, T. (2012). Securing recipiency in workplace meetings: Multimodal practices. *Discourse Studies,* 14 (1): pp. 11-30.

Ford, C. E. (2004). Contingency and units in interaction. *Discourse Studies,* 6 (1): pp. 27-52.

Ford, C., Fox, B. & Thompson A. S. (1996). Practices in the construction of turns: The TCU revisited. *Pragmatics,* 6 (3): pp. 427-454.

Frischen, A., Andrew P. B., & Steven P. T. (2007). Gaze Cueing of Attention: Visual Attention, Social Cognition, and Individual Differences. *Psychological Bulletin*, 133 (4), pp. 694-724.

Gardner, R. (2001). *When Listeners Talk: Response tokens and listener stance.* Amsterdam: J. Benjamins Publishing.

Gavioli, L. (2012). Minimal responses in interpreter-mediated medical talk. In Baraladi, L. & Gavioli, C. (Eds.), *Coordinating Participation in Dialogue Interpreting.* Amsterdam & Philadelphia: Benjamins Publishing, pp. 201-228.

Gerwing, J. & Bavelas, J. (2013). The social interactive nature of gestures: Theory,

assumptions, methods, and findings. In Müller, C., Cienki, A., Fricke, E. Ladewig, S.H., McNeil, D. & Tessendorf, S. (Eds.), *Body-language-communication, Volume I, Contemporary approaches* Berlin: Mouton De Gruyter, 4 (51): pp. 816-831.

Goffman, E. (1963). *Behavior in public places: Notes on the social organization of gatherings*. New York: The Free Press.

Goffman, E. (1981). *Forms of talk.* Oxford: Basil Blackwell.

Goodwin, C. & Goodwin, M.H. (2004). Participation. In Duranti, A. (Ed.), *A Companion to Linguistic Anthropology*. Oxford: Basil Blackwell, pp. 222-243.

Goodwin, C. (1979). The Interactive Construction of a Sentence in Natural. Conversation. In Psathas, G. (Ed.), *Everyday Language: Studies in Ethnomethodology*. New York, Irvington Publishers, pp. 97-121.

Goodwin, C. (1980). Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Sociological Inquiry*, 50: pp. 272-302.

Goodwin, C. (1981). *Conversational Organization. Interaction between Speakers and Hearers.* New York, London: Academic.

Goodwin, C. (1984). Notes on story structure and the organization of participation. In J.M. Atkinson and J. Heritage (ed.) *Structures of social action. Studies in Conversation Analysis*. Cambridge: Cambridge University Press, pp. 225-246.

Goodwin, C. (1986). Between and within alternative treatments of continuers and assessments. *Human Studies,* 9: 205-217.

Goodwin, C. (2002). Time in action. *Current Anthropology*, 43 (4): pp. 19-35.

Goodwin, C. (2013). The co-operative, transformative organization of human action and knowledge. *Journal of Pragmatics,* 46: pp. 8-23.

Gravano, A., Hirschberg, J. & Beňuš, Š. (2013). Affirmative Cue Words in Task-Oriented Dialogue. *Computational Linguistics,* 38 (1).

Gullberg, M. & Kita, S. (2009). Attention to speech-accompanying gestures: Eye movements and information uptake. *Journal of Nonverbal Behaviour* 33: pp. 251-277.

Haddington, P. (2006). The organization of gaze and assessments as resources for stance taking. *Text & Talk,* 26 (3): pp. 281-328.

Hjalmarsson, A. (2010). The vocal intensity of turn-initial cue phrases in dialogue. *Proceedings of the (SIGIDIAL) 2010 Conference, the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, Tokio, Japan.

Hansen, J. P. B. (2016). Interpreting at a distance: A comparative analysis of turn-taking in video remote interpreting and on-site interpreting. Master thesis. University of Oslo.

Hayashi, M. (2013). Turn Allocation and Turn Sharing. In Sidnell, J. & Stivers, T. (Eds.), *The Handbook of Conversation Analysis*. Oxford: Wiley Blackwell, pp. 167-190.

Healey, P.G.T., Lavelle, M., Howes, C., Battersby, S. & McCabe, R. (2013). How listeners respond to speaker's troubles. CogSci.

Heath, C. & Luff, P. (2013). Embodied action and organizational interaction: Establishing contact on the strike of a hammer. *Journal of Pragmatics,* 46 (1): pp. 24-38.

Heath, C. (1984). Talk and recipiency: Sequential organization in speech and body movement. In Atkinson, J.M. & Heritage, J. (Eds.), *Structure of social action*. Cambridge: Cambridge University Press, pp. 247-65.

Heath, C. (1986). *Body Movement and Speech in Medical Interaction*. Cambridge: Cambridge University Press.

Heath, C. (1992). Gesture's discreet tasks: Multiple relevancies in visual conduct and in the contextualization of language. In Auer, P. & di Luzio, A. (Eds.), *The contextualization of language.* Amsterdam: John Benjamins, pp. 101-128.

Heinemann, T. & Koivisto, A. (2016). Indicating a change-of-state in interaction: Cross-linguistic explorations. *Journal of Pragmatics,* 104: pp. 83-88.

Hepburn, A. & Bolden, G.B. (2013). Transcription. In Stivers, T. and Sidnell, J. (Eds.), *The handbook of Conversation Analysis*. Oxford: Wiley-Blackwell, pp. 57-76.

Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In

Atkinson, M. & Heritage, J. (Eds.), *Structures of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press, pp. 299-345.

Heritage, J. (2011). Territories of knowledge, territories of experience: Empathic moments in interaction. In Stivers, T., Mondada, L. & Steensig, J. (Eds.), *The morality of knowledge in conversation.* Cambridge: Cambridge University Press, pp. 159-183.

Heritage, J. (2012). Epistemics in action: Action Formation and Territories of Knowledge, *Research on Language and Social Interaction*, 45 (1): pp. 1-29.

Heritage, J. (2013). Epistemics in conversation. In Sidnell, J. & Stivers, T. (Eds.), *Handbook of Conversation Analysis*. Boston: Wiley-Blackwell, pp. 370-394.

Holler, J. & Kendrick, K. H. (2015). Unaddressed participants' gaze in multi-person interaction: Optimizing recipiency. *Frontiers in Psychology*, 6: pp. 98.

Holmqvist, K. & Andersson, R. (2017). *Eye tracking: A comprehensive guide to methods, paradigms and measures*, Lund, Sweden: Lund Eye-Tracking Research Institute.

Hömke, P., Holler, J. & Levinson, S.C. (2017). Eye Blinking as Addressee Feedback in Face-To-Face Conversation, *Research on Language and Social Interaction*, 50 (1): pp. 54-70.

Hoey, E., & Kendrick, K. H. (2017). Conversation Analysis. In A. M. B. De Groot, & Hagoort, P. (Eds.), Research methods in psycholinguistics and the neurobiology of language: A practical guide, Hoboken: Wiley, pp. 151-173.

Houtkoop, H. & Mazeland, H. (1985). Turns and discourse units in everyday conversation. *Journal of Pragmatics*, 9: pp. 595-619.

Jefferson, G. (1984). Notes on some orderlinesses of overlap onset. In D'Urso, V. & Leonardi, P. (Eds.), *Discourse analysis and natural rhetoric*. Padua: Cleup Editore, pp. 11-38.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In Gene

Lerner (ed.), *Conversation analysis: Studies from the first generation*. Amsterdam: John Benjamins, pp. 13-31.

Jokinen, K. (2010). Non-verbal signals for turn-taking & feedback. *Proceedings Of 7th International Conference on Language Resources & Evaluation (LREC) International Universal Communication Symposium.*

Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica,* 26: pp. 22-63.

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kendrick, K.H. & Holler, J. (2017). Gaze direction signals response preference in conversation. *Research on Language and Social Interaction*, 50: pp. 12-32.

Krystallidou, D. (2012). On Mediating Agents' Moves and How They Might Affect Patient-centredness in Mediated Medical Consultations. *Linguistica Antverpiensia*, 11: pp. 75-93.

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): pp. 399-417.

Lang, R. (1978). Behavioral aspects of liaison interpreters in Papua New Guinea: some preliminary observations. In Gerver, D. & Sinaiko, H. W. (Eds.), *Language Interpretation and Communication.* New York & London: Plenum Press, pp. 231-244.

Lerner, G. H. (1991). On the Syntax of Sentences in Progress. *Language in Society,* 20: pp. 441-458.

Lerner, G. H. (1992). Assisted Storytelling: Deploying Shared Knowledge as a Practical Matter. *Qualitative Sociology*, 15 (3): pp. 247-271.

Lerner, G. H. (2003). Selecting next speaker: The context-sensitive operation of a context-free organization. *Language in Society,* 32: pp. 177-201.

Levinson, S. C. (2013). Action formation and ascription. In Sidnell, J. & Stivers, T. (Eds.), *The Handbook of Conversation Analysis*. Malden: Wiley-Blackwell, pp.103-131

Levinson, S. C. & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. In Holler, J., Kendrick, K.H., Casillas, M. & Levinson, S.C. (Eds.), *Turn-taking in human communicative interaction, Frontiers in Psychology*, pp. 10-27.

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

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

Licoppe, C., Verdier M., & Veyrier, C.A. (2018). Voice, control and turn-taking in multi-lingual, consecutively interpreted courtroom proceedings with video links. In Skinner, R., Napier, J. & Braun, S. (Eds.), *Remote Interpreting*. Washington: Gallaudet University Press, pp. 299-322.

Linell, P., Linda J. & Cecilia W. (1992). Establishing Communicative Contact through a Dialogue Interpreter. In A. Grindsted, A. & Wagner, J. (Eds.), *LSP Communication/Fachsprachliche Kommunikation.* Tübingen: Narr Verlag, pp. 125-142.

Llewellyn-Jones, P. & Lee, R. G. (2014). *Redefining the Role of the Community Interpreter: The Concept of role-space*. Lincoln: SLI Press.

Lund, K. (2007). The importance of gaze and gesture in interactive multimodal explanation. *Language Resources and Evaluation,* 41: pp. 289-303.

Mandelbaum, J. (2013). Storytelling in Conversation. In Sidnell, J. & Stivers, T. (Eds.), *The Handbook of Conversation Analysis*. Malden: Wiley-Blackwell, pp. 493-507.

Mason, I. (Ed.) (1999). *Dialogue interpreting*. Special issue of The Translator 5 (2).

Mason, I. (Ed.) (2009). Role, Positioning and Discourse in Face-to-Face Interpreting. In de Pedro Ricoy, R., Perez, I. & Wilson, C. (Eds.), *Interpreting and Translating in Public Service Settings. Policy, Practice, Pedagogy*. Manchester: St Jerome, pp. 52-73.

Mason, I. (2012). Gaze, positioning and identity in interpreter-mediated dialogues. In

Baraladi, L. & Gavioli, C. (Eds.), *Coordinating Participation in Dialogue Interpreting.* Amsterdam & Philadelphia: Benjamins Publishing, pp. 177-199.

Mazeland, H. (1990). 'Yes', 'no' and 'mhm': Variations in acknowledgment choices. *Communication-Technologie-Societe,* 8 (1): pp. 251-282.

Merlini, R. & Favaron, R. (2005). Examining the "voice of interpreting" in speech pathology. *Interpreting, 7* (2): pp. 263-302.

Merlini, R. (2015). Dialogue Interpreting. In Pöchhacker, F. (Ed.), *The Routledge Encyclopedia of Interpreting Studies*. London & New York: Routledge, pp. 102-103.

Merlino, S. & Mondada, L. (2014). Identités fluides dans le travail interactionnel du traducteur improvisé. In Greco, L., Mondada, L. & Renaud, P. (Eds.), *Identités en interaction*. Faits de Langue, pp. 87-144.

Merlino, S. (2014). Traduction orale et organisation de la parole: la gestion multimodale des transitions. In Mondada, L. (Ed.), *Corps en interaction: participation, spatialité, mobilité,* Lyon: ENS Editions, pp. 65-105.

Mondada, L. (2007). Multimodal resources for turn-taking: pointing and the emergence of possible next speakers. *Discourse Studies, 9 (2)*: pp. 195-226.

Mondada, L. (2011). Understanding as an embodied, situated and sequential achievement in interaction. *Journal of Pragmatics*, 43 (2): pp. 542-552.

Mondada, L. (2013). Video as a tool in the social sciences. In Müller, C., Cienki, A., Fricke, E., Ladewig, S.H., McNeill, D. & Teßendorf, S. (Eds.), *Body - Language Communication: An International Handbook on Multimodality in Human Interaction, Vol. I.* Berlin: De Gruyter Mouton, pp. 982-992.

Mondada, L. (2014). Conventions for multimodal transcription. Basel: Romanisches Seminar der Universität, https://franzoesistik.philhist.unibas.ch/fileadmin/user_upload/franzoesistik/mondada_multimodal_conventions.pdf

Mondada, L. (2016). Challenges of multimodality: Language and the body in social interaction. *Journal of Sociolinguistics,* 20 (3): pp. 336-366.

Mondada, L. (2018). Multiple Temporalities of Language and Body in Interaction:

Challenges for Transcribing Multimodality. *Research on Language and Social Interaction* 51 (1): pp. 85-106.

Müller, F.E. (1996). Affiliating and disaffiliating with continuers: Prosodic aspects of recipiency. In Couper-Kuhlen, E. & Selting, M. (Eds.), *Prosody in conversation: Interactional studies.* Cambridge: Cambridge University Press, pp. 131-176.

Muntigl, P., Knight, N. & Watkins, A. (2012). Working to keep aligned in psychotherapy: using nods as a dialogic resource to display affiliation. *Language and Dialogue,* 2 (1): pp. 9-27.

Muntigl, P. & Horvath, A. O. (2014). The therapeutic relationship in action: How therapists and clients co-manage relational disaffiliation. *Psychotherapy Research,* 24 (3): pp. 327-345.

Nielssen, G. (1962). *Studies in self-confrontation*. Copenhagen: Munksgaard.

Norrick, N. R. (2010). Listening practices in television celebrity interviews. *Journal of Pragmatics,* 42: pp. 525-543.

Norrick, N.R. (2012). Listening practices in English conversation: The responses responses elicit, *Journal of Pragmatics*, 44 (5): pp. 566-576.

Norris, S. (2006). Multiparty Interaction: A Multimodal Perspective on Relevance. *Discourse Studies*, 8 (3): pp. 401-421.

Oben, B. (2015). *Modeling interactive alignment: a multimodal and temporal account.* Unpublished PhD dissertation, KU Leuven.

Oben, B. & Brône; G. (2015). What you see is what you do. On the relationship between gaze and gesture in multimodal alignment. *Language and Cognition,* 7: pp. 546-562.

Oben, B. & Brône, G. (2015). InSight Interaction: a multimodal and multifocal dialogue corpus. *Language Resources and Evaluation,* 49: pp. 195-214.

Pasquandrea, S. (2011). Managing multiple actions through multimodality: Doctors' involvement in interpreter-mediated interactions. *Language in Society,* 40 (4): pp. 455-481.

Pasquandrea, S. (2012). Co-constructing dyadic sequences in healthcare interpreting: A multimodal account. *New Voices in Translation Studies,* 8: pp. 132-157.

Peräkylä, A. & Ruusuvuori, J. (2012). Facial Expression and Interactional Regulation of Emotion. In Peräkylä, A. & Sorjonen, M.L. (Eds.), *Emotion in interaction*. Oxford: Oxford University Press, pp. 64-91.

Peräkylä, A. (2013). Conversation Analysis in Psychotherapy. In Sidnell, J. & Stivers, T. (Eds.), *The Handbook of Conversation Analysis*. Chichester: Wiley-Blackwell, pp. 551-574.

Pöchhacker, F. & Schlesinger, M. (Eds.) (2007). Healthcare interpreting. Amsterdam: John Benjamins.

Pöchhacker, F. (2018). "Going video": Mediality and Multimodality in Interpreting. In Salaets, H. & Brône, G. (Eds.), *Linking up with Video: Perspectives on Interpreting Practice and Research*. John Benjamins.

Pöllabauer, S. (2004). Interpreting in asylum hearings: Issues of role, responsibility and power. *Interpreting*, 6 (2): pp. 143-180.

Pomerantz, A.M. & Heritage, J. (2013). Preference. In Sidnell, J, & Stivers, T. (Eds.), *Handbook of Conversation Analysis*. New York: Wiley-Blackwell, pp. 210-228.

Raymond, C.W. (2014). Epistemic Brokering in the Interpreter-Mediated Medical Visit: Negotiating "Patient's Side" and "Doctor's Side" Knowledge, *Research on Language and Social Interaction*, 47 (4): pp. 426-446.

Reddy, M. (1979). The Conduit Metaphor: A case of frame conflict in our language about language. In Ortony, A. (Ed.), *Metaphor and Thought (1993)*. Cambridge: Cambridge University Press, pp. 164-201.

Rossano, F. (2012). *Gaze behavior in face-to-face interaction*. PhD Thesis, Radboud University Nijmegen, The Netherlands.

Rossano, F. (2013). Gaze Conversation. In Sidnell, J. & Stivers, T. (Eds.), *The Handbook of Conversation Analysis*. Malden: Wiley-Blackwell, pp. 308-329.

Rossano, F., Brown, P. & Levinson, S. C. (2009). Gaze, questioning, and culture. In Sidnell, J. (Ed.), *Conversation analysis: Comparative perspectives*. Cambridge: Cambridge University Press, pp. 187-249.

Roy, C. (2000). *Interpreting as a discourse process*. New York: Oxford University

Press.

Ruusuvuori, J. & Peräkylä, A. (2009). Facial and verbal expressions in assessing stories and topics. *Research on Language and Social Interaction,* 42 (4): pp. 377-394.

Sacks, H., Schegloff, E. A. & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language,* 50: pp. 696-735.

Schegloff, E.A., Jefferson, G. & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language,* 53: pp. 361-382.

Schegloff, E.A. (1982). Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. In Tannen, D. (Ed.), *Analyzing Discourse: Text and Talk.* Georgetown University Roundtable on Languages and Linguistics 1981. Washington: Georgetown University Press, pp. 71-93.

Schegloff, E. A. (1996). Turn Organization: One Intersection of Grammar and Interaction. *Interaction and Grammar,* pp. 52-133.

Schegloff, E. A. (2000). Overlapping talk and the organization of turn-taking for conversation. *Language in Society,* 29: pp. 1-63.

Schmitt, R. (2005). Zur multimodalen Struktur von turn-taking. *Gesprächsforschung* 6: pp. 17-61.

Seeber, K. (2013). What eye-tracking can tell us about simultaneous interpreting with text. *CRELA- Cultures de recherche en linguistique appliquée.*

Seeber, K. (2015). Eye-tracking. In Pöchhacker, F., Grbic, N., Mead, P. & Setton, R. (Eds.) *Routledge Encyclopedia of Interpreting Studies.* Oxford and New York: Routledge.

Selting, M. (2000). The construction of units in conversational talk. *Language in Society*, 29 (4): pp. 477-517.

Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J., Bergmann, P., Birkner, K., Couper-Kuhlen, E., Deppermann, A., Gilles, P., Günthner, S., Hartung, M., Kern, F., Mertzlufft, C., Meyer, C., Morek, M., Oberzaucher, F., Peters, J., Quasthoff, U., Schütte, W., Stukenbrock, A., & Uhmann, S. (2011). A system for transcribing talk-in-interaction: GAT

2. *Gesprächsforschung - Online - Zeitschrift zur verbalen Interaktion*, 12: pp. 1-51.

Stivers, T. & Nick J. E. (2010). A coding scheme for question-response sequences in conversation. *Journal of Pragmatics,* 42 (10): pp. 2620-2626.

Stivers, T. & Rossano, F. (2010). Mobilizing Response. *Research on Language and Social Interaction,* 43 (1): pp. 3-31.

Stivers, T. (2008). Stance, alignment and affiliation during story telling: when nodding is a token of preliminary affiliation. *Research on Language in Social Interaction,* 41: pp. 29-55.

Stivers, T. (2013). Sequence organization. In Sidnell, J. and Stivers, T. (Eds.), *The Handbook of Conversation Analysis.* Chichester: Wiley-Blackwell, pp. 191-209.

Stivers, T. (2015). Coding social interaction: A heretical approach in conversation analysis? *Research on Language and Social Interaction,* 48 (1): pp. 1-19.

Stivers, T., Mondada, L. & Steensig, J. (2011). Knowledge, morality and affiliation in social interaction. In Stivers, T., Mondada, L. & Steensig, J. (Eds.), *The morality of knowledge in conversation.* Cambridge: Cambridge University Press, pp. 3-26.

Streeck, J. (2009). *Gesturecraft - The manufacture of meaning*. Amsterdam & Philadelphia: John Benjamins.

Streeck, J. (2014). Mutual gaze and recognition. Revisiting Kendon's Gaze direction in two-person conversation. In Seyfeddinpur, M. & Gullberg, M. (Eds.), *From Gesture in Conversation to Visible Action as Utterance: Essays in honor of Adam Kendon,* John Benjamins Publishing Company, pp. 35-56.

Streeck, J. (2011). The Changing Meanings of Things: Found Objects and Inscriptions in Social Interaction. In Goodwin, C., Streeck, J. & LeBaron, C. (Eds.), *Embodied Interaction: Language and Body in the Material World.* New York: Cambridge University Press, pp. 67-78.

Stukenbrock, A. (2018). Blickpraktiken von SprecherInnen und AdressatInnen bei der

Lokaldeixis: Mobile Eye Tracking-Analysen zur Herstellung von joint attention. *Gesprächsforschung - Online - Zeitschrift zur verbalen Interaktion* 19: pp. 132-168.

Tebble, H. (2012). Interpreting or interfering? in Baraldi, C. & Gavioli, L. (Eds.), *Coordinating Participation in Dialogue Interpreting*. Amsterdam: John Benjamins, 1: pp. 23-44.

Thompson, S.A., Fox, B.A. & Couper-Kuhlen, E. (2015). *Grammar in Everyday Talk: Building Responsive Actions*. Cambridge: Cambridge University Press.

Tolins, J., & Fox Tree, J. E. (2014). Addressee backchannels steer narrative development. *Journal of Pragmatics,* 70: pp. 152-164.

Van De Mieroop, D. & Mazeland, H. (2009). Modificaties van vraag-antwoordsequenties in getolkte arts-patiëntgesprekken (Modifications of question-answer sequences in interpreted doctor-patient interactions). *Tijdschrift voor Taalbeheersing*. 31 (2): pp. 124-141.

Van De Mieroop, D. (2012). The Quotative 'He/She Says' in Interpreted Doctor-Patient Interaction. *Interpreting. An International Journal of Research and Practice in Interpreting,* 14 (1): pp. 91-116.

Vanhoomissen, T. (2012). *Inductieve statistiek voor de gedragswetenschappen*. Leuven: Acco.

Vertegaal, R., Slagter, R., Van der Veer, G. & Nijholt, A. (2001). Eye gaze Patterns in Conversations: There is More to Conversational Agents Than Meets the Eyes. *Proceedings of the Conference on Human Factors in Computing Systems*, pp. 301-308.

Vranjes, J., Bot, H., Feyaerts, K. & Brône, G. (forthcoming). Affiliation in interpreter-mediated therapeutic talk: on the relationship between gaze and head nods. *Interpreting. An International Journal of Research and Practice in Interpreting.*

Vranjes, J., Bot, H., Feyaerts, K. & Brône, G. (2018a). Displaying recipiency in an interpreter-mediated dialogue: an eye-tracking study**.** In Oben B. & Brône G. (Eds.), *Eye-tracking in interaction. Studies on the role of eye gaze in dialogue*. Amsterdam: Benjamins.

Vranjes, J., Brône, G. & Feyaerts, K. (2018b). Dual feedback in interpreter-mediated interactions: on the role of gaze in the production of listener responses. *Journal of Pragmatics*, 134: pp. 15-30.

Vranjes, J., Brône, G. & Feyaerts, K. (2018c). On the Role of Gaze in the Organization of Turn-Taking and Sequence Organization in Interpreter-Mediated Dialogue. *Language and Dialogue*, 3.

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

Wadensjö, C. (2001). Interpreting in Crisis: The Intepreters' Position in Therapeutic Encounters. In Mason, I. (Ed.), *Triadic Exchanges. Studies in Dialogue Interpreting*. Manchester: St. Jerome Publishing, pp. 71-85.

Wadensjö, C. (2015). Current and Future Themes in the Study of Interpreter-Mediated Interaction. Paper presented at IMDI (Integrating Multimodality in the study of Dialogue Interpreting) research seminar, Guidford, UK.

Weiß, C. & Auer, P. (2016). Das Blickverhalten des Rezipienten bei Sprecherhäsitationen: eine explorative Studie. *Gesprächsforschung - Online - Zeitschrift zur verbalen Interaktion*, 17: pp. 132-167.

Whitehead, K. A. (2011). Some uses of head nods in 'third position' in talk-in-interaction. *Gesture*, 11 (2): pp. 103-122.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A. & Sloetjes, H. (2006). ELAN: a Professional Framework for Multimodality Research. In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*.

Yngve, V. H. (1970). On getting a word in edgewise. In *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, pp. 567-577.

Samenvatting

Wij leven in een geglobaliseerde wereld, waarin elke dag mensen met verschillende talige achtergronden met elkaar in contact komen. Om wederzijds begrip mogelijk te maken tussen gesprekspartners die elkaars taal niet verstaan worden er vaak tolken ingeschakeld. Het tolkproces werd traditioneel opgevat als reproductie van 'teksten' en de tolk als een onzichtbare 'spreekbuis' die de boodschappen omzet van de ene taal naar de andere. Deze opvatting stemt echter niet overeen met de complexe realiteit van getolkte gesprekken, die eerder als 'interacties' gezien moeten worden. In getolkte interacties speelt het non-verbale gedrag, en in het bijzonder waar gesprekspartners naar kijken, een belangrijke rol voor de coördinatie van het gesprek. Het is echter niet duidelijk *hoe* blikrichting in getolkte interacties georganiseerd is in correlatie met taal en andere non-verbale signalen en welke rol het speelt in de interactie. Systematisch onderzoek naar de correlatie tussen blikrichting, gesproken taal en andere non-verbale signalen in getolkte interacties werden lange tijd belemmerd door moeilijkheden in het verzamelen van videocorpora, die een dergelijke studie mogelijk zouden maken. Dit proefschrift levert een bijdrage in die richting door te laten zien welke rol blikrichting speelt in twee processen die universeel zijn voor elke vorm van sociale interactie: (a) de productie van feedback signalen (*backchannel responses*) en (b) de regeling van beurtwisseling (*turn-taking*). De studies gepresenteerd in dit proefschrift zijn gebaseerd op een uniek videocorpus van getolkte gesprekken, die met mobiele eye-trackingtechnologie werden opgenomen. Ik hanteerde daarbij een mixed-methods approach, steunend op de inzichten uit tolkwetenschap, (multimodale) conversatieanalyse (CA) en de 'joint action theory'.

Hoofdstuk 1 situeert het proefschrift in de context van de literatuur over getolkte interacties (tolkwetenschap), blikrichting, organisatie van menselijke interacties en multimodaliteit. Het hoofdstuk introduceert de veronderstellingen die aan de basis

liggen van dit interdisciplinaire onderzoek en presenteert de specifieke bijdragen van de volgende hoofdstukken.

Hoofdstuk 2 introduceert de dataset bestaande uit negen getolkte student-ombuds gesprekken in een universitaire context die consecutief werden getolkt door een aanwezige tolk. De gesproken talen waren Russisch en Nederlands. Het hoofdstuk bespreekt verder de methode (mobiele eye-tracking) en dataverwerking. Het sluit af met een toelichting over de gehanteerde transcriptiesmethode voor blikrichting voor de kwalitatieve analyses van dit proefschrift.

Hoofdstuk 3 geeft een kwantitatief overzicht over de verdeling van visuele aandacht van alle deelnemers in de opgenomen gesprekken. De resultaten laten opvallende verschillen zien in de verdeling van visuele aandacht van de 'primaire' gesprekspartners (student en ombuds) die gelinkt blijken te zijn aan de verschillen in hun positie en professionele status in het gesprek. Het hoofdstuk legt een link met de resultaten uit vorige studies over blikrichting in (getolkte) interacties en nuanceert ze tegelijkertijd.

Hoofdstuk 4 analyseert de rol van blikrichting in beurtwisseling van getolkte interacties. De studie is gericht op op het proces van 'chunking', waarbij de primaire gesprekspartner zijn/haar langere beurten opsplitst, zodat de tolk kortere stukken zou kunnen vertalen. De studie laat zien hoe de tolk, door middel van haar blikverschuiving van de ene gesprekspartner naar de andere bijdraagt aan het vlotte verloop van 'chunking'. De studie bekrachtigt de opvatting van de tolk als een co-participant, in plaats van een spreekbuis, in het gesprek.

In Hoofdstuk 5 wordt de rol van blikrichting in de productie van *backchannel responses* zoals 'mm hm', 'ja' en hoofdknikken, bestudeerd. Het hoofdstuk begint met een overzicht van de relevante literatuur over *backchannel responses* in gesproken interacties. Het analytische deel van het hoofdstuk is opgesplitst in drie delen. Het eerste deel geeft een kwantitatief overzicht van de productie van *backchannel responses* in de dataset. Het tweede deel focust op de relatie tussen (wederzijds) blikcontact en *backchannel responses*, in navolging van vorige studies die een verband daartussen hadden aangetoond. De analyses laten zien dat er een bijkomend onderscheid tussen non-verbale *backchannel responses* (zoals hoofdknikken) en verbale backchannel

responses nodig is, en dat non-verbale backchannel responses vaker geproduceerd worden na wederzijds blikcontact met de spreker. Bovendien blijkt dat in de huidige dataset enkel het geval te zijn voor de tolk . Het derde deel analyseert een terugkerend fenomeen van backchannel responses in de dataset, dat ik 'dual feedback' heb genoemd. Dat zijn *backchannel responses* die samengaan met een blikverschuiving van de huidige spreker naar de andere participant (en terug). De analyse geeft aan dat zulke blikverschuivingen systematisch voorkomen met *backchannel responses* en dat ze 'composite signals' vormen. De analyse laat verder zien dat dual feedback een korte verschuiving in het participatiekader veroorzaakt, waardoor de momenteel 'uitgesloten' gesprekspartner in het gesprek wordt betrokken. Het toont ook aan dat luisteraars gevoeligheid vertonen voor de verschillen in participatiestatus en *epistemische* posities van de gesprekspartners. Die blijkt bijzonder evident tijdens de productie van *newsmarkers* (zoals 'Ah ja', 'Oh!') en affiliërende *backchannel responses*.

Hoofdstuk 6 presenteert een slotstuk van de analyses geïntroduceerd in het vorige hoofdstuk. Hierin wordt een case study van een getolkt therapeutisch gesprek in een psychotherapeutische instelling besproken. Deze studie laat zien hoe sommige fenomenen die in Hoofdstuk 5 werden beschreven, samenkomen in de context van een authentiek, therapeutisch gesprek en levert daarmee een argument voor de ecologische validiteit van deze bevindingen.

In Conclusie worden de belangrijkste bevindingen van dit onderzoek overlopen. Daarna volgt een korte discussie over de implicaties van deze bevindingen voor het onderzoek en voor de (tolk)praktijk. Tenslotte worden de beperkingen van deze studie aangegeven alsook suggesties voor vervolgonderzoek.