Anticipation and timing of turn-taking in dialogue interpreting:

a quantitative study using mobile eye-tracking data

Jelena Vranjes (UGent) and Bert Oben (KU Leuven)

This paper presents the results of an exploratory study on the timing of turn-taking in face-to-face dialogue interpreting based on a corpus of interpreted interactions that were recorded with mobile eye trackers. Our aim was to investigate (1) the timing of interpreter's turns in dialogic interaction and identify (2) features that have an impact on the interpreter's turn-taking speed. These include input processing factors (including turn type and turn duration) and gaze, which have been shown to play an important role in turn-taking. The analysis shows that, although interpreters in our data tend to orient to the maxim 'one speaker at a time', turn transitions between the primary speaker and the interpreter contain more gaps and longer overlaps than in same-language interactions. It also shows that the type of turn produced by the primary speaker (question vs non-question), the primary speaker's speech rate and to a certain extent turn duration affect the interpreter's turn-taking speed. Thus, the present study aims to contribute to a better understanding of the processes that impact the timing of turn-taking in face-to-face dialogue interpreting.

Keywords*:* dialogue interpreting, turn-taking, timing, corpus analysis, eye gaze, eye-tracking

1. **Introduction**

Turn-taking is a fundamental aspect of dialogue interpreting. Typically, dialogue interpreting is conducted in consecutive mode and the interpreter participates in the exchange by taking every second turn (Bot 2005, Mason 2009). In this process, the interpreter does not merely translate primary speakers' utterances, but is an active participant who is critically engaged in the negotiation of meaning and in the coordination of the interaction (Wadensjö 1998, Angelelli 2000, Davidson 2002, Gavioli 2015, Wadensjö 2018, Gavioli and Wadensjö 2020). Although the interpreter participates in the interaction by taking turns, finding appropriate moments to take the floor is not an easy task. It is important that this process runs smoothly, as problems in turn-taking between the interpreter and the primary participants may affect the accuracy and fluency of the interpreter's rendition and the overall success of communication (see Braun 2013, De Boe 2020, Vranjes & Bot, 2021). Moreover, while taking part in the exchange, the interpreter is at the same time engaged in a cognitively demanding activity involving comprehension, memorisation and reproduction of the incoming turn in a target language (see Gile 1997). This may affect the interpreter's turn-taking behaviour: for instance, it has been argued that interpreters usually benefit from shorter primary speakers' turns, since turns that are too long could overburden the interpreter's memory capacity (Bot 2005, see also Englund Dimitrova 1997). Both the interpreter's position in the exchange as a 'mediating' third party and her involvement in a cognitively complex task will inevitably have an impact on her decision when to take the turn.

While previous studies in interpreting have examined turn-taking with reference to the construction of conversational meaning (Davidson 2002), interpreter's coordination of talk (Wadensjö 1998, Davitti 2013) and multimodal cues involved in the negotiation of turn transfers (Pasquandrea 2011, Davitti 2013, Vranjes & al. 2018, Vranjes & Brône 2020, Vranjes & Bot 2021), *timing* of turn-taking has not been the object of systematic analysis, despite its importance in the interpreting process and a growing number of studies focusing on timing in same-language interactions. Research on the conversational organization of monolingual, spontaneous interactions has pointed out that interlocutors tend to minimize gaps (silences between turns) and overlaps (simultaneous talk) when taking turns (Sacks et al. 1974, Stivers et al. 2009, Levinson & Torreira 2015). This is achieved through the process of anticipation that allows the participant to "gear up in advance to begin talking at just the right moment" (De Ruiter et al. 2006, 517). Moreover, recent psycholinguistic studies have identified features of the speaker's turn that can affect the listener's turn-taking times, such as turn length, speed, frequency of words, type of turn and nonverbal signals, arguing that duration of gaps between turns reflects the amount of processing required to comprehend the previous turn and plan the upcoming turn, (De Ruiter et al. 2006, Heldner & Edlund 2010; Roberts et al. 2015). The question is whether interpreters display similar tendencies (i.e. avoidance of gaps and overlaps) during the process of dialogue interpreting and which factors impact the interpreter's turn-taking speed.

Drawing on the insights from Conversation Analysis and psycholinguistic research of the turn-taking organization, this paper presents a systematic analysis of the timing of turn-taking in dialogue interpreting based on a video corpus recorded with mobile eye-trackers. More specifically, the following questions will be addressed: (1) How fast do dialogue interpreters take the turn? (2) Which features of the primary speaker's turn impact the speed with which the interpreter takes the turn? The present study takes both processing factors and factors related to the sequential organization of talk into account. Moreover, the study adopts a multimodal approach by including eye gaze behaviour in the analysis, since gaze has been shown to play an important role in turn-taking in dialogue interpreting (see Mason 2012, Davitti 2013, Vranjes et al. 2018a). In this way, the present study aims to contribute to a better understanding of the socio-cognitive processes underlying interpreter's turn-taking in dialogic interaction.

The paper is structured as follows. We first provide a brief overview of the relevant studies on the organization of turn-taking in interaction (Section 2). After an introduction to the dataset that was used for the present study (Section 3), we describe the method used in annotating and analysing the data. In the analytical part of the paper, we present the findings regarding the timing of turn-taking and the role of gaze in this process (Section 4.1), followed by factors that impact the interpreter's turn-taking speed in our dataset (Sections 4.2). To conclude, we discuss how the analysis of the timing of turn-taking in dialogue interpreting sheds insight into language processing in dialogue interpreting and what are the implications of these findings for the interpreting practice and for our overall understanding of the mental processes underlying dialogue interpreting.

2. Timing of turn-taking in interaction

A key feature of conversation is that people take turns at talking. Over the last four decades, the organization of turn-taking has been described extensively in conversation analytic (CA) research. Sacks et al. (1974) presented in their seminal work the first systematic analysis of turn-taking in spontaneous conversation, arguing that people generally observe the conversational rule of 'one speaker at the time' (1974, 708) and that when overlap occurs, it tends to be brief. Thus, as mentioned above, the turn-taking system as proposed by Sacks et al. (1974) appears to minimise gaps and discourage overlaps between interlocutors, a tendency which was confirmed across languages (see Stivers et al. 2009). There is a clear functional basis for the turns to be immediately adjacent; a timely response expresses a direct link with the preceding turn, shows how the prior utterance was understood and allows for quick correction if necessary (Stivers et al. 2009, 10587). Rapid turn-taking is also important when interlocutors are competing for speaking space, as the speaker who takes the floor first "acquires rights to a turn" (Sacks et al. 1974, 704). However, gaps and overlaps are not necessarily avoided and may carry added conversational implications: gaps that are too long may signal a troubled speaker transition or a disfluency that might need conversational repair (Jefferson 1974); they could also signal non-compliance (Levinson 1983) or an undesired response (Kendrick & Torreira 2015). On the other hand, constant interruptions of the current speaker can be treated as rude or violations that require some form of repair (Levinson & Torreira 2015). Furthermore, conversational significance of gaps and overlaps depends on the context in which they are produced: for instance, a delayed response by a suspect in the context of a police interview does not have the same conversational implication as a delayed response by a patient in the context of psychotherapy.

The tendency for rapid turn transitions in spontaneous conversations was corroborated in recent systematic corpus analyses of turn-taking in spontaneous monolingual interactions. The distribution of gaps and overlaps in those studies is usually plotted as a histogram of FTO's (floor transfer offsets, i.e. the time between the end of a current speaker's turn and the start of a new speaker's turn), negative values corresponding to overlaps and positive values corresponding to gaps (see De Ruiter et al. 2006). Corpus research established that modal response times between turns in different languages fall between 0 and 200 ms (Stivers et al. 2009, Heldner & Edlund 2010, Levinson & Torreira 2015). This is extremely fast, considering the fact that the next speaker needs at least 600 ms to prepare a simple next turn (Roberts et al. 2015). In order to achieve such rapid turn exchanges, recipients must be planning at least part of their upcoming turn while the current speaker is still talking (Sacks et al. 1974, see also Stivers et al. 2009, Levinson & Torreira 2015). Thus, the recipients project the content of the current turn and *anticipate* when the current speaker will complete the turn on the basis of syntactic structure, prosody and pragmatics (Sacks et al. 1974, Ford & Thompson 1996, Selting 1996, Auer 2005). A case in point are phenomena such as anticipatory turn-completions, whereby the recipient utters what the speaker was going to say before the current turn has come to an end (Hayashi 2013, 183). Also, recent findings on eye movements in conversational settings have shown that listeners anticipate upcoming speaker change with their gaze, i.e. by shifting their gaze towards the next speaker well before that speaker actually takes the turn (Holler & Kendrick 2015, Lammertink et al. 2015).

Important to note here is the use of the word ‘anticipation’, which has the object of long-standing research in conference interpreting (Setton 1999, Seeber 2011). While ‘anticipation’ in CA research refers to a basic process underlying turn-taking in conversational interaction, that is predicting what the speaker is going to say, ‘anticipation’ in interpreting research refers to a specific tactic of simultaneous interpreting, entailing “simultaneous interpreter’s production of a constituent (a word or a group of words) in the target language before the speaker has uttered the corresponding constituent in the source language.” (Van Besien 1999; 250). However, even in monolingual comprehension, it has been argued that interpreters are better at predicting what the speaker is going to say and that “anticipation ability is an important goal in the training of interpreters” (Van Besien 1999, 252).

In the present study, we set out to systematically analyse *the timing of turn-taking* in the process of dialogue interpreting and on the factors explaining the variation in interpreter’s turn-taking speed. Although previous studies have suggested that interpreter’s turn-taking behaviour is different from that of interlocutors in same-language interaction (see Englund Dimitrova 1997, Bot 2005), to our knowledge, no previous study has investigated this systematically with reference to interpreter’s turn-taking speed. We will analyse the following features of the primary speaker’s turn as having a potential effect on the interpreter’s turn-taking speed: turn duration, speech rate, word frequency, lexical density, turn type, and speaker gaze. *Turn duration* has been shown to affect turn transitions in monolingual interactions: longer turns result in longer FTO’s as they require more processing from the recipient (Roberts et al. 2015). *Speech rate* has also been reported to affect turn-transition times. Some studies have suggested that listeners adapt their turn-taking speed to the rhythm of the speaker's speech and that faster speech would be followed by shorter FTO's (Couper-Kuhlen 1993, see also Edlund et al. 2009). However, recent findings by Roberts et al. (2015) point to the opposite direction, showing that a faster speech rate in same-language conversations leads to longer gaps between turns, since the recipient has less time to prepare his/her turn. In the context of (simultaneous) interpreting, speech rate is a predictor of cognitive load that has an effect on the interpreter's output (Seeber 2011, Plevoets & Defrancq 2016), which would lead to longer gaps between turns. *Lexical density* of the source speech was also found to increase cognitive load in (simultaneous) interpreting (see Plevoets & Defrancq 2018) and may thus lead to longer FTO's. *Word frequency*, on the other hand, has a positive effect on processing and reduces response times (see Balota et al 2007, Roberts et al. 2015), which suggests that turns containing high frequency words are comprehended faster, leading to shorter FTO's.

A further explanation for the variation in the turn-taking speed is associated with the sequential organization of talk. In conversational settings, speech is not produced in a vacuum but with reference to what came before and what comes next. Certain types of turns (such as questions, offers) set up an expectation for a specific type of reaction (answer, uptake) and thus help the listener to predict what is to come. Roberts et al. (2015, 122) argue that "if initiating actions can be recognised easily, then responding actions may be produced closer to the turn end". *Questions* in particular have been shown to increase predictability and lead to shorter turn transition times in comparison to non-questions (Stivers et al. 2009, Roberts et al. 2015).

Finally, speaker's *gaze behaviour* will be included in the analysis. Gaze is not only used to indicate who is being addressed, but it also signals when the speaker is going to finish his turn: speakers tend to look away from the recipient when taking the turn and look back towards the recipient towards the end of the turn (Kendon 1967, see also Brône et al. 2017). Moreover, speaker's gaze was shown to elicit faster responses from listeners and thus reduce gaps between turns in conversational settings (Stivers et al. 2009). In the context of dialogue interpreting, the primary speakers are usually instructed to gaze at each other, so that gaze shift towards the interpreter towards the end of the turn can be viewed as a marked gaze event with a turn-giving function (see also Vranjes et al. 2018a). As a consequence, we would expect shorter FTO's when such gaze shifts occur (compared to when they do not).

In the following, we will describe the data set and the method used for the analysis of the data. We then discuss how FTO's were extracted and analysed statistically.

3. Method & Data

The data set for this study consists of 9 video recordings of interpreter-mediated interactions in a university setting between a Russian-speaking foreign student, a Dutch speaking university counsellor and an interpreter. We focus on spontaneous interactions that impose no special constraints on talk (who can speak and when), in contrast to highly structured forms of interaction (e.g. in courtrooms) that are determined by formal rules and legal provisions (Komter 2013). 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 and concerns they had. The students and counsellors were previously unacquainted and had very limited knowledge of each other's language. Each conversation was interpreted consecutively by a professional interpreter. In total, three counsellors, three students and three interpreters participated in the study (see also Vranjes 2018a for a detailed description of the dataset). The interpreters were originally from Russia with Russian as their mother tongue. The participants were seated in a triangular formation, with the interpreter in the middle and on more or less equal distance from the primary speakers, who were seated opposite to each other.

The interactions were recorded with a video camera and each participant was also wearing mobile eye-tracking glasses (*Arrington Gig-E60* and *Tobii*), which allow for a fine-grained analysis of gaze direction (e.g. rapid gaze shifts) in ongoing interaction and provide objective information for the analysis of gaze. The screenshot in Figure 1 illustrates the seating arrangement of the participants and the recording set-up. The image on the left is the recording from the mobile eye-tracking glasses worn by the interpreter, with the interpreter's gaze fixation indicated with a green dot. Each conversation was around 20 minutes in length, which amounts to approximately 3 hours of recorded data. The study was approved by the local ethics committee of the Faculty of Arts and Philosophy at KU Leuven (G-2015 02 173). All participants signed an informed consent form prior to the experiment, which ensured their anonymity and stated how the data were going to be used and presented.

INSERT FIG 2 HERE

Figure 1. Recording set-up

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. It is therefore difficult to determine the level of intrusiveness of the mobile eye-trackers on the ongoing talk (Vranjes et al. 2018b). We can assume that the participants were primarily oriented at achieving certain interactional goals during the talk: the students and the counsellors were oriented towards exchanging information and establishing an interpersonal relationship, whereas the interpreters were focused on their task of understanding and rendering the talk. Furhermore, if a participant was concerned about having his/her gaze behavior tracked, arguably it would be difficult to maintain those concerns when faced with socio-cognitive demands associated with interacting with previously unacquainted participants (see Rogers et al. 2018, 5).

For the purpose of this study, we restricted our analysis on turn transitions between the counsellor and the interpreter in order to avoid potential differences in turn-taking times related to interpreter's comprehension in L1 (Russian) and L2 (Dutch). The conversations were transcribed and annotated in the annotation software ELAN (Wittenburg et al. 2006) according to the following principles:

- Each turn was coded for its function ('turn type' in the analysis). We made a distinction between *questions* (Q's) and *non-questions* (NQ's) in our data. Criteria for identifying questions in our dataset were based on the coding scheme proposed by Stivers and Enfield (2010), who state that "a question had to be either (or both) a formal question (i.e., it had to rely on lexico-morpho-syntactic or prosodic interrogative marking) or a functional question (i.e., it had to effectively seek to elicit information, confirmation or agreement whether or not they made use of an interrogative sentence type)" (p. 2621). Interpreter's requests for clarification (repair-initiations) were not taken into account, since, as argued by Holler et al. (2015, 6), "turns dealing with problems of speaking, hearing, and understanding (i.e., other-initiations of repair) are governed by different timing principles and can thus break the common pattern of minimal gaps between turns". Using this coding scheme, we determined 96 questions in our data set and matched that with an equal number of randomly selected non-questions. This random subset was then annotated on the levels described below.

- **Turn onset:** We calculated the FTOs from the end of a primary speaker's turn to the vocal onset of the interpreter's turn (including turn-preserving placeholders such as 'euhm' followed by speech). In-breaths preceding speech were not included in the analysis since they were not always clearly audible in our recordings. To check for inter-coder agreement of our dependent variable FTO, we performed a consistency check on a subset of the data. Both authors calculated the FTO value for 50 randomly selected turns (i.e. roughly 25% of the entire data set). Both the Pearson correlation coefficient and the intraclass correlation coefficient between the two coders was a more than satisfactory 0.99. On average there was a 54 ms difference between the two coders.

- **Turn completion point:** corresponds to the places at which transition from current speaker to interpreter occurs.

- Transitions involving extreme gaps and overlaps (larger than 2200 ms) were not taken into account in the analysis (in line with Roberts et al. 2015).

- The **annotations of gaze** were segmented on the basis of gaze fixations, viz. the moments when eyes stand relatively still, according to a limited tag set, such as 'counsellor's face', student's face', 'wall' etc. We calculated the timing of the interpreter's gaze aversion with reference to the end of the counsellor's turn (see section 4.1). To assess the role of the counsellor's gaze in the interpreter's turn-taking speed, we calculated the timing of the start of the counsellor's gaze shift and the start of the interpreter's turn (see section 4.3)

In our study we want to relate a number of paraverbal factors to our dependent variable FTO. For every turn in our data set we determined the following measures. We calculated speech rate as the number of transcribed characters per second. For lexical density, we calculated the number of content words (i.c. nouns, full verbs, adjectives and adverbs) over the total number of words. To compute word frequency, we used the reference corpus for spoken Dutch (CGN: Corpus Gesproken Nederlands, Oostdijk 2000): departing from the frequency list of the CGN-corpus we calculated (i) the average frequency of all the content words in the turn and (ii) the average rank for all the content words in the turn. Both measures indicate word frequency, but (i) is based on absolute frequency (with very frequent or very infrequent words weighing heavily on the average value) and (ii) is based on a more relative type frequency (with a smoothed average value as a result). To avoid spurious results in computing the paraverbal factors outlined above, we omitted values that were based on less than 10 (content) words. It would be, for example, unfair to calculate an average word frequency for very short turns in which only 3 content words are uttered. This omission had to be carried out for 39 turns.

To answer our research questions, we first zoom in on the distribution of FTO values in our data and provide some qualitative analyses to account for that distribution. Next, we zoom in on relevant factors (cf. supra) that might be good predictors for the FTO's we observe. We do so by first testing the effect of the individual factors, and subsequently by using multi-level analyses for more stringent testing of the relation between our dependent variable (FTO), a series of fixed factors as outlined above (i.c. turn type, turn length, turn duration, counsellor gaze, speech rate, lexical density and word frequency), and the interpreter as a random factor. Models were fit using the R-package lme4 (Bates et al. 2014).

4. Analyses

4.1. Timing of turn-taking in dialogue interpreting

In our corpus of interpreter-mediated dialogues, the average FTO was 258 ms (see dashed red line in Figure 2) and the median value was 314 ms (dashed blue line in Figure 2). This is comparable to what has been observed in other dialogue types such as monolingual face-to-face conversations and telephone conversations (cf. Brady 1968, Heldner & Edlund 2010, Levinson & Torreira 2015). As can be seen in the density plot in Figure 2, and as could be expected from the relation between median and mean, the data are slightly skewed to the left. This means there are more overlaps than to be expected on the basis of a normal distribution of the data.

INSERT FIG 2 HERE

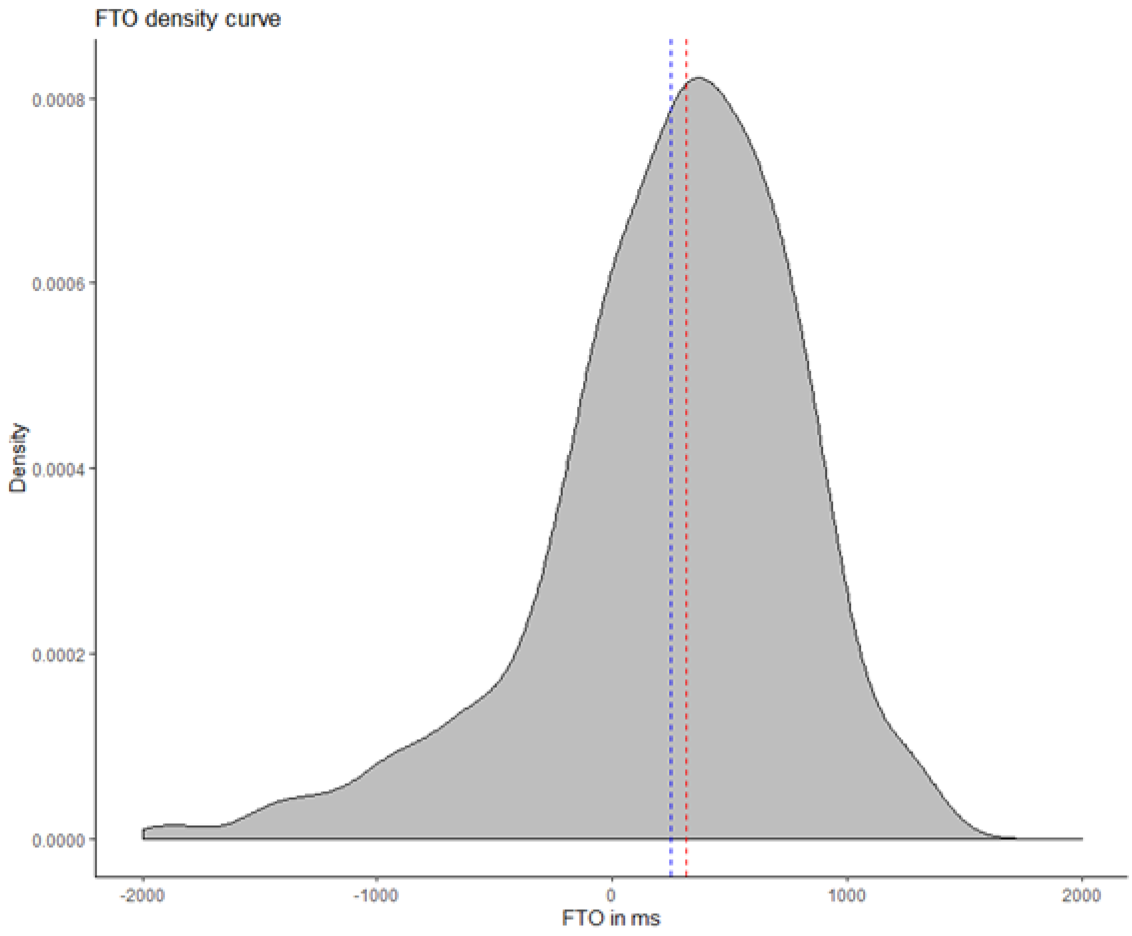

Figure 2. Density plot of floor transfer offsets (FTO's)

At first sight, the FTO distribution in interpreted-mediated dialogue does not differ from spontaneous, monolingual interaction[1] (cf. histograms and density plots in Heldner & Edlund 2010, Levinson & Torreira 2015, Stivers et al. 2009). However, if we zoom in further, we do observe some deviations from the literature on the duration of gaps and pauses. In our data set there was only 1 occurrence of a perfect transition between speakers, i.e. 1 case of an FTO of exactly 0 ms in which two speakers seamlessly switch turns. The remaining FTO's consisted

[1] In contrast to our data, these studies are not based on institutional conversations. Although some institutional interactions involve very specific and systematic transformations in conversational turn-taking procedures (e.g. debates, courtroom hearings etc.), there are many kinds of institutional interaction that use the same turn-taking organization as ordinary conversations (see Heritage 1997: 225). The counselling conversations studied in this paper are clearly "institutional", in the sense of that they involve participants in specific goals orientations which are tied to their institution-relevant identities (see Heritage 1997), however no special turn-taking organization is involved in the conversations.

for 27% out of overlaps and 73% out of gaps. In Table 1 below we compare the distribution of our data to the distributions presented in Heldner & Edlund (2010), which cover data from the Spoken Dutch Corpus (CGN), the HCRC Map Task Corpus (MTC), and the Swedish Map Task Corpus (SMTC).

Table 1. Overview of the distribution of gaps and overlaps in our corpus and those presented in Heldner & Edlund (2010)

	Present study	Heldner & Edlund (2010)
FTO < 0 ms	27%	40%
0 ms < FTO < 200 ms	12%	14-19%
-200 ms < FTO < 200 ms	24%	55-59%
FTO > 200 ms	61%	41-45%
FTO < 500 ms	65%	70-82%
FTO < 1000 ms	97%	82-95%

The most obvious observation from a comparison on the basis of Table 1 is the smaller amount of overlaps (FTO < 0 ms) and larger amount of gaps (FTO > 0 ms) in our interpreter-mediated data, compared to the non-mediated data presented in Heldner & Edlund (2010). In addition, turn transitions between the primary speaker and the interpreter appear to be less 'smooth', i.e. less unnoticeable gaps (gaps shorter than 200 ms) and unnoticeable overlaps (overlaps shorter than 200 ms): in our data only 24% of the turn transitions occur in such a smooth gap/overlap, whereas for the non-mediated data this is 55-59%.

Similar results can be found when comparing our data to Levinson and Toreira (2015), who used the NXT Switchboard Corpus (Calhoun et al. 2010) for their analyses. Levension and Torreira (2015) also report a density plot that is skewed to the left, with more overlaps/short gaps compared to large gaps. However, zooming in further, they report different results for gaps and for overlaps. Concerning gaps, we found a mean FTO of 499 ms, with a median of 475 ms, and modal values between 200 and 400 ms, whereas Levinson and Torreira[2] (2015: 12) report that gaps "are most typically short, with modal values for FTO's falling between 100 and 200 ms". For overlaps, we found a larger mean (447 ms) and median (279 ms) than Levinson and Torreira (resp. 275 ms and 205 ms). In sum, this paints a picture in which the overall FTO is quite comparable in terms of median values, however, the mediated interaction in our study displays longer average overlaps and even more clearly so also larger gaps, compared to the non-mediated data analysed in Levinson & Torreira (2015) or Heldner and Edlund (2010). In other words, interpreter-mediated interactions appear to adhere less to the no gaps/no overlaps-principle proposed by Sacks et al. (1974) and thus seem to be less smooth than non-mediated monolingual interactions.

To try to account for the larger amount of noticeable overlaps in our data, we investigated whether overlaps can be explained as an interpreter's strategy to take the floor (for instance, if the speaker's turns are too long), as had been suggested in the literature (see Englund Dimitrova 1997, Licoppe & Veyrier 2017). Interestingly, our data reveal that noticeable overlaps tend to occur in the majority of cases with questions (66%) and less with non-questions (34%). This is mostly the case with questions that are extended beyond their completion point (74%). As an illustration, consider sequence 1 below:

(1) Turn extension beyond completion point

[2] Levinson and Torreira (2015) do not report mean and median values for gaps.

1	CNS	**want wat is euh de einddatum die je**
		because what is uh the final date that you
	int	--gaze at Cns ------------------------------->
2	CNS	**momenteel voor ogen hebt?**
		are currently having in mind?
	int	---gaze at Stu-------------------->
3		**(0.1)**
4	CNS	**[Of voor ogen had?**
		Or you had in mind?
5	INT	**[uh да к какому**
		Uh yes, what
	int	------------gaze at Cns-->
6		**(0.3)**
7	INT	**Да, к какому временни вы должны**
		Yes, what time should you
	int	---gaze away------------------>

Turn-extension

As the counsellor approaches a point in her turn that can be seen as complete (‘because what is uh the final date that you are currently having in mind?’) the interpreter starts shifting her gaze towards the student (‘gaze at Stu’), thus anticipating the end of the question and displaying her readiness to take the floor. After a slight pause, we see that the interpreter starts rendering the turn (line 5). However, as it happens, the counsellor produces a turn-extension “Or you had in mind?”, which results in overlapping talk. The overlap duration is 1491 ms. This occurrence of overlap is resolved quickly, as the interpreter suddenly cuts off her turn in progress and looks at the counsellor. As the counsellor makes no further attempt to speak, we see that the interpreter recycles the preceding utterance (“yes, what”) and continues with her turn. Thus, overlap does not result from the interpreter’s strategy to interrupt or cut off the counsellor’s turn in order to start rendering the talk, but rather from the interpreter’s early anticipation of turn completion.

By way of comparison, we will present here an example of ‘chunking’, which is defined in the literature as a “practice aiming to split either party’s input into manageable processing units, mostly with a view to keeping the flow of the dialogue” (Davitti 2018). In our data, ‘chunking’ is found in non-questions and it is used by the interpreter to segment the primary speaker's talk into shorter turn units or to stop the current turn from becoming too long. An example is provided below:

(2) Chunking

1	CNS	**Ja, ik kan dat begrijp[en, want** *yes, I can understand that, because*
2	INT	**[Я e:e** *I euh*
	int	---gaze at Cns----------gaze at Stu--->
3	INT	**могу понять,** *can understand,*
4	CNS	**jij woont hier, [en en** *you live here, and and*
5	INT	**[Живёте [здесь,** *you live here*
	int	-----------------------------gaze at Cns-->
6	CNS	**[hoevaak per jaar of zo ga je dan terug** *how often do you go back*

In this excerpt, the counsellor reacts to the student's wish to learn Dutch. Although there is no indication on the counsellor's part that she is yielding the turn, the interpreter starts shifting her gaze away from the counsellor and takes the turn in overlap to render the counsellor's agreement ('Yes, I can understand that') to the student. The interpreter thus anticipates a possible transition relevance place (TRP) at the end of the counsellor's syntactic unit and takes the opportunity to initiate her rendition. The counsellor then continues with her turn (line 4) by motivating why learning Dutch is useful ('you live here'), which is again immediately rendered by the interpreter. In this way, the interpreter not only 'chunks' the counsellor's turn in short turn units, but she also promotes direct contact between the primary speakers. Note that the counsellor in this session displays a tendency to produce long turns, which may impact the interpreter's turn-taking speed. In contrast to the previous example, the interpreter uses overlap as a strategy to split up the turn in shorter interpreting units in order to interpret the turn piecemeal[3].

4.2. Factors explaining turn-taking speed

So far, we have presented results on the distribution of FTO's in interpreter-mediated interaction. In this section we try to account for some of the variation in FTO by exploring the explanatory power of both linguistic (i.c. speech rate, turn duration and lexical density) and non-linguistic behavioural features (i.c. gaze shifts) of the primary speaker. To test our hypotheses, we first look for individual effects of our independent variables and finally present a mixed effects regression model that is able to provide a more strict and thorough view on the interplay between fixed and random factors under scrutiny.

In our dataset we annotated all the questions (N=96) and a subset of randomly selected non-questions (N = 96). A t-test revealed that FTO's were significantly lower for questions (M = 81 ms) than for non-questions (M = 418 ms), $t(190) = 4.50$, $p < 0.001$, i.e. interpreters are significantly faster at starting their turn after a questions than after a non-question. Furthermore, a correlation test indicated that FTO's were positively ($r = 0.45$, $p < 0.001$) correlated with the turn duration. In other words, interpreters were slower in taking the turn after longer turns. Turn duration and turn type, however, did not appear to be independent. That is, the questions (M = 5.6 sec) were significantly shorter than the non-questions (M = 11.7 sec), $t(190) = 4.95$, $p < 0.001$. For that reason we reran our analyses on a subset of the data, i.e. we omitted all turns that were shorter than 10 seconds. This resulted in omitting 37 data points. In this subset there

[3] The discussion of the two excerpts serves primarily as a qualitative illustration of the different causes for overlaps in our dataset. A more in-depth, qualitative analyses of the causes of overlaps in DI would exceed the scope and the main purpose of the present paper.

no longer was a significant difference in turn duration between questions and non-questions, but crucially, there no longer was a correlation between FTO and turn duration. To be able to truly discriminate between the factors turn type and turn duration, we would need a larger data set with more instances of long questions. Within the present three hour corpus, already all the questions had been annotated, so an expansion of the data set in that respect was not possible. For the more complex regression analyses below, and to avoid issues of non-independence within our independent variables, we will continue our statistical testing with the subset of our data in which all turns by primary speakers are at least 10 seconds long.

For speech rate, we found a significant negative correlation ($r = -0.42$, $p < 0.001$) with FTO: the faster the speech rate in a primary speaker's turn, the faster the interpreter starts his/her turn. For the remaining para-verbal factors lexical density and word frequency[4] we did not find a correlation with FTO. Also for the non-verbal factor eye-gaze no significant effect was found: a t-test revealed that interpreters are not significantly faster (as hypothesized) in taking the turn when the primary speaker shifts his/her gaze towards the interpreter before turn-completion, compared to when such a shift does not occur[5]. To allow for a more direct comparison with the results in Stivers et al. (2009, 10588), who did find an effect of speaker gaze ("responses were delivered earlier if the speaker was looking at the recipient while the question was asked"), we also performed the analysis on the subset of questions in our data. A t-test, however, indicated that also for questions alone, a gaze shift by the primary speaker towards the interpreter does not significantly speed up the FTO. Analysing gaze shifts for questions and non-questions separately did yield a relevant observation: gaze shifting itself (more precisely, turn-final gaze shifts of the counsellor from the primary speaker towards the interpreter) occur less frequently

[4] This was regardless of the manner of calculating word frequency. For both the rank-based and raw frequency-based calculation (see section 3) there was no correlation.

[5] Note that interpreters in this study tend to look at the speaker while listening (see Vranjes 2018). In Belgium, interpreters are not instructed to avoid eye contact with primary speakers, but we acknowledge that this might be different in other parts of the world.

for questions than for non-questions. As is apparent from Table 2 below, such gaze shifts occur in less than one fourth of the cases, whereas for non-questions this is exactly half of the cases. A chi-square test indicated that the amount of gaze shifts during questions differed significantly from that amount during non-questions, $\chi^2(1, N=192) = 7.11$, p = 0.007.

Table 2. Overview of turn-final gaze shifts for questions and non-questions

	Questions	Non-questions
With gaze shift	16	32
Without gaze shift	80	64

To combine the individual observations sketched above in a multifactorial model, and to maximally account for idiosyncratic behaviour of the individual interpreters, we built a mixed effects model with interpreter as a random factor, and turn type, turn duration and speech rate as fixed factors. Those fixed factors were chosen because they emerged as relevant in our single-factor analyses (cf. supra) and because an AIC-based stepwise selection procedure indicated those were the relevant factors as well. In the stepwise selection procedure we explored all possible two-way interactions, but none were found. To assure that the assumptions under which our multilevel model can be conducted were met, we first checked for issues of heteroscedasticity. When plotting the residuals against the fitted values, a visual diagnostic confirmed a homoscedastic pattern. This was backed-up by running a non-constant variance test from the R-package *car* (Fox & Weisberg 2019), which turned out to yield a non-significant result, thus indicating the errors vary in a constant manner. Second, we tested for independence of the predictor variables in our model by calculating VIF-scores. We found that none of the factors exceeded a VIF-score of 1.3, which allows us to claim there are no issues of multicollinearity. Third, a Shapiro-Wilk test indicated there was a mild violation of the assumption that residuals should be normally distributed ($W = 0.98$, $p = 0.03$). A visual inspection of the corresponding Q-Q plot indeed showed a limited deviation of normality towards the tails, but because of the limited nature of the deviation and because all other assumption tests were met, we decided not to weigh or transform our data and perform our multilevel model on the original data.

Building on that model, we found that turn type (t = 3.05, p = 0.003) and speech rate (t = -3.77, p < 0.001) were good predictors and turn duration was on the verge of a significant predictor (t = 1.92, p = 0.057). In addition, no interactions between factors appeared to be significant. From our mixed effects model we can conclude that, notwithstanding the peculiarities of the individual interpreters, speech rate and turn type are good predictors for FTO's: interpreters are faster at starting their turns if they are interpreting a question (compared to a non-question) and if they are interpreting a primary turn that was delivered at a fast speech rate. Shorter FTO's also occurred when interpreters were tackling turns of shorter duration. Given the fact that turn duration was only marginally significant, caution should be taken not to overinterpret this particular finding.

5. Discussion

The present study sought to investigate the timing of turn-taking in dialogue interpreting using a data set of interpreted interactions that were recorded with mobile eye-trackers. The study advances previous research in three ways: first, we provided a systematic analysis of the precise timing of turn-taking in dialogue interpreting and compared it with the findings reported on the timing of turn-taking in same-language interactions. Second, we analysed features that could have an impact on the interpreter's turn-taking speed. And third, we investigated the role of gaze, which has been argued to contribute to smooth turn-taking in face-to-face interaction.

First of all, our analyses suggest that, although interpreter's median and mean turn-taking times are quite comparable to turn-taking times of participants in monolingual spontaneous interactions (Helder & Edlund 2010, Levinson & Torreira 2015), a closer examination of their exact distribution shows that turn transitions appear to be less smooth. Interpreters in our data generally appear to orient to the maxim 'one speaker at a time', but we found more and longer gaps in our corpus of interpreted interactions. Whereas longer gaps between turns may be experienced as awkward or may carry certain implications in some situations, we assume that they are more acceptable when an interpreter is taking the turn given the interpreter's specific role in the conversation. This, however, remains to be investigated further, but our data suggest that interpreter's turn-taking speed is guided by different mechanisms than those of the interlocutors in monolingual interactions. Moreover, we found less, though longer average overlaps in our dataset. A closer inspection of the data revealed that noticeable overlaps in most cases occur when the counsellor is asking a question. Importantly, in such cases, overlaps usually result from the interpreter's *early anticipation* of turn completion, rather than from a turn-taking strategy (such as 'chunking'). Overall, it appears that dialogue interpreters in our corpus tend to avoid noticeable overlaps with the current speaker.

Our analysis of the factors impacting the timing of turn-taking revealed that interpreter's turn-taking speed is affected by turn type: interpreters are significantly faster at starting their turn after questions than after non-questions. This suggests that interpreters are able to *anticipate* faster when the primary speaker's turn is going to end when listening to questions, which is line with previous research showing that questions increase predictability and lead to reduced FTO's (Stivers et al. 2009, Roberts et al. 2015, see also Casillas & Frank 2013).

Perhaps unexpectedly, faster speech rate did not lead to longer transitions times. The results reveal that faster speech rate is followed by faster turn-taking rate, suggesting that interpreters adapt their turn-taking speed to the rhythm of the speaker's speech. This is in contrast to recent corpus research on turn-taking times in monolingual interactions that found that faster speech rate leads to longer FTO's (Roberts et al. 2015). Our finding suggests that interpreters work towards *maintaining the conversational flow,* which points to the interactional role of the interpreter.

We expected that turn duration would have an impact on the interpreter's turn taking speed. We found some support in the data for the hypothesis that longer turns do increase turn transition times, which suggests that, when talking with the aid of an interpreter, shorter turns will reduce gaps between primary speaker's and interpreter's turns and thus increase the flow of the conversation. However, the results were on the verge of significance and more data will be needed to obtain more conclusive results. Furthermore, neither lexical density nor word frequency were found to affect the interpreter's turn-taking speed. Although they do not have an effect on interpreter's turn-taking times, it remains to be investigated in future studies if they will have an effect on other indicators of cognitive load (such as disfluencies and reformulations a.o.). Overall, our findings indicate that dialogue interpreter's turn-taking times are somewhat different from what is to be expected in same-language conversation and that timing cannot be attributed to processing factors (such as turn duration and turn type) alone but also to other factors (such as rhythm of the talk).

In addition, we analysed the role of speaker's gaze in the timing of interpreter's turn-taking. It has been argued that speaker's gaze is a powerful turn yielding signal (see Kendon 1967, Bavelas et al. 2002) that elicits faster responses from the listener (Stivers et al. 2009). Moreover, since counsellors in our data gaze significantly more at the student than at the interpreter (Vranjes 2018), gaze shift towards the interpreter can be treated as a marked gaze event. Our results, however, reveal that current speaker's gaze does not affect the interpreter's turn-taking speed: interpreters in our data were not faster in taking the turn if the current speaker was looking in their direction before turn-completion. Interestingly, we found that counsellors display significantly less turn-final gaze shifts towards the interpreter in questions than in non-questions. This may be explained by the fact that questions inherently make speaker-change relevant and contain more cues for the listener to anticipate the end of the turn. Overall, although counsellor's gaze towards the interpreter at turn end may function as a turn-yielding signal, the interpreter's turn-taking speed appears to be independent from such visual cues.

As a final point, we need to emphasize the limitations of this study. Since the study is based on a limited dataset of nine interpreter-mediated interactions, all of which were recorded in one specific context, the data are not necessarily representative of the behaviour of all interpreters in certain settings and cannot be generalised. Also, the small corpus size affects statistical power and may be a reason for not finding significant results on the role of turn duration on interpreter's turn-taking speed. However, by focusing on one specific context, we were able to reduce the number of variables that could potentially have an impact on the interpreter's turn-taking times. This systematic approach enabled us to draw some conclusions regarding the timing of turn-taking in our specific dataset. Further, we focused only on the broad distinction between 'questions' and 'non-questions', without differentiating any further between different subtypes of (non-)questions. An additional subclassification would have raduced the number of cases and consequently reduced the statistical power of our analysis. There is thus scope for even further granularity in the analysis, preferably in a larger corpus. Furthermore, the present study focused only on transitions between the counsellor and the interpreter, and not between the student and interpreter, thus excluding potential differences in turn-taking speed related to interpreter's linguistic competence (L1 and L2) (see Tiselius & Sneed 2020) or differences related to asymmetries in power relationship between the primary speakers. Finally, although the comparison of our data with a corpus of monolingual conversations is not ideal, we had no other point of comparison since there have been no other quantitative studies on the timing of turn-taking in dialogue interpreting. Despite the inevitable limitations, this study represents the first attempt at carrying out a systematic, quantitative analysis of interpreter's turn-taking speed in dialogic context. We hope that this research will give impetus for further empirical investigations and that it will provide an interesting starting point for comparative future analyses with datasets from different contexts of dialogue interpreting.

6. Conclusion

In this explorative study, we have used a corpus-analytic approach to investigate the timing of turn-taking in dialogue interpreting. On the one hand, our analysis has shown in what ways the interpreter's turn-taking speed differs from timing of turn-taking in same-language, spontaneous interaction. On the other hand, we have identified specific factors that appear to have an impact on this timing both on the level of input processing and conversational organisation. The present study also has implications for research on turn-taking in dialogue interpreting, which has so far been primarily qualitative in nature and has usually focused on turn-taking as part of the interpreter's coordinative role in the exchange (Wadensjö 1998, Davidson 2002). As such, our study contributes to recent research efforts aiming to uncover specific aspects of processing involved in dialogue interpreting (Englund Dimitrova & Tiselius 2016), especially in comparison to simultaneous interpreting that has been the object of more systematic scrutiny in the past. We have found that anticipation is an important mechanism in dialogue interpreting, although it has to be understood differently from what has been suggested in earlier interpreting research. Also, although gaze plays an important role in monitoring the progression of the turn and in allocating turns (see Davitti 2012, Mason 2012, Vranjes et al. 2018a), our data show that speaker's turn-final gaze does not affect the interpreter's turn-taking speed. Further, we have also shown that interpreters not only organize their turn-taking according to their own processing requirements, but also by taking into consideration the expectations with regard to turn-taking in spontaneous interaction and aspects such as the 'rhythm' of the conversation. Together, these observations illustrate the complexity of the dialogue interpreting process, that does not only involve comprehension, memorisation and conversion of the source language into a target language, but also a high level of awareness of and compliance with the 'unwritten' rules of conversational organisation in order to enable smooth conversational flow. Finally, uncovering the distinct processes involved in dialogue interpreting may not only help inform the interpreting practice, but also enhance the position of dialogue interpreting as an object of systematic academic scrutiny and as a professional practice.

References

Auer, Peter. 2005. "Projection in Interaction and Projection in Grammar." *Interdisciplinary Journal for the Study of Discourse* 25 (1): 7–36.

Balota, David A., Melvin J. Yap, Michael J. Cortese, Keith A. Hutchinson, Brett Kessler, Bjorn Loftis, James H. Neely, Douglas L. Nelson, Greg B. Simpson, and Rebecca Treiman. 2007. "The English Lexicon Project." *Behavioral Research Methods* 39: 445-459.

Bavelas, Janet B, Linda Coates, and Trudy Johnson. 2002. "Listener Responses as a Collaborative Process: The Role of Eye Gaze." *Journal of Communication* 52 (September): 566–80.

Bates, Douglas, Martin Maechler, Ben Bolker and Steven Walker. 2014. lme4: Linear mixed-effects models using Eigen and S4.

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

Brady, Paul T. 1968. "A Statistical Analysis of On-Off Patterns in 16 Conversations." *Bell Labs Technical Journal* 47 (1): 73-91. https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1538-7305.1968.tb00031.x

Braun, Sabine. 2013. "Keep Your Distance? Remote Interpreting in Legal Proceedings: A Critical Assessment of a Growing Practice." *Interpreting* 15 (2): 200–228.

Brône, Geert, Bert Oben, Annelies Jehoul, Jelena Vranjes, and Kurt Feyaerts. 2017. "Eye Gaze and Viewpoint in Multimodal Interaction Management." *Cognitive Linguistics* 28 (3): 448-483.

Casillas, Marisa, and Michael C. Frank. 2013. "The Development of Predictive Processes in Children's Discourse Understanding." *Proceedings of the 35th Annual Meeting of the Cognitive Science Society,* 299–304. http://langcog.stanford.edu/papers/CF-cogsci2013.pdf.

Couper-Kuhlen, Elizabeth. 1993. *English speech rhythm: Form and function in everyday verbal interaction*. Amsterdam: John Benjamins.

Davidson, Brad. 2002. "A Model for the Construction of Conversational Common Ground in Interpreted Discourse." *Journal of Pragmatics* 34 (9): 1273–1300.

Davitti, Elena. 2013. "Dialogue Interpreting as Intercultural Mediation: Interpreter's Use of Upgrading Moves in Parent-Teacher Meetings." *Interpreting* 15 (2): 168–99.

Davitti, Elena. 2018. "Methodological explorations of interpreter-mediated interaction: novel insights from multimodal analysis." *Qualitative Research. Special Issue: Multimodality: Methodological Explorations* 19 (1): 7-29.

De Boe, Esther. 2020. *Remote Interpreting in Healthcare Settings: a comparative study on the influence of telephone and video link use on the quality of interpreter-mediated communication*. PhD dissertation. University of Antwerpen.

de Ruiter, Jan-Peter, Holger Mitterer, and Nick J. Enfield. 2006. "Projecting the End of a Speaker's Turn: A Cognitive Cornerstone of Conversation." *Language* 82 (3): 515-535.

Edlund, Jens, Mattias Heldner, and Julia Hirschberg. 2009. "Pause and Gap Length in Face-to-Face Interaction". *Proceedings Interspeech: 2779-2782.*

Englund Dimitrova, Birgitta. 1997. "Degree of Interpreter Responsibility in the Interaction Process in Community Interpreting." In *Critical Link: Interpreters in the Community*, edited by S.E. Carr, 147–64. Amsterdam/Philadelphia: Benjamins.

Englund Dimitrova, Birgitta, and Elisabet Tiselius. 2016. ”Cognitive aspects of community interpreting. Toward a process model” In *Reembedding Translation Process Research*, edited by Martín R. Muñoz, 195-214. Amsterdam: Benjamins.

Ford, C. E., and Sandra A Thompson. 1996. “Interactional Units in Conversation: Syntactic, Intonational, and Pragmatic Resources for the Management of Turns.” *Interaction and Grammar* 1: 134–84.

Fox J, Weisberg S (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks CA.

Gavioli, Laura. 2015. Conversation Analysis. In *Researching Translation and Interpreting*, edited by Claudia Angelelli, Brian James Baer, 185-194. London: Routledge

Gavioli, Laura, and Cecilia, Wadensjö. 2020. “Reflections on Doctor Question – Patient Answer Sequences and on Lay Perceptions of Close Translation.” *Health Communication:* 1-11.

Gile, Daniel. 1997. “Conference Interpreting as a Cognitive Management Problem.” In *Cognitive Processes in Translation and Interpreting*, edited by Joseph H. Danks, Gregory M.

Shreve, Stephen B. Fountain, and Michael McBeath, 196–214. Thousand Oaks, CA: SAGE Publications.

Hayashi, Makoto. 2013. Turn Allocation and Turn Sharing. In *The Handbook of Conversation Analysis*, edited by Jack Sidnell and Tanya Stivers, 167-190. Oxford: Wiley Blackwell.

Heldner, Mattias, and Jens Edlund. 2010. “Pauses, Gaps and Overlaps in Conversations.” *Journal of Phonetics* 38 (4): 555–68.

Heritage, John. 1997. Conversation analysis and institutional talk: analysing data. In *Qualitative Research: Theory, Method and Practice,* edited by David Silverman, 161-182. London: Sage Publications.

Holler, Judith, and Kobin H. Kendrick. 2015. "Unaddressed Participant's Gaze in Multi-Person Interaction: Optimizing Recipiency." *Frontiers in Psychology* 6 (1919): 98.

Holler, Judith, Kobin H. Kendrick, Marisa Casillas, and Stephen C. Levinson. 2015. "Editorial: Turn-Taking in Human Communicative Interaction." *Frontiers in Psychology* 6= 1919.

Kendon, Adam. 1967. "Some Functions of Gaze-Direction in Social Interaction." *Acta Psychologica* 26: 22–63.

Kendrick, Kobin H., and Francisco Torreira. 2015. "The Timing and Construction of Preference: A Quantitative Study." *Discourse Processes* 52 (4): 255-289.

Komter, Martha. 2013. "Conversation analysis in the courtroom." In *Handbook of Conversation Analysis,* edited by Jack Sidnell and Tanya Stivers, 612-629. Chichester, West Sussex, UK: Wiley-Blackwell.

Lammertink, Imme, Marisa Casillas, Titia Benders, Brechtje Post, and Paula Fikkert . 2015. Dutch and English toddlers' use of linguistic cues in predicting upcoming turn transitions. *Frontiers in Psychology* 6: 495.

Levinson, Stephen C. (1983). *Pragmatics.* Cambridge: Cambridge University Press.

Levinson, Stephen C., and Francisco Torreira. 2015. "Timing in Turn-Taking and Its Implications for Processing Models of Language." *Frontiers in Psychology* 6: 1–17.

Licoppe, Christian, and Clair-Antoine Veyrier. 2020. “The interpreter as a sequential coordinator in courtroom interaction. “Chunking” and the management of turn shifts in extended answers in consecutively interpreted asylum hearings with remote participants.” *Interpreting* 22 (1): 56-86.

Mason, Ian. 2012. “Gaze, positioning and identity in interpreter-mediated dialogues.” In *Coordinating Participation in Dialogue Interpreting,* edited by Laura Baraldi & Claudia Gavioli, 177-199. Amsterdam & Philadelphia: Benjamins Publishing.

Oostdijk, Nelleke. 2000. “Het Corpus Gesproken Nederlands.” *Nederlandse Taalkunde* 3: 280-284.

Pasquandrea, Sergio. 2011. “Managing Multiple Actions through Multimodality: Doctors’ Involvement in Interpreter-Mediated Interactions.” *Language in Society* 40 (4): 455–81.

Plevoets, Koen, Bart Defrancq. 2016. “The effect of informational load on disfluencies in interpreting: A corpus-based regression analysis.” *Translation and Interpreting Studies* 11 (2): 202–224.

Plevoets, Koen, Bart, Defrancq. 2018. “The cognitive load of interpreters in the European Parliament: A corpus-based study of predictors for the disfluency uh(m).” *Interpreting* 20 (1): 1–29.

Rogers, Shane, L., Craig, P., Speelman, Oliver, Guidetti, and Melissa, Longmuir. 2018. “Using dual eye tracking to uncover personal gaze patterns during social interaction.” *Scientific Reports* 8 **(**4271): 1-8.

Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson. 1974. “A Simplest Systematics for the Organization of Turn-Taking for Conversation.” *Language* 50 (4): 696-735.

Seeber, Kilian. 2011. “Cognitive Load in Simultaneous Interpreting: Existing Theories – New Models.” *Interpreting* 13 (2): 176–204.

Selting, Margret. 1996. “On the interplay of syntax and prosody in the constitution of turn constructional units and turns in conversation.” *Pragmatics* 6 : 371 – 388 .

Setton, Robin. 1999. *Simultaneous interpretation: A cognitive-pragmatic analysis.* Philadelphia: John Benjamins.

Stivers, Tanya, and Nick J. Enfield. 2010. “A Coding Scheme for Question-Response Sequences in Conversation.” *Journal of Pragmatics* 42 (10): 2620–26.

Stivers, Tanya, Nick J. Enfield, Penelope Brown, Christina Englert, Makoto Hayashi, Trine Heinemann, Gertie Hoymann, et al. 2009. “Universals and Cultural Variation in Turn-Taking in Conversation.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (26): 10587–92.

Tiselius, Elisabet, and Kayle, Sneed. 2020. “Gaze and eye movement in dialogue interpreting: An eye-tracking study.” *Bilingualism: Language and Cognition* 23 (4): 780-787.

Van Besien, Fred. 1999. “Anticipation in Simultaneous Interpretation.” *Meta* 44 (2): 250–9.

Vranjes, Jelena. 2018. *On the role of eye gaze in the coordination of interpreter-mediated interactions an eye-tracking study.* PhD diss. KU Leuven.

Vranjes, Jelena, Geert Brone, and Kurt Feyaerts. 2018. “On the Role of Gaze in the Organization of Turn-Taking and Sequence Organization in Interpreter-Mediated Dialogue.” *Language and Dialogue* 8 (3): 439–67.

Vranjes, Jelena, Hanneke Bot, Kurt Feyaerts, and Geert Brône. 2018b. “Displaying Recipiency in an Interpreter-Mediated Dialogue : An Eye-Tracking Study.” In *Eye-Tracking in Interaction : Studies on the Role of Eye Gaze in Dialogue*, edited by Geert Brône and Bert Oben, 10:303–24. Amsterdam: John Benjamins.

Vranjes, Jelena, and Geert Brône. 2020. “Eye-Tracking in Interpreter-Mediated Talk : From Research to Practice.” In *Linking up with Video : Perspectives on Interpreting Practice and Research*, edited by Heidi Salaets and Geert Brône, 149:203–33. John Benjamins.

Vranjes, Jelena, and Hanneke Bot. 2021. “A Multimodal Analysis of Turn-Taking in Interpreter-Mediated Psychotherapy.” *Translation & Interpreting. The International Journal of Translation and Interpreting* 13 (1): 101–17.

Wadensjö, Cecilia. 1998. *Interpreting as Interaction*. London and New York: Longman.

Wadensjö, Cecilia. 2018. “Involvement, trust and topic control in interpreter-mediated healthcare encounters.” *Communication & Medicine* 15 (2): 165-176.

Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann, and Hans Sloetjes. 2006. “ELAN: a Professional Framework for Multimodality Research.” In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*.