The Interpreters' Newsletter Special Issue 2025 (30), 85-124
DOI: 10.13137/2421-714X/37797 ISSN 1591-4127 (print) ISSN 2421-714X (online)
https://www.openstarts.units.it/dspace/handle/10077/2119

Simultaneous interpreting (with text): impact of condition, speech difficulty, speech rate and visual attention on young professional interpreters' performance

Isabelle S. Robert
Esther de Boe
Priya Saxena
University of Antwerp

Abstract

In simultaneous interpreting (SI) with text (SIMTXT), interpreters receive input through auditory (via the speaker) and visual (via the transcript of the speech) channels, rendering the activity multimodal. Whereas some researchers indicate that transcript use may facilitate performance, others claim it may add a layer of complexity to the intricate cognitive processes involved in SI. To fill this research gap, we conducted a quasi-experimental within-subject design study involving 12 English-Dutch junior interpreters to investigate the impact on performance of the condition (SI versus SIMTXT), the speech difficulty, the speech rate and the visual attention allocation to the transcript. The study combined data from eye-tracking glasses (visual attention) and audio recordings (performance). Performance was measured via two methods: Overall Quality Indicator and Problem Trigger Accuracy Rate. Results from a Linear Mixed Model analysis show that interpreting performance was superior in the SIMTXT condition. In addition, high speech rate and speech difficulty negatively affected performance in both conditions, with the extent of the effect depending on the method of measurement. Finally, in SIMTXT, a positive – be it limited – effect of visual attention to the transcript on performance was established, although a great variability was observed between the participants.

Keywords

Simultaneous interpreting, SIMTXT, eye tracking, interpreting performance, visual attention

Introduction

Simultaneous interpreting (SI) is a complex cognitive activity which requires language-comprehension tasks and language-production tasks to be carried out more or less concurrently. At international meetings (see, e.g. Seeber/Delgado Luchner 2020) and on the private market (Ivanov *et al*. 2014), interpreters are often provided with the written transcript of the speech, be it on paper or in electronic form. This type of interpreting is called SI with text or SIMTXT. As in SI, the input in SIMTXT is multimodal, that is, both auditory (speaker's voice) and visual (speaker's appearance/paralanguage *and* transcript). On the one hand, this may add to interpreters' cognitive load, but on the other hand, having access to a transcript may facilitate interpreters' work, since it can support their memory effort, especially for so-called problem triggers, that is, elements that may increase interpreters' processing capacity requirements such as speech rate, proper names, numbers, acronyms and other low-redundancy speech segments (Gile 2009, 2020; Mankauskienė 2016).

So far, results from the scarce SIMTXT-related conceptual and empirical research remain inconclusive (Seeber 2017; Ma/Cheung 2020). Conceptual models tend to suggest an increased effort or cognitive load (CL). Gile's effort model (2009, 2020) describes four efforts (listening and analysis, memory, production and coordination). When using a transcript of the speech, a fifth effort is added (reading), which may add up to CL. However, Gile (2009, 2020) also mentions a possible facilitating effect of transcript use, when it supports the interpreter in providing information that was not retrieved from the auditory input, particularly problem triggers. Seeber's (2017) model of SIMTXT corroborates the assumption that adding a transcript may increase the amount of "task interference" and overall cognitive load, although no "reliable empirical data" substantiates this prediction (*Ibid*.: 474). Empirical studies on SIMTXT are indeed scarce and show inconsistent results. This exploratory study attempts to partially fill this gap by investigating the facilitating effect of a transcript as revealed in the interpreting performance.

The second gap we address relates to the methodology of past studies on the topic. So far, only a few studies (e.g. Chmiel *et al*. 2020; Zou *et al*. 2022) have used eye-tracking technology to accurately and objectively determine interpreters' visual attention to a transcript. In addition, while previous studies focused on one single problem trigger (e.g. numbers in Desmet *et al*. 2018; Chmiel *et al*. 2020), our study combines the effect of several problem triggers. Besides, to our knowledge, no study has looked into the interaction effects of speech rate and visual attention.

To fill these research gaps, we designed a quasi-experimental within-subject design study, in which 12 young professionals performed two interpreting tasks, one in SI and one in SIMTXT. Data were collected by means of audio recordings, eye tracking glasses, wristbands, post-task questionnaires and post-experiment interviews. In this paper, we will *only* report on the quantitative results, in particular concerning performance and visual attention, focusing on the impact of the main condition (SI versus SIMTXT) on interpreting performance. This is measured by: (1) Overall Quality Indicator (OQI) and; (2) Problem Trigger Accuracy Rate (PTAR), while manipulating two independent variables (IVs). These are speech rate and one particular aspect of speech difficulty, i.e. low-redundancy speech segments such as numbers, names and idioms, which, for the sake of clarity, we will refer to as “speech difficulty”. Additionally, we investigate the relationship between visual attention to the transcript in SIMTXT and interpreting performance. The data were quantitatively analysed, based on Linear Mixed Models (LMMs).

To investigate the effect of text in SI among professional interpreters in real-life conditions we focused on the following research questions:

- RQ1: To what extent does the interpreting condition (SI versus SIMTXT) impact on interpreting quality as measured via OQI and PTAR?
- RQ2: If one interpreting condition appears to generate better performance, do speech rate and speech difficulty modulate the comparative advantage of that condition over the other?
- RQ3: In SIMTXT, is there a relationship between visual attention to the transcript and interpreting quality as measured via OQI and PTAR?

This paper is structured as follows. Section 1 presents a review of relevant studies and concepts related to SIMTXT, including speech difficulty, speech rate and visual attention. Subsequently, the methodology is explained in Section 2. The results of the quantitative analysis are presented in Section 3. Finally, Section 4 discusses the main conclusions from the study and their potential impact on interpreting research and training.

1. Literature review

1.1 SIMTXT

As explained in the introduction, SIMTXT is increasingly common in current interpreting practice but academic research on this topic is still scarce and there is little agreement on the facilitating effect of SIMTXT (Seeber 2016; Seeber/Delgado Luchner 2020), which was confirmed by our review. Lambert (2004) conducted an experiment with 14 translation trainees performing three tasks: sight translation (ST), sight interpretation (SIT, which equals SIMTXT) and SI. Results show that both ST and SIT yielded significantly higher performance scores than SI. Besides, performance rates in ST and SIT were virtually identical. Ma/Cheung (2020) analysed an authentic corpus of 7 SIMTXT and 8 SI speeches interpreted by United Nations interpreters, finding evidence in support of better performance in SIMTXT, as manifested in less source-language interference. This was explained by Ma/Cheung (*Ibid*.: 452) by the fact that access to the transcript liberates interpreters’ cognitive resources for target-language production. Elsewhere, Yang *et al*. (2020) assessed the interpreting product of two groups of 27 trainees between SIMTXT and SI. They found that SIMTXT yielded significantly higher interpreting quality than SI, although participants in the SIMTXT condition had a significant longer ear-voice span (EVS)[1] com-

1 Ear-voice span’ (EVS) or time lag, also referred to by the French term *décalage*, refers to the delay between the speaker’s delivery and the interpreter’s output.

pared to the SI group. Finally, Zou *et al*. (2022) used fixed eye tracking (i.e. an eye tracker attached to a computer screen) in SIMTXT experiments with 9 professional interpreters to investigate the relationship between attention allocation and interpreting performance. Their analysis showed that interpreters who dedicated most visual attention to the text generated the highest interpreting quality.

Other studies on SIMTXT found a partially facilitating effect, or none. Lamberger-Felber (2001) had 12 professionals performing 3 tasks: SIMTXT without preparation, SIMTXT with preparation and SI. She found a facilitating effect of text use for names and numbers in SIMTXT. According to Lamberger-Felber (*Ibid*.: 50) these results neither confirmed nor disproved the hypothesis that SIMTXT facilitates interpreting, which may be due to the unexpectedly high variability between the participants. Nevertheless, most interpreters performed better when provided with a prepared manuscript, although a longer average time lag was found in SIMTXT. Yet, these results must be interpreted with care, given the "non-comparability of the speeches" (*Ibid*.: 45) and the lack of counterbalancing between the tasks. In a follow-up study based on the same data, this time with a focus on interference, Lamberger-Felber and Schneider (2008) found neither a facilitating nor a negative effect of the text on performance: all 12 interpreters produced interferences in the different conditions, albeit of considerably varying frequency, indicating that the condition had less impact on the frequency of interferences than individual interpreting strategies applied by different professional interpreters with comparable qualifications. Finally, Chmiel *et al*. (2020) conducted an experiment with 24 simultaneous interpreters, analysing accuracy of interpreted numbers, names and control words and visual attention. They found no evidence of a facilitating effect for congruent items (numbers or names identical in the oral and written text) but did identify an impeding effect on performance by the presence of the transcript for incongruent items, which were linked to the Colavita effect (visual stimuli taking precedence over auditory ones) and risk-avoiding strategic behaviour by the interpreters.

1.2 Speech difficulty and speech rate in SI and SIMTXT

Speech difficulty is directly related to the concept of input variables, which refer to content, form and delivery of the original speech (Mead 2015). In terms of content, an important variable is the information density of the original. As far as form is concerned, four series of features have been identified as relevant: (1) lexical density; (2) syntactic complexity; (3) terminology and; (4) low-redundancy speech segments such as names, acronyms and numbers, which are all potential problem triggers.

In empirical research on SI/SIMTXT, investigating the effect of input variables on performance is prominent and a number of studies on numbers as problem triggers have been carried out. For example, both Desmet *et al.* (2018) and Defrancq/Fantinuoli (2021) looked into the effect of using technological support, such as automatic speech recognition, on the accuracy of rendering numbers, finding a positive impact. Stachowiak-Szymczak/Korpal (2019) also focused on the accuracy of the rendition of numbers, finding that providing visual materials such as a slide presentation led to higher accuracy in number rendition, while Korpal/Stachowiak-Szymczak

(2020) combined numbers with high speech rate to observe that a high speech rate compromised the accuracy of number rendition.

Speech rate is indeed “a key input variable […] and a major determinant of how much information the interpreter has to process in a given time” (Mead 2015: 191). Yet, few studies investigate its impact on SI performance and results are inconsistent (Yang *et al*. 2020). Some studies found that high speech rates enhanced interpreting performance (Shlesinger 2003), while others such as Pio (2003) and Barghout *et al*. (2015) show a negative impact of high speech rate on performance due to increased omission or disfluency. Pio (2003) found that participants tended to produce more omissions during fast speech rate compared with slow speech rate, as well as a higher number of filled pauses and corrections during fast speech rate. Barghout *et al*. (2015) found that high and extremely high speech rate led interpreters to omit more redundant information. Similarly, Korpal (2017) showed that a high delivery rate increased the level of stress experienced by the participants and had a detrimental effect on accuracy. However, not all studies confirm a negative effect of speed. In the study by Liu *et al*. (2004), comparing professional interpreters and students with speech rate as one of the independent variables, the two groups responded similarly to increased speech rates. Chang (2005) found that the accuracy of renditions was undermined by a fast speech rate, but not the target language quality. Last, Han/Riazi (2017) found that high speech rate affected interpreting performance negatively for information completeness and target language quality but led to higher fluency of delivery.

It must be noted that in Interpreting Studies there is no consensus on an optimal speech rate for SI. According to Gile (2009), a speech rate between 100 words per minute (wpm) and 130 wpm is comfortable for interpreters. However, these rates are often exceeded in practice (Seeber 2015a), for example, in international conferences, the average speech rate is 140 wpm (Seeber/Delgado Luchner 2020). Fast input rates (>160 wpm) may pose a problem for trained interpreters when combined with high information density (Setton/Dawrant 2016: 323).

As far as SIMTXT is concerned, one of the few studies addressing speech rate is Yang *et al.* (2020), who concluded that use of a transcript had a significant effect on the quality of interpreting during high speech rates. By contrast, Robert *et al.* (2025) did not find a relationship between speech rate and performance in SIMTXT, but their study was a pilot case study. In Yang *et al.* (2020), no link between high speech rate and visual attention to the transcript was investigated, whereas Robert *et al.*’s (2025) case study indicated that a higher speech rate led to a greater visual attention to the transcript provided. Given this scarce empirical evidence, the present study further investigates the relationship between speech rate and performance in SIMTXT.

1.3 Visual attention

As explained in Section 1.1, there is little agreement on the extent to which visual attention to the transcript of the speech helps the interpreter. Besides, this was measured in only a few studies. Visual attention is usually analysed from the perspective of the Eye-Mind Hypothesis, which presupposes a close relationship between eye movements and their fixations, and what is being processed by the mind (Just/Carpenter

1980). Within this (limited) system of visual attention, relevant and irrelevant visual information is filtered to achieve efficient information processing (Carrasco 2011). Visual attention can be measured most objectively by means of eye tracking, that is, measurement of the position and movement of the human eye with tools called eye trackers (Seeber 2015b). Eye tracking is a harmless and non-invasive method, which sheds light on visual attention and cognitive processes in translation and interpreting (Korpal 2015; Stachowiak-Szymczak/Korpal 2019). It was used by Seeber *et al*. (2020) who compared SIMTXT with a reading-while-listening task. They found that in SIMTXT, participants showed a visual lag behind aural stimuli, which is different from reading while listening and concluded that visual input was not used to support comprehension, but to facilitate production of output, contrary to previous assumptions. So far, only few empirical studies have used eye tracking to explore the possible facilitating effect of SIMTXT compared with SI. Chmiel *et al*. (2020) and Zou *et al*. (2022) both used a screen with a fixed eye tracker, showing the transcript of the speech without visual input from the speaker. Yet, according to Chen (2017: 645), interpreters estimate the visibility of the speaker crucial in interpreting. Such a setting therefore lacks ecological validity (Stachowiak-Szymczak/Korpal 2019: 248).

2. Methodology

2.1 Research design, implementation and participants

To create an ecologically valid setting, we designed a simulated real-life quasi-experimental within-subject study combining quantitative and qualitative data collection. This paper reports on the quantitative part of the study and focuses on performance and visual attention only.[2] Ethical clearance was obtained through the Ethics Committee for the Social Sciences and Humanities of the University of Antwerp (SHW_2023_153_1). The quasi-experiments took place in 2023 in four interpreting booths at our university's campus. Via conference interpreters' networks we recruited 12 MA-trained conference interpreters with a maximum of 3 years of experience and working from English (B) into Dutch (A). Although some participants received training as conference interpreters, they were not all active interpreters (See Appendix 1). However, this was taken into account in the data analysis (see Section 3.2.1). Two days before the study, they were informed about the topic of each speech described in a few words. All interpreters signed an informed consent form and filled out a short demographic survey (Appendix 1). Before the first task, the interpreters were provided with eye tracking Tobii Pro Glasses 3. During the first task (SI), interpreters were allowed to take notes, which they did (e.g. numbers). After the first task, and a short break, they received the transcript of the second speech on paper (Calibri 14, line spacing 1,5, one-sided sheets). They were allowed 10 minutes to prepare the speech (mark the transcript and look up terms), after which the second task (SIMTXT) started, delivered by the same speaker. The interpreters were allowed to take notes and

2 Further qualitative and quantitative analyses are work in progress.

consult their transcript. It needs to be noted that for organisational and budgetary reasons, task order was not counterbalanced. We could only plan 3 sessions with 4 interpreters at the same time, whereas counterbalancing task order would have required 8 participants to start with Task 1 compared to 4 participants starting with Task 2 or vice versa. Task-speech combination was fixed for the same reason, i.e. Task 1 was always performed with Speech 1 and Task 2 with Speech 2. For further details, see Section 4.

2.2 Material

Since we worked with a within-group design, special attention was paid to the creation of two highly comparable source speeches. These were based on two existing speeches from the European Commission Speech repository.[3] These were manipulated, that is, similar numbers and problem triggers were added to create well-balanced materials (see below). The first speech dealt with the economic crisis in Greece in 2012; the second one with the demographic shift and pensions. We also made sure that the speeches were approximately the same length (1,500 words and 13 minutes) and divided each speech into 3 parts of approximately the same length (see Table 1), in order to deliver each part at a different speech rate: (1) 'slow' (110 wpm); (2) 'fast' (150 wpm) and; (3) 'moderate' (130 words wpm). These rates correspond to what Setton/Dawrant (2016) categorised as 'easy', 'average', and 'challenging'. Each part was again divided into two subparts, one containing 'few problem triggers' (PT) (N=7) and one containing 'many problem triggers' (N=13), in order to create an 'easy' and a 'difficult' subpart in each part of the speech. For maximum comparability of the speeches, we used several readability indexes (Flesh index Speech 1: 61; Speech 2: 60.6; Flesh-Kincaid index 8/8.5; Gunning-Gof: 10.5/ 10.9; Coleman Liau: 9.8/ 9.3).

Speech rate	Speech difficulty (based on the number of problem triggers – 'PT')	Subpart of the speech	Number of words Speech 1	Number of words Speech 2
Slow 110 wpm	Easy = Few PT (N = 7)	1	232	255
	Difficult = Many PT (N = 13)	2	238	260
Fast 150 wpm	Easy = Few PT (N = 7)	3	245	247
	Difficult = Many PT (N = 13)	4	247	256
Moderate 130 wpm	Easy = Few PT (N = 7)	5	255	267
	Difficult = Many PT (N = 13)	6	252	256

Table 1: Structure and features of each speech

3 <https://speech-repository.webcloud.ec.europa.eu/speech/greece-doldrums>; <https://speech-repository.webcloud.ec.europa.eu/speech/demographic-shift-europe>.

2.3 Data analysis

To measure quality of the interpreting performance, we first transcribed the 24 interpreting tasks verbatim (including disfluencies, such as filled pauses and incomplete phrases), based on the audio recordings. These transcriptions were analysed, and disfluencies and errors were first coded by one researcher-interpreter, co-author of the present paper, and then validated by a second researcher, co-author of the present paper, conference interpreter, PhD in interpreting studies and experienced interpreter trainer. Where the coding diverged, both researchers discussed the issue to reach a consensus, so no inter-rater reliability was calculated. The coding of the disfluencies was mainly based on Bóna/Bakti (2020) and on Shreve *et al*. (2010). The error coding was based on established quality assessment scales by Setton/Dawrant (2016). We distinguished between content-related errors (omissions, additions, inaccuracies), language-related errors (lexical and/or terminological errors, non-idiomatic formulations, grammatical errors, inappropriate register), and delivery-related errors (articulation, voice quality issues, and disfluencies). In addition, we attributed weight to the errors by distinguishing between 'minor' and 'major' errors. Weighing errors is common in quality assessments to account for interpreting strategies such as strategic omissions (Gieshoff/Albl-Mikasa 2022). Drawing on these categories, we designed 2 indicators for interpreting quality: 'overall quality indicator' (OQI) and 'problem trigger accuracy rate' (PTAR). OQI was calculated as follows: 1 point per minor error, 2 points for major errors (disfluencies excluded), and 0.25 point per disfluency. Consequently, a high OQI means a low performance. As to the PTAR, only the *number* of correct renderings was taken into account, so no weight was attributed. Here, a high PTAR means a high performance.

To answer the first research question, we analysed the impact of the condition, while to answer the second research question, we also analysed the impact of speech rate and speech difficulty. Finally, to answer the third research question, visual attention registered by means of eye tracking was measured through different metrics related to three areas of interest (AOI): (1) transcript (in SIMTXT); (2) speaker and; (3) notes taken by interpreters on notebook or paper. No additional AOIs for particular words in the transcript were created, given that drawing AOIs in Tobii Pro Lab is extremely time-consuming. The participants sat in a booth and had their notes and transcript in front of them on their desk, approximately at 40-70 centimetres while the speaker was seated at a distance of approximately 6 to 7 meters.

Eye tracking offers a wide variety of measures related to eye movement, position, and other eye activities (Holmqvist *et al.* 2015) but interpreting scholars have frequently used measures such as fixation duration, fixation count and gaze time (Stachowiak-Szymczak 2019) to investigate visual attention and/or CL. The metrics we used were generated in and downloaded from Tobii Pro Lab in Excel format ("Excel report", Tobii Pro Lab Product Description v 1.171.1, 2021:28): total fixation duration (TFD), average fixation duration (AFD), fixation count (FC), total visit[4] du-

4 In eye tracking, a visit corresponds to all data between the start of the first fixation inside an AOI until the end of the last fixation in the same AOI (including saccades, blinks or invalid gaze data) (Tobii 2021).

ration (TVD), average visit duration (AVD) and visit count (VC). As to data cleaning for eye movements, fixations shorter than 60 ms were filtered out by the Tobii Pro Lab default settings. As demonstrated in Section 3, metrics were not combined but were used separately as indicators of visual attention.

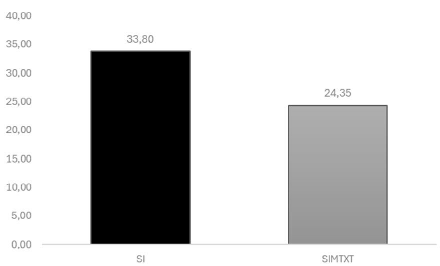
3. Results

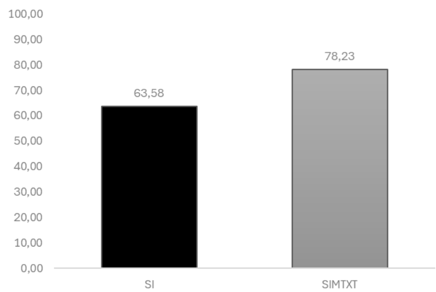
This section provides an overview of the results of the quantitative analyses comparing SI and SIMTXT by means of descriptive statistics (Section 3.1), and inferential statistics (Section 3.2).

3.1 Descriptive statistics

3.1.1 Interpreting performance

Figure 1 represents the scores for the first dependent variable, i.e. the OQI and shows that interpreters performed better in SIMTXT, since the score was lower. Figure 2 represents the second dependent variable (PTAR) and also shows that participants performed better in SIMTXT, since the PTAR is higher than in SI.





Figures 1 and 2: Interpreting performance indicators (left: OQI; right: PTAR, in percentages)

3.1.2 Visual attention

As illustrated in Figures 3 to 5, in SI, visual attention to the speaker based on *fixation* metrics is always higher than to the notes, whatever metrics are used (AFD, TFD and FC). In SIMTXT, visual attention changes drastically, with less attention paid to the notes and the speaker, in favour of the transcript, which receives most visual attention drawing on total fixation duration and fixation count, but not on *average* fixation duration. This aligns with Perego *et al.*’s (2010) study on subtitling processing. They observed that more fixation time was devoted to subtitle reading than to the visual analysis of film scenes, but that fixations on the latter were longer than on subtitles.

Besides, research in characteristics of fixation durations in scene viewing and reading shows that fixation durations in scene perception tend to have a longer average duration than in reading (Nuthmann/Henderson 2012).

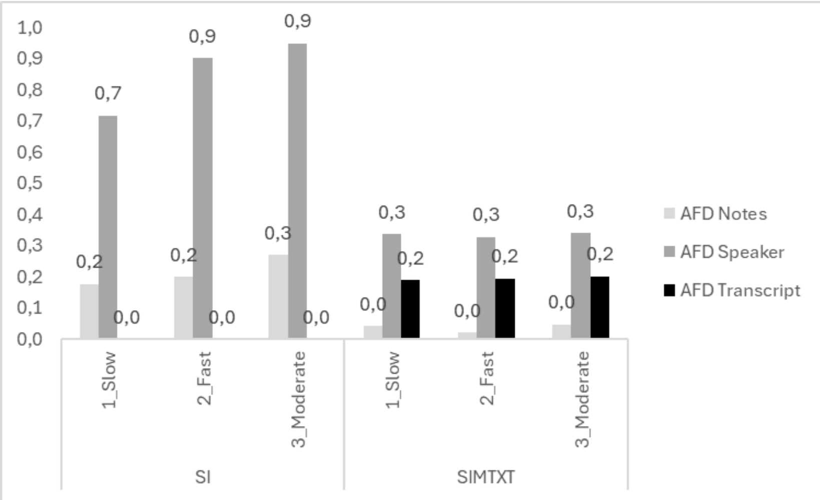


Figure 3: Average fixation duration (AFD, in seconds) in each AOI, in both conditions

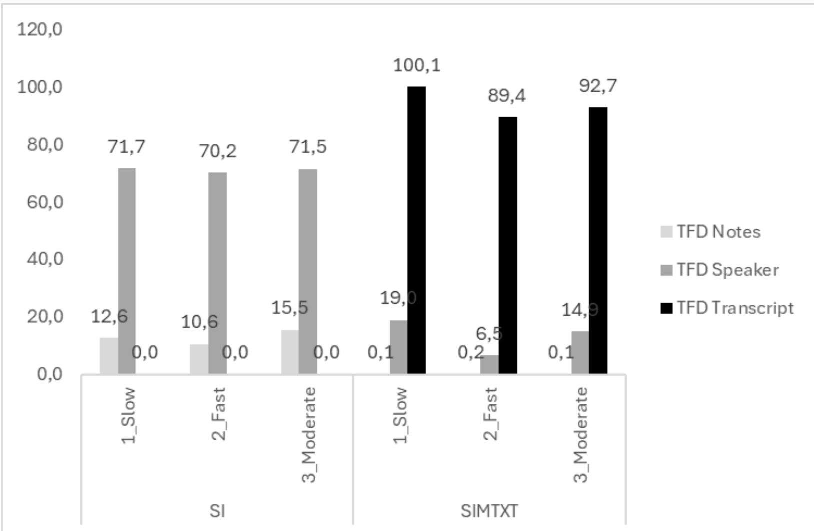


Figure 4: Total fixation duration (TFD, in seconds) in each AOI, in both conditions

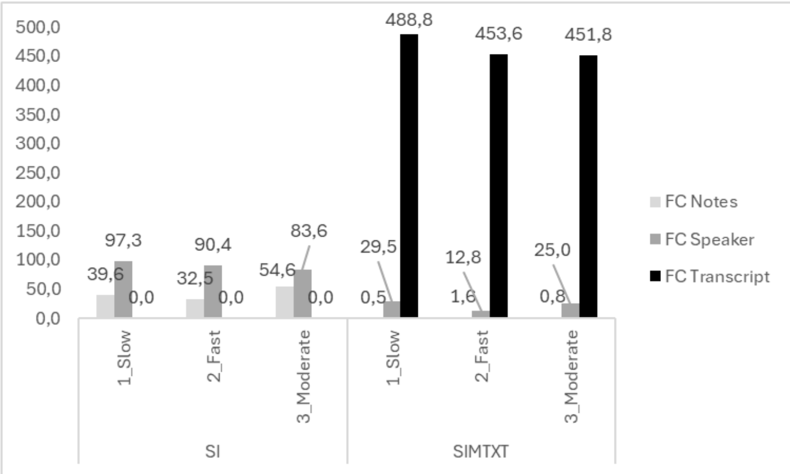


Figure 5: Fixation count (FC) in each AOI, in both conditions

Visual attention metrics based on *visits* show similar trends (Figures 6 to 8). In SI, there is more attention paid to the speaker than to the notes. In SIMTXT, visual attention to the notes and to the speaker decreases in favour of visual attention to the transcript, whatever the metrics (AVD, TVD and VC). These results are similar to those from previous studies investigating gaze behaviour in another type of visual verbal support in interpreting, i.e. subtitles (Li/Chmiel 2024), in which interpreters devoted more time on looking at the subtitles than at the speaker, which in turn confirms Seeber *et al*. (2020) on SIMTXT.

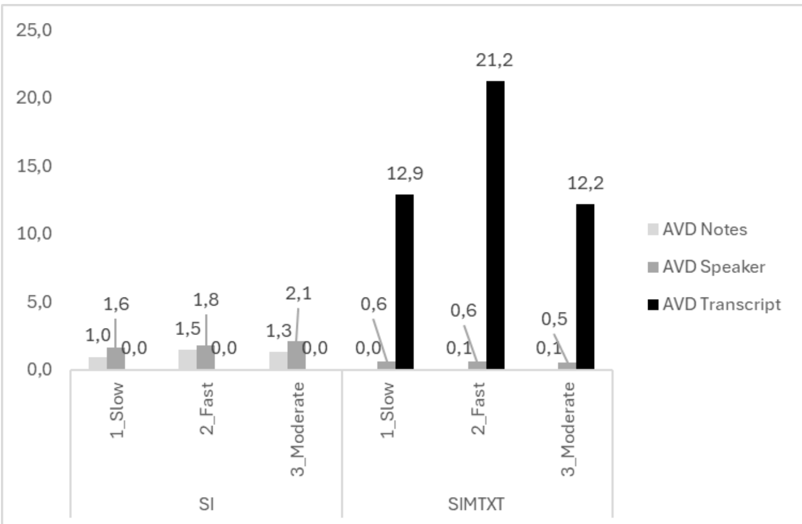


Figure 6: Average visit duration (AVD, in seconds) in each AOI, in both conditions

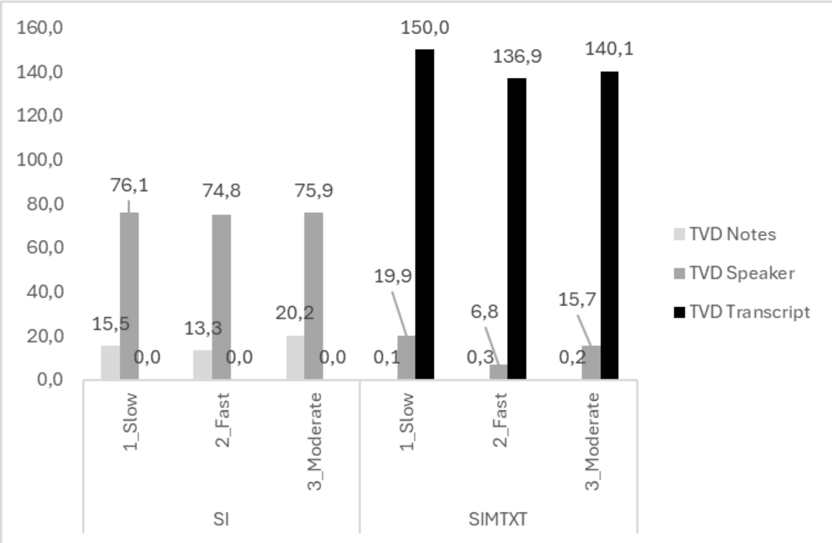


Figure 7: Total visit duration (TVD, in seconds) in each AOI, in both conditions

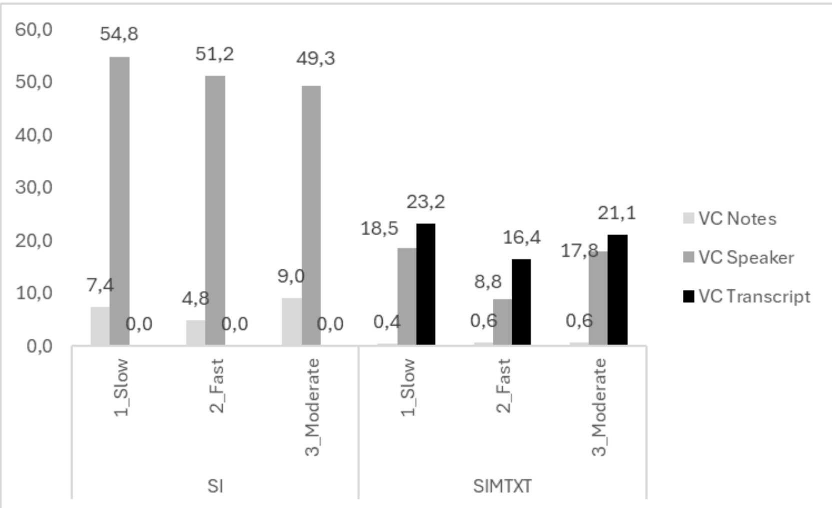


Figure 8: Visit count (VC) in each AOI, in both conditions

Participants behaved very differently as far as visual attention was concerned, confirming Robert *et al.*'s (2025) case study. This diversity in gaze behaviour may be related to the interpreter's style, as noted by Rennert (2008), who observed, for example, that some interpreters prefer to close their eyes during difficult parts (2008: 216).

3.2 Inferential statistics: Linear Mixed Models (LMM)

3.2.1 RQ1: Condition

To answer the first research question, we conducted a Linear Mixed Model Analysis in SPSS[5] version 29. We started with a first model which included the OQI as dependent variable (DV), the condition as predictor variable (fixed effect), and the participants[6] as random effect, since the same participants carried out both quasi-experiments. Furthermore, we included the speech as additional random effect. Although we attempted to design two very comparable speeches, the topic in each was different and, consequently, the potential effect of the speech was therefore included in the model.

In this first model (Appendix 2, Model A1: -2 Log=1070.28; 4 parameters), the condition appeared as a significant predictor of the DV ($F(1,24) = 13.20$, $p=0.001$). On average, Condition 1 (SI) has a mean outcome that is 9.45 units higher (95% CI: 4.08 – 14.82) compared to Condition 2 (SIMTXT), which means that the OQI is on average 9.45 higher in SI as compared to SIMTXT, showing that participants made significantly more errors in SI.

We conducted the same test with the PTAR as DV (Appendix 2, Model B1: -2 Log=1212.41; 4 parameters), which was again significant ($F(1,24) = 22.16$, $p<.001$). On average, Condition 1 (SI) has a mean outcome that is 14.65 units lower (95% CI: -21.08 – -8.23) compared to Condition 2 (SIMTXT), which means that the PTAR score in SI was 14.65 lower in SI than in SIMTXT and thus, that participants achieved a significantly higher accuracy rate when interpreting problem triggers in SIMTXT.

3.2.2 RQ2: Speech rate and speech difficulty

To answer the second research question, the next step was to add the two predictor variables as additional fixed effects and determine: (1) whether this improved the model; (2) if so, whether the additional variable was a significant predictor of the DV (performance) and; (3) if so, whether there was an interaction effect between the additional variables and the effect of the main predictor (condition).

Following Field (2009: 737), to determine whether a new model is better than the previous one, we looked into the difference in -2 Log Likelihood of each model and, depending on the difference in the number of parameters included in the two models (e.g. 7 parameters compared to 5 means a difference of 2 parameters), we determined whether the difference in -2 Log Likelihood was high enough to be significant. To do so, we consulted Field (2009: 808, Table A.4). For example, if two models differ by one parameter, a difference in -2 Log Likelihood of 3.84 is necessary to be significant at $p=.05$. In addition, we also looked at the adjusted versions of the log-likelihood, in particular the AICC, corrected for model complexity and designed for small samples (Field 2009: 737).

5 In SPSS 29, degrees of freedom are estimated using Satterthwaite's approximation.
6 To compensate the heterogeneity of the participants, we also included the number of years of experience in SI as additional fixed effect, but it was never significant in any of the models reported here (See Appendix 2).

When considering OQI as DV, the best model (Appendix 2, Model A2: -2 Log=1029.15; 9 parameters) consisted of the following significant fixed effects: (1) condition, and (2) interaction between speech rate and speech difficulty.[7] The random effects were the same as in the previous section (participants and speeches). The model dimensions as well as the tests of fixed effects and the estimates of fixed effects can be found in the tables in Appendix 2, Model A2. Results show that the condition remains a significant predictor of performance (F(1,24) =13.20, *p*=0.001), with a significant increase of the OQI in SI (on average 9.45 units higher, 95% CI: 4.08 – 14.82) compared with SIMTXT, thus a significant better performance in SIMTXT. However, the interaction effect between speech rate and difficulty level is significant and must be taken into account in the prediction (F(2,120)=8.96, *p*<0.001). Independent of the condition, there is a significant decrease in the number of errors in the 'slow speech rate and easy speech' subpart. On average, the OQI is 2.27 units[8] lower compared with the 'moderate speech rate and difficult speech' subpart (which is the reference by default in SPSS), whereas there is a significant increase in the 'fast speech rate and difficult speech' subpart (on average, the OQI is 11.73[9] units higher). There is also a significant increase in the three other subparts compared with the reference, but this increase is rather similar and less sharp (resp. 5.84, 5.22, 4.10 units higher[10] than the reference by default), as shown in Figures 9 and 10. As explained previously, this holds true for both SI and SIMTXT since there is no interaction with the main predictor variable (condition).

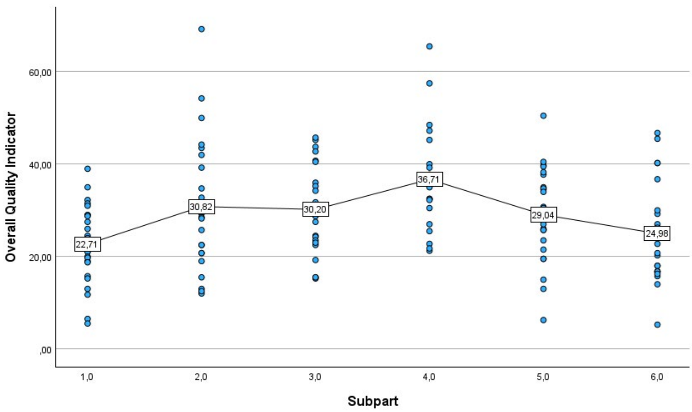


Figure 9: Mean OQI in each subpart, irrespective of the condition

7 When interaction between two effects is significant, we only mention interaction.

8 Effects are not simply additive. The effect size of the interaction is calculated as follows: sum of the estimate of [Speech Rate=1], [Difficulty_nb_RP=1] *and* [Speech Rate=1]*[Difficulty_nb_RP =1] from Table 4 of Appendix 2, Model A2 (marked in grey), that is 5,84 (95% CI: 1.47 – 10.22) + 4.06 (95% CI: -0.32 – 8.44) + -12.18 (95% CI: -18.37 – -5.99).

9 Effect size: 11.73 (95% CI: 7.35 – 16.11) + 0 + 0 (see Table 4 of Appendix 2, Model A2).

10 Effect size, respectively: 5.84 (95% CI: 1.47 – 10.22) + 0 + 0; 11.73 (95% CI: 7.35 – 16.11) + 4.06 (95% CI: -0.32 – 8.44) + -10.57 (95% CI: -16.76 – -4.38); and 4.06 (95% CI: -0.32 – 8.44) + 0 + 0 (see Table 4 of Appendix 2, Model A2).

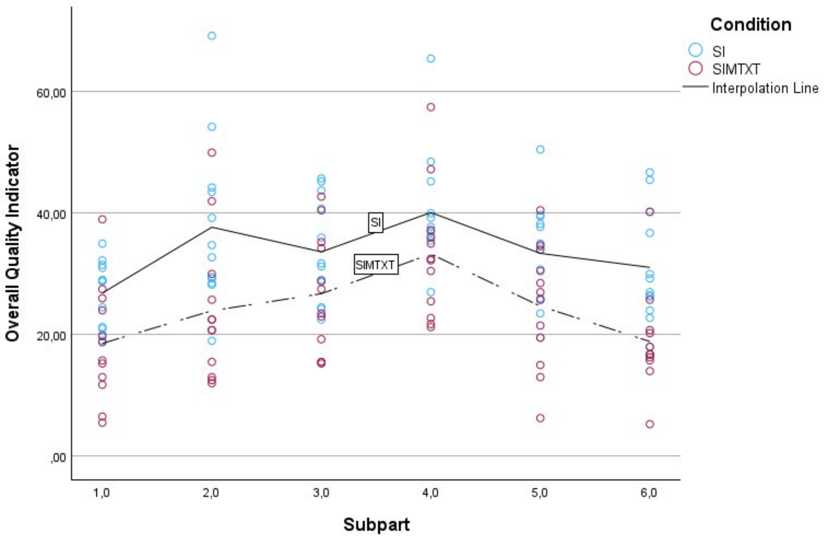


Figure 10: Mean OQI in each subpart, per condition

As observed in Section 3.2.1, the main condition was a significant predictor of performance measured via the PTAR too. The best model (Appendix 2, Model B2; -Log=1156.43; 9 parameters) consisted of 2 significant effects: speech difficulty ($F(1,120)$ =10.34, p=0.002) and interaction between condition and speech rate ($F(2,120)$ =5.37, p=0.006). In other words, contrary to the best model related to the OQI, there is no interaction effect between speech rate and speech difficulty, but there is an interaction between the condition and speech rate as shown in Figure 11. A decrease in performance occurred in the difficult part of the speech (6.73 units lower, 95% CI: 2.59 – 10.87; see Table 4, Model B2, Appendix 2). In addition, we observed a decrease in performance in the fast part of the speech, which was more pronounced in SI than in SIMTXT. Compared with the moderate speech rate part of the task in SIMTXT (reference by default in SPSS), the PTAR is on average 28.79 units[11] lower in the fast speech rate part in SI, whereas it is on average only 8.44 units lower[12] in the fast speech rate part in SIMTXT. There is little difference between the moderate speech rate part in SIMTXT (reference by default in SPSS) and the slow speech rate parts in both conditions (-2.95 units in SI and 2.06 units in SIMTXT compared to the reference),[13] but the difference is again high in the moderate part of the speech in SI (-18.60 units[14] compared to the same speech rate in SIMTXT).

11 Effect size: -18.60 (95% CI: -27.16 – -10.04) + -8.44 (95% CI: -15.62 – -1.27) + -1.75 (95% CI: -11.90 – 8.40) (see Table 4, Appendix 2, Model B2).
12 Effect size: -8.44 (95% CI: -15.62 – -1.27) + 0 + 0.
13 Effect size, respectively: -18.60 (95% CI: -27.16 – -10.04) + 2.06 (95% CI: -5.12 – 9.23) + 13.59 (95% CI: 3.44 – 23.74); and 2.06 (95% CI: -5.12 – 9.23) + 0 + 0 (see Table 4, Appendix 2, Model B2).
14 Effect size: -18.60 (95% CI: -27.16 – -10.04) + 0 + 0 (see Table 4, Appendix 2, Model B2).

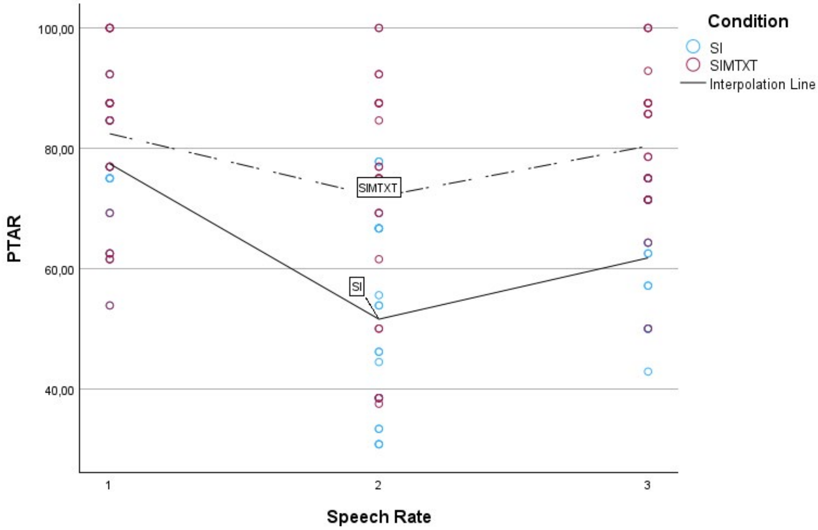


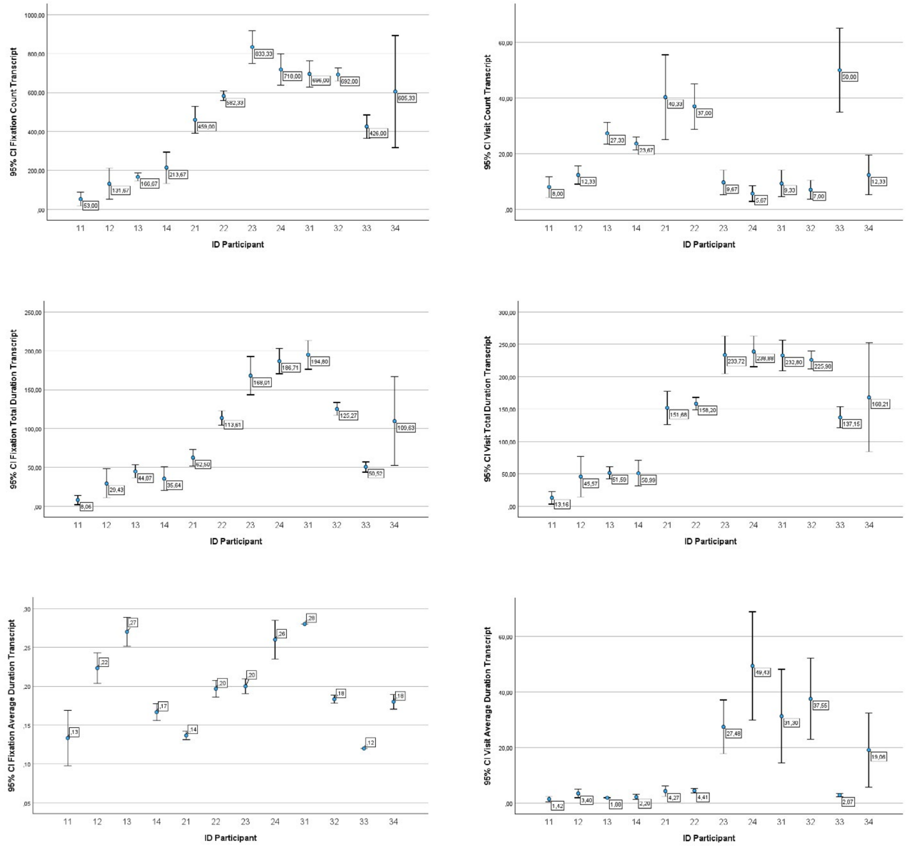
Figure 11. PTAR per condition, in relation with speech rate

3.2.3 RQ3: Visual attention in SIMTXT

Before conducting a LMM analysis of visual attention to the transcript, we conducted a test of normality to determine participants' differences in gaze behaviour. This test was significant in all cases (see Appendix 3), indicating that metrics were not normally distributed. We also created an error bar graph per metrics which clearly show that participants behaved differently (Figures 12–17). For example, as far as visits are concerned, the average visit duration was rather low for 7 participants (11–14, 21, 22, 33), and much higher for the 5 others. This clearly shows a high variation in visual attention paid to the transcript among participants, echoing results from the case study by Robert *et al.* (2025) and likely to be related to interpreters' individual style (Rennert 2008).

When considering OIQ as DV, the best model in SIMTXT (Appendix 2, Model C1: -2 Log=494.38; 9 parameters) consisted of the following significant fixed effects: (1) the interaction effect between speech rate and speech difficulty ($F(2,58.45) = 6.66$, $p=0.002$) and; (2) the average visit duration to the transcript ($F(1,43.56)=4.30$ $p=0.04$). The interaction effect of speech rate and speech difficulty is in line with the previous results in Section 3.2.2. In the 'fast and difficult' subpart, the OQI increases by 15.84 units[15] compared to the reference (moderate and difficult). The increase is

15 Effect size: 15.84 (95% CI: 10.38 – 21.30) + 0 + 0 (see Table 4, Appendix 2, Model C1).



Figures 12, 13, 14, 15, 16, 17: Metrics of visual attention to the transcript, from left to right

less sharp in the 'fast and easy' subpart (9.30 units)[16], even lower and comparable in 'moderate and easy' (5.81 units)[17], and 'slow and difficult' (5.17 units)[18]. It is negative and almost equal to the reference in the 'slow and easy' (-0.21 units)[19] subparts. In addition, as shown in Figure 18, the longer the average visit duration to the transcript, the better the performance (0.16 fewer errors per additional second, 95% CI: -0.31 – -0.004), irrespective of the speech rate or speech difficulty.

We conducted the same test for SIMTXT with the PTAR as DV. The best model (Appendix 2, Model C2: -Log=588.22; 6 parameters) consisted of two significant fixed

16 Effect size: 15.84 (95% CI: 10.38 – 21.30) + 5.81 (95% CI: 0.54 –11.01) + -12.35 (95% CI: -19.83 – -4.88) (see Table 4, Appendix 2, Model C1).
17 Effect size: 5.81 (95% CI: 0.54 –11.01) + 0 + 0 (see Table 4, Appendix 2, Model C1).
18 Effect size: 5.17 (95% CI: -0.12 – 10.46) + 0 + 0 (see Table 4, Appendix 2, Model C1).
19 Effect size: 5.17 (95% CI: -0.12 – 10.46) + 5.81 (95% CI: 0.54 –11.01) + -11.19 (95% CI: -18.67 – -3.71 (see Table 4, Appendix 2, Model C1).

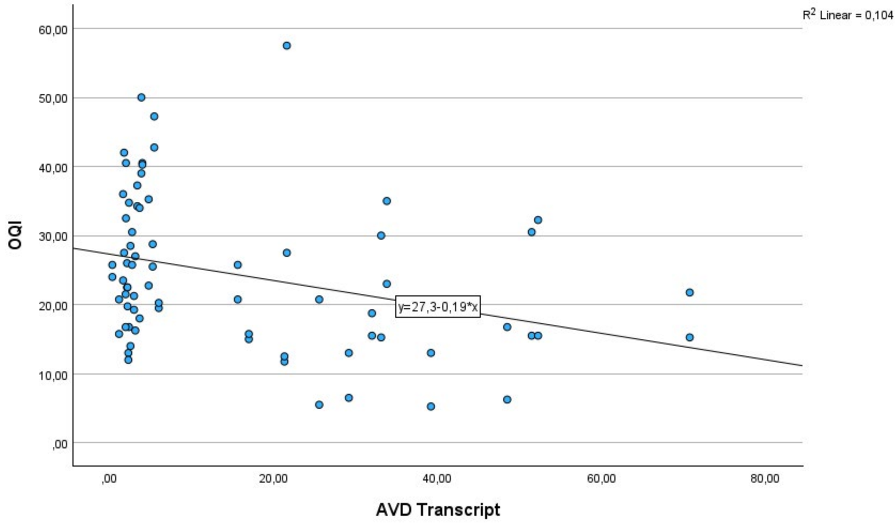


Figure 18: Overall quality indicator in relation to visual attention to the transcript

effects, i.e. speech rate (F(2,60)=4.24, *p*=0.019) and speech difficulty (F(1,60)=5.67, *p*=0.020). There was no interaction between the two predictors and no visual attention metrics were significant and/or improved the model. In other words, when interpreting performance was measured by drawing on the PTAR only, performance in SIMTXT only varied with speech rate or speech difficulty, with a worse performance under fast speech rate (decrease of 8.44 units compared to the reference, i.e. moderate speech rate, 95% CI: -16.09 – -0.80; see Table 4, Appendix 2, Model C2) or when there are many problem triggers (7.43 units lower, 95% CI: -1.19 – 13.67; see Table 4, Appendix 2, Model C2). Visual attention to the transcript did not impact performance.

4. Discussion and conclusion

The aim of this paper was to determine whether the availability of the transcript of a speech impacts performance in SI and if so, whether other variables (speech rate, speech difficulty and visual attention) predict performance as well. To do so, 12 young conference interpreters performed two similar speeches simultaneously, without transcript (SI) and with transcript (SIMTXT). Data were collected through audio recordings and eye-tracking glasses.

Results show that the interpreting performance of the participants was superior in SIMTXT. This is in line with Lambert (2004), Ma/Cheung (2020), Yang *et al.* (2020) and Zou *et al.* (2022). However, there were other significant predictors of performance as well. When interpreting performance was measured as the number and weight of errors (Overall Quality Indicator, OQI), speech rate, a well-known problem trigger, interacted with speech difficulty: when the speech rate was high and the speech difficult, interpreters performed less well. This was the case in both conditions, indicating that the interaction effect was similar in SI and SIMTXT. When interpreting performance was measured drawing only on the accurate rendition of problem triggers (Problem Trigger Accuracy Rate, PTAR), speech difficulty was found to be a significant predictor of performance, with a poorer performance when the speech was difficult. In addition, an interaction effect was found between condition and speech rate: the performance was poorer when the speech rate was fast, even more so in SI. These results are in line with Pio (2003), Barghout *et al*. (2015), and Korpal (2017).

As to visual attention in SIMTXT, the eye-tracking data reveal great variety in participants' visual attention to the transcript. Whereas certain participants devoted substantial attention to it, others did so markedly less. This confirms previous research by Robert *et al.* (2025). In the SIMTXT condition, at least when performance was measured via OQI, visual attention to the transcript measured by average visit duration (AVD) shows that the more visual attention interpreters paid to the transcript, the better they performed, even if the effect remains limited. No other metrics of visual attention were a significant predictor: there was neither an effect of the number of fixations or visits, nor an effect of total fixation or visit duration, nor of average fixation duration. Apparently, devoting attention to the script frequently did not suffice, interpreters needed to take the time to process the transcript by 'visiting' it long enough. There was no interaction of visual attention allocation with speech rate or difficulty. However, as indicated before, during high speech rate and difficulty, interpreters performed less well. When the evaluation was based on PTAR, visual attention did not impact performance. However, speech rate and speech difficulty were again significant predictors of performance, but without interaction.

Our research has several limitations. The main one is the absence of counterbalancing of condition, speech rate and speech difficulty. Due to the small number of participants and the design consisting of three groups of four interpreters, each group started with the mode they were most familiar with (SI). Counterbalancing speech rate and speech difficulty would have required creating six groups (Latin Square design) to avoid the results being skewed by a potential fatigue effect. However, this was not feasible within the scope of this project for both practical and budgetary reasons. Nevertheless, our results demonstrate that fatigue does not seem to have played an important role: participants performed better in the second task and never performed worse in the last two subparts of each task. Therefore, this does not seem to have been detrimental to the study's design. A second limitation is the relatively heterogeneous group of participants, caused by the difficulty to recruit interpreters willing to participate, although great care was taken to recruit similar profiles. Last, the transcript was congruent with the speaker's words, which is not always the case in real life.

The results are relevant for both training and future research. We have shown that the transcript seemed to facilitate the interpreting task for young professional interpreters, but only when used efficiently, that is, when the transcript was visited sufficiently long enough to process its content. This implies that future interpreters need to be trained on how to use this particular type of support in a strategic way. On a theoretical level, since the results of this study provide evidence for better performance in SIMTXT, they confirm Gile's (2009) Effort Model in the sense that the verbal support by means of the transcript reduces the memory effort and liberates mental energy that can be devoted to other efforts. At the same time, they undermine Seeber's

(2017) theory that predicts increased CL as a result of interfering efforts (reading and listening). However, this is valid only for CL as expressed by means of performance and does not take into account further measurements.

Funding information

This research has been funded by the University of Antwerp, through the BOF-SIMPRO funding scheme (project-ID 46247).

Acknowledgement

We would like to thank Professor Erik Fransen (University of Antwerp) for his advice on the statistical analyses in this paper.

References

Barghout A. / Ruiz Rosendo L. / Varela García M. (2015) "The influence of speed on omissions in simultaneous interpretation: an experimental study", *Babel* 61/3, 305-334.

Bóna J. / Bakti M. (2020) "The effect of cognitive load on temporal and disfluency patterns of speech: evidence from consecutive interpreting and sight translation", *Target* 32/3, 482-506.

Carrasco M. (2011) "Visual attention: the past 25 years", *Vision Research* 51/13, 1484-1525.

Chang C. C. (2005) *Directionality in Chinese/English Simultaneous Interpreting: Impact on Performance and Strategy Use*, unpublished doctoral dissertation, The University of Texas at Austin.

Chen S. (2017) "The construct of cognitive load in interpreting and its measurement", *Perspectives* 25/4, 640-657.

Chmiel A. / Janikowski P. / Lijewska A. (2020) "Multimodal processing in simultaneous interpreting with text: interpreters focus more on the visual than the auditory modality", *Target* 32/1, 37-58.

Defrancq B. / Fantinuoli C. (2021) "Automatic speech recognition in the booth: assessment of system performance, interpreters' performances and interactions in the context of numbers", *Target* 33/1, 73-102.

Desmet B. / Vandierendonck M. / Defrancq B. (2018) "Simultaneous interpretation of numbers and the impact of technological support", in C. Fantinuoli (ed.), *Interpreting and technology,* Berlin: Language Science Press, 13-27.

Field A. (2009) *Discovering Statistics Using SPSS*, London, SAGE Publications Ltd.

Gieshoff A. C. / Albl-Mikasa M. (2022) "Interpreting accuracy revisited: a refined approach to interpreting performance analysis", *Perspectives* 32/2, 210-228.

Gile D. (2009) *Basic Concepts and Models for Interpreter and Translator Training*, Amsterdam: John Benjamins.

Gile D. (2020) “The effort models and gravitational mode: clarifications and update [PowerPoint slides]. <https://www.researchgate.net/publication/325485989_The_Effort_Models_and_Gravitational_Model_Clarifications_and_Update_01_June_2018_version>.

Han C. / Riazi M. (2017) “Investigating the effects of speech rate and accent on simultaneous interpretation: a mixed-methods approach”, *Across Languages and Cultures* 18, 237-259.

Holmqvist K. / Nystrom M. / Anderson, R. / Dewhurst R. / Jarodzka H. / van de Weiter J. (2015) *Eye Tracking: A Comprehensive Guide to Methods and Measures,* Oxford, England: University Press.

Ivanov K. / Davies K. / Naimushin B. (2014) “Teaching simultaneous interpreting with text”, in S. Goncharov / A. Antonova (eds) *Fighting the Fog of Multiculturalism*, St. Petersburg, Herzen University Press, 48-61.

Just M. A. / Carpenter P. A. (1980) “A theory of reading: from eye fixations to comprehension”, *Psychological Review* 87/4, 329-354.

Korpal P. (2015) “Eye-tracking in translation and interpreting studies: the growing popularity and methodological problems”, in Ł. Bogucki / M. Deckert (eds.) *Accessing Audiovisual Translation,* Berlin, Peter Lang, 199-212.

Korpal P. (2017) *Linguistic and Psychological Indicators of Stress in Simultaneous Interpreting*, Poznań, Adam Mickiewicz University Press.

Korpal P. / Stachowiak-Szymczak K. (2020) “Combined problem triggers in simultaneous interpreting: exploring the effect of delivery rate on processing and rendering numbers”, *Perspectives*, 28/1, 126-143.

Lamberger-Felber H. (2001) “Text-oriented research into interpreting: examples from a case-study”, *Hermes - Journal of Linguistics* 14/26, 39-64.

Lamberger-Felber H. / Schneider J. (2008) “Linguistic interference in simultaneous interpreting with text: a case study”, in G. Hansen / A. Chesterman / H. Gerzymisch-Arbogast (eds) *Efforts and Models in Interpreting and Translation Research: A Tribute to Daniel Gile,* Amsterdam, John Benjamins, 215-236.

Lambert S. (2004) “Shared attention during sight translation, sight interpretation and simultaneous interpretation”, *Meta* 49/2, 294-306.

Li T. / Chmiel A. (2024) “Automatic subtitles increase accuracy and decrease cognitive load in simultaneous interpreting”, *Interpreting* 26/2, 253-281.

Liu M. / Schallert D. L. / Carroll P. J. (2004) “Working memory and expertise in simultaneous interpreting”, *Interpreting* 6/1, 19-42.

Ma X. / Cheung A. K. (2020) “Language interference in English-Chinese simultaneous interpreting with and without text”, *Babel* 66/3, 434-456.

Mankauskienė D. (2016) “Problem trigger classification and its applications for empirical research”, *Procedia - Social and Behavioral Sciences* 231, 143-148.

Mead P. (2015) “Input variables”, in F. Pöchhacker (ed.) *Routledge Encyclopedia of Interpreting Studies*, London, Routledge, 191-192.

Nuthmann, A. / Henderson J. (2012) "Using CRISP to model global characteristics of fixation durations in scene viewing and reading with a common mechanism", *Visual Cognition* 20/4–5, 457-494.

Perego E. / Del Missier F. / Porta M. / Mosconi M. (2010) "The cognitive effectiveness of subtitle processing", *Media Psychology,* 13/3, 243-272.

Pio S. (2003) "The relation between ST delivery rate and quality in simultaneous interpretation", *The Interpreters' Newsletter* 12, 69-100.

Rennert S. (2008) "Visual Input in Simultaneous Interpreting", *Meta*, 53/1, 204-217.

Robert I. S. / De Boe E. / Saxena P. / Crijns M. (2025) "The impact of speech rate in simultaneous interpreting with text: an exploratory case study using eye tracking", *TransKom* 18/1, 224-256.

Seeber K. G. (2015a) "Simultaneous interpreting", in H. Mikkelson / R. Jourdenais (eds) *The Routledge Handbook of Interpreting*, London, Routledge, 79-95.

Seeber K. G. (2015b) "Eye tracking", in F. Pöchhacker (ed.) *Routledge Encyclopedia of Interpreting Studies,* London, Routledge, 157.

Seeber K. G. (2016, September 29-30) "Modeling multimodal processing in simultaneous interpreting", Poster session, the FTI's 75th Anniversary, Geneva.

Seeber K. G. (2017) "Multimodal processing in simultaneous interpreting", in J. W. Schwieter / A. Ferreira (eds) *The Handbook of Translation and Cognition*, Hoboken, NJ, John Wiley & Sons Inc, 461-475.

Seeber K. G. / Delgado Luchner C. (2020) "Simulating simultaneous interpreting with text: from training model to training module", in M. D. R. Melchor / I. Horváth / K. Ferguson (eds) *The Role of Technology in Conference Interpreter Training*, Oxford, Peter Lang, 70-86.

Seeber K. G. / Keller L. / Hervais-Adelman A. (2020) "When the ear leads the eye – the use of text during simultaneous interpretation", *Language, Cognition and Neuroscience* 35/10, 1480-1494.

Setton R. / Dawrant A. (2016) *Conference Interpreting: A Trainers' Guide*, Amsterdam, John Benjamins.

Shlesinger M (2003) "Effects of presentation rate on working memory in simultaneous interpreting", *The Interpreter's Newsletter 12*, 37-49.

Shreve G. M. / Lacruz I. / Angelone E. (2010) "Cognitive effort, syntactic disruption, and visual interference in a sight translation task", *Translation and Cognition* 15, 63-84.

Stachowiak-Szymczak K. (2019) "Eye movements and gestures in embodied language processing and interpreting", in M. Pawlak (ed.) *Second Language Learning and Teaching: Eye Movements and Gestures in Simultaneous and Consecutive Interpreting,* Cham: Springer, 63-84.

Stachowiak-Szymczak K. / Korpal P. (2019) "Interpreting accuracy and visual processing of numbers in professional and student interpreters: an eye-tracking study", *Across Languages and Cultures* 20/2, 235-251.

Tobii (2021) "A complete guide to the fundamentals of eye tracking". https://go.tobii.com/eye-tracking-fundamentals.

Yang S. / Defeng L. / Lai Cheng, L. V. (2020) "The impact of source text presence on simultaneous interpreting performance in fast speeches: will it help trainees or not?", *Babel* 66/4–5, 588-603.

Zou L. / Carl M. / Feng J. (2022) "Patterns of attention and quality in English-Chinese simultaneous interpreting with text", *International Journal of Chinese and English Translation & Interpreting* 2, 1-23.

Appendices

Appendix 1

ID	Gender	Age	Graduation EMCI	Main job	% of workload as an active interpreter	Experience with SI (in years)	Experience with SIMTXT
11	Male	28	2021	Interpreter	70	1.5	Quite a lot
12	Female	28	2019	Interpreter	15	1.5	Some
13	Female	24	2022	Employee	0	Only during internship	Some
14	Male	26	2021	Interpreter	30	1	Some
21	Female	25	2021	Interpreter	20	1.5	None
22	Female	23	2022	Student	0	0.1	None
23	Female	23	2022	Teacher	0	0.1	None
24	Male	23	2022	Employee	0	Only during internship	Some
31	Female	29	2019	Interpreter	60	2.5	Quite a lot
32	Female	27	2019	Subtitler–audio describer	5	3	None
33	Female	23	2022	Interpreter	10	0.5	Some
34	Male	24	2022	Interpreter	55	0.8	Some

Table 1: Profile description of the participants

Simultaneous interpreting (with text): Impact of condition, speech difficulty, speech rate and visual attention on young professional interpreters' performance

Isabelle S. Robert
Esther de Boe
Priya Saxena
University of Antwerp

Appendix 2

Model A1 – OQI

Model Dimension[a]					
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	Condition	2		1	
Random Effects	Intercept	1	Variance Components	1	ID_Participant * Speech
Residual				1	
Total		4		4	
a. Dependent Variable: OQI.					

Table 1: Model dimension

Information Criteria[a]	
-2 Log Likelihood	1070,27781955
Akaike's Information Criterion (AIC)	1078,27781955
Hurvich and Tsai's Criterion (AICC)	1078,56558933
Bozdogan's Criterion (CAIC)	1094,15707275
Schwarz's Bayesian Criterion (BIC)	1090,15707275
The information criteria are displayed in smaller-is-better form.	
a. Dependent Variable: OQI.	

Table 2: Information criteria

Type III Tests of Fixed Effects[a]				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	24	499,676	<,001
Condition	1	24	13,199	,001
a. Dependent Variable: OQI.				

Table 3: Tests of fixed effects

Estimates of Fixed Effects[a]							
						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	24,351	1,840	24	13,237	<,001	20,554	28,147
[Condition=1]	9,451	2,602	24	3,633	,001	4,082	14,821
[Condition=2]	0[b]	0	.	.	.	.	.
a. Dependent Variable: OQI.							
b. This parameter is set to zero because it is redundant.							

Table 4: Estimates of fixed effects

Estimates of Covariance Parameters[a]							
						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Residual		82,631	10,668	7,746	<,001	64,159	106,423
Intercept [subject = ID_Participant * Speech]	Variance	26,835	11,856	2,263	,024	11,288	63,795
a. Dependent Variable: OQI.							

Table 5: Estimates of covariance parameters

Model A1b – OQI

Model Dimension[a]					
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	Condition	2		1	
	Experience_SI_Years	1		1	
Random Effects	Intercept	1	Variance Components	1	ID_Participant * Speech
Residual				1	
Total		5		5	
a. Dependent Variable: OQI.					

Table 1: Model dimension

Information Criteria[a]	
-2 Log Likelihood	1067,93118677
Akaike's Information Criterion (AIC)	1077,93118677
Hurvich and Tsai's Criterion (AICC)	1078,36596938
Bozdogan's Criterion (CAIC)	1097,78025327
Schwarz's Bayesian Criterion (BIC)	1092,78025327
The information criteria are displayed in smaller-is-better form.	
a. Dependent Variable: OQI.	

Table 2: Information criteria

Type III Tests of Fixed Effects[a]				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	24	288,989	<,001
Condition	1	24	14,555	<,001
Experience_SI_Years	1	24	2,465	,129
a. Dependent Variable: OQI.				

Table 3: Tests of fixed effects

Estimates of Fixed Effects[a]							
						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	26,477	2,214	24	11,957	<,001	21,907	31,048
[Condition=1]	9,451	2,477	24	3,815	<,001	4,338	14,564
[Condition=2]	0[b]	0	.	.	.	.	.
Experience_SI_Years	-2,042	1,300	24	-1,570	,129	-4,725	,642
a. Dependent Variable: OQI.							
b. This parameter is set to zero because it is redundant.							

Table 4: Estimates of fixed effects

Estimates of Covariance Parameters[a]								
						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Residual		82,631	10,668	7,746	<,001	64,159	106,423
Intercept [subject = ID_Participant * Speech]	Variance	23,053	10,778	2,139	,032	9,221	57,636
a. Dependent Variable: OQI.							

Table 5: Estimates of covariance parameters

Model A2 – OQI

Model Dimension[a]					
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	Condition	2		1	
	SpeechRate	3		2	
	Difficulty_nb_RP	2		1	
	SpeechRate * Difficulty_nb_RP	6		2	
Random Effects	Intercept	1	Variance Components	1	ID_Participant * Speech
Residual				1	
Total		15		9	
a. Dependent Variable: OQI.					

Table 1: Model dimension

Information Criteria[a]	
-2 Log Likelihood	1029,14624567
Akaike's Information Criterion (AIC)	1047,14624567
Hurvich and Tsai's Criterion (AICC)	1048,48952925
Bozdogan's Criterion (CAIC)	1082,87456536
Schwarz's Bayesian Criterion (BIC)	1073,87456536
The information criteria are displayed in smaller-is-better form.	
a. Dependent Variable: OQI.	

Table 2: Information criteria

Type III Tests of Fixed Effects[a]				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	24	499,676	<,001
Condition	1	24	13,199	,001
SpeechRate	2	120	11,770	<,001
Difficulty_nb_RP	1	120	7,609	,007
SpeechRate * Difficulty_nb_RP	2	120	8,956	<,001
a. Dependent Variable: OQI.				

Table 3: Tests of fixed effects

Estimates of Fixed Effects[a]							
						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	20,253	2,328	57,421	8,699	<,001	15,592	24,915
[Condition=1]	9,451	2,602	24	3,633	,001	4,082	14,821
[Condition=2]	0[b]	0	.	.	.	.	.
[SpeechRate=1]	5,844	2,211	120	2,643	,009	1,466	10,221
[SpeechRate=2]	11,729	2,211	120	5,305	<,001	7,352	16,106
[SpeechRate=3]	0[b]	0	.	.	.	.	.
[Difficulty_nb_RP=1]	4,063	2,211	120	1,838	,069	-,315	8,440
[Difficulty_nb_RP=2]	0[b]	0	.	.	.	.	.
[SpeechRate=1] * [Difficulty_nb_RP=1]	-12,177	3,127	120	-3,895	<,001	-18,367	-5,987
[SpeechRate=1] * [Difficulty_nb_RP=2]	0[b]	0	.	.	.	.	.
[SpeechRate=2] * [Difficulty_nb_RP=1]	-10,573	3,127	120	-3,382	<,001	-16,763	-4,383
[SpeechRate=2] * [Difficulty_nb_RP=2]	0[b]	0	.	.	.	.	.
[SpeechRate=3] * [Difficulty_nb_RP=1]	0[b]	0	.	.	.	.	.
[SpeechRate=3] * [Difficulty_nb_RP=2]	0[b]	0	.	.	.	.	.
a. Dependent Variable: OQI.							
b. This parameter is set to zero because it is redundant.							

Table 4: Estimates of fixed effects

Estimates of Covariance Parameters[a]								
						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Residual		58,652	7,572	7,746	<,001	45,540	75,539
Intercept [subject = ID_Participant * Speech]	Variance	30,832	11,790	2,615	,009	14,571	65,238
a. Dependent Variable: OQI.							

Table 5: Estimates of covariance parameters

Model B1 – PTAR

Model Dimension[a]					
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	Condition	2		1	
Random Effects	Intercept	1	Variance Components	1	ID_Participant * Speech
Residual				1	
Total		4		4	
a. Dependent Variable: PTAR.					

Table 1: Model dimension

Information Criteria[a]	
-2 Log Likelihood	1212,40570051
Akaike's Information Criterion (AIC)	1220,40570051
Hurvich and Tsai's Criterion (AICC)	1220,69347029
Bozdogan's Criterion (CAIC)	1236,28495370
Schwarz's Bayesian Criterion (BIC)	1232,28495370
The information criteria are displayed in smaller-is-better form.	
a. Dependent Variable: PTAR.	

Table 2: Information criteria

Type III Tests of Fixed Effects[a]				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	24	2075,943	<,001
Condition	1	24	22,163	<,001
a. Dependent Variable: PTAR.				

Table 3: Tests of fixed effects

Estimates of Fixed Effects[a]							
						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	78,228	2,201	24	35,546	<,001	73,686	82,770
[Condition=1]	-14,652	3,112	24	-4,708	<,001	-21,076	-8,229
[Condition=2]	0[b]	0	.	.	.	.	.
a. Dependent Variable: PTAR.							
b. This parameter is set to zero because it is redundant.							

Table 4: Estimates of fixed effects

Estimates of Covariance Parameters[a]							
						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Residual		251,409	32,457	7,746	<,001	195,205	323,795
Intercept [subject = ID_Participant * Speech]	Variance	16,217	17,628	,920	,358	1,926	136,529
a. Dependent Variable: PTAR.							

Table 5: Estimates of covariance parameters

Model B1b – PTAR

Model Dimension[a]					
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	Condition	2		1	
	Experience_SI_Years	1		1	
Random Effects	Intercept	1	Variance Components	1	ID_Participant * Speech
Residual				1	
Total		5		5	
a. Dependent Variable: PTAR.					

Table 1: Model dimension

Information Criteria[a]	
-2 Log Likelihood	1210,15030073
Akaike's Information Criterion (AIC)	1220,15030073
Hurvich and Tsai's Criterion (AICC)	1220,58508334
Bozdogan's Criterion (CAIC)	1239,99936723
Schwarz's Bayesian Criterion (BIC)	1234,99936723
The information criteria are displayed in smaller-is-better form.	
a. Dependent Variable: PTAR.	

Table 2: Information criteria

Type III Tests of Fixed Effects[a]				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	24	966,732	<,001
Condition	1	24	24,347	<,001
Experience_SI_Years	1	24	2,365	,137
a. Dependent Variable: PTAR.				

Table 3: Tests of fixed effects

Estimates of Fixed Effects[a]							
						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	75,732	2,654	24	28,533	<,001	70,254	81,210
[Condition=1]	-14,652	2,969	24	-4,934	<,001	-20,781	-8,523
[Condition=2]	0[b]	0	.	.	.	.	.
Experience_SI_Years	2,397	1,559	24	1,538	,137	-,820	5,614
a. Dependent Variable: PTAR.							
b. This parameter is set to zero because it is redundant.							

Table 4: Estimates of fixed effects

Estimates of Covariance Parameters[a]							
						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Residual		251,409	32,457	7,746	<,001	195,205	323,795
Intercept [subject = ID_Participant * Speech]	Variance	11,004	16,202	,679	,497	,614	197,170
a. Dependent Variable: PTAR.							

Table 5: Estimates of covariance parameters

Model B2 – PTAR

Model Dimension[a]					
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	Condition	2		1	
	SpeechRate	3		2	
	Difficulty_nb_RP	2		1	
	SpeechRate * Condition	6		2	
Random Effects	Intercept	1	Variance Components	1	ID_Participant * Speech
Residual				1	
Total		15		9	
a. Dependent Variable: PTAR.					

Table 1: Model dimension

Information Criteria[a]	
-2 Log Likelihood	1156,43517775
Akaike's Information Criterion (AIC)	1174,43517775
Hurvich and Tsai's Criterion (AICC)	1175,77846133
Bozdogan's Criterion (CAIC)	1210,16349745
Schwarz's Bayesian Criterion (BIC)	1201,16349745
The information criteria are displayed in smaller-is-better form.	
a. Dependent Variable: PTAR.	

Table 2: Information criteria

Type III Tests of Fixed Effects[a]				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	24	2075,943	<,001
Condition	1	24	22,163	<,001
SpeechRate	2	120	25,121	<,001
Difficulty_nb_RP	1	120	10,340	,002
SpeechRate * Condition	2	120	5,366	,006
a. Dependent Variable: PTAR.				

Table 3: Tests of fixed effects

Estimates of Fixed Effects[a]							
						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	76,992	3,212	86,763	23,968	<,001	70,607	83,377
[Condition=1]	-18,600	4,295	74,807	-4,331	<,001	-27,157	-10,044
[Condition=2]	0[b]	0	.	.	.	.	.
[SpeechRate=1]	2,055	3,625	120	,567	,572	-5,122	9,233
[SpeechRate=2]	-8,442	3,625	120	-2,329	,022	-15,619	-1,265
[SpeechRate=3]	0[b]	0	.	.	.	.	.
[Difficulty_nb_RP=1]	6,730	2,093	120	3,216	,002	2,586	10,874
[Difficulty_nb_RP=2]	0[b]	0	.	.	.	.	.
[SpeechRate=1] * [Condition=1]	13,592	5,127	120	2,651	,009	3,442	23,742
[SpeechRate=1] * [Condition=2]	0[b]	0	.	.	.	.	.
[SpeechRate=2] * [Condition=1]	-1,747	5,127	120	-,341	,734	-11,897	8,403
[SpeechRate=2] * [Condition=2]	0[b]	0	.	.	.	.	.
[SpeechRate=3] * [Condition=1]	0[b]	0	.	.	.	.	.
[SpeechRate=3] * [Condition=2]	0[b]	0	.	.	.	.	.
a. Dependent Variable: PTAR.							
b. This parameter is set to zero because it is redundant.							

Table 4: Estimates of fixed effects

Estimates of Covariance Parameters[a]							
						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Residual		157,694	20,358	7,746	<,001	122,441	203,098
Intercept [subject = ID_Participant * Speech]	Variance	31,836	17,117	1,860	,063	11,098	91,323
a. Dependent Variable: PTAR.							

Table 5: Estimates of covariance parameters

Model C1 – OQI, SIMTXT only

Model Dimension[a]					
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	SpeechRate	3		2	
	Difficulty_nb_RP	2		1	
	SpeechRate * Difficulty_nb_RP	6		2	
	Visit_AverageDuration_Transcript	1		1	
Random Effects	Intercept	1	Variance Components	1	ID_Participant
Residual				1	
Total		14		9	
a. Dependent Variable: OQI.					

Table 1: Model dimension

Information Criteria[a]	
-2 Log Likelihood	494,37606068
Akaike's Information Criterion (AIC)	512,37606068
Hurvich and Tsai's Criterion (AICC)	515,27928648
Bozdogan's Criterion (CAIC)	541,86605575
Schwarz's Bayesian Criterion (BIC)	532,86605575
The information criteria are displayed in smaller-is-better form.	
a. Dependent Variable: OQI.	

Table 2: Information criteria

Type III Tests of Fixed Effects[a]				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	15,033	151,063	<,001
SpeechRate	2	62,501	16,000	<,001
Difficulty_nb_RP	1	58,453	1,779	,188
SpeechRate * Difficulty_nb_RP	2	58,453	6,662	,002
Visit_AverageDuration_Transcript	1	43,559	4,298	,044
a. Dependent Variable: OQI.				

Table 3: Tests of fixed effects

Estimates of Fixed Effects[a]							
						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	20,789	2,672	34,615	7,779	<,001	15,362	26,216
[SpeechRate=1]	5,170	2,643	58,476	1,956	,055	-,120	10,460
[SpeechRate=2]	15,837	2,730	62,062	5,801	<,001	10,380	21,295
[SpeechRate=3]	0[b]	0	.	.	.	.	.
[Difficulty_nb_RP=1]	5,812	2,643	58,453	2,200	,032	,524	11,101
[Difficulty_nb_RP=2]	0[b]	0	.	.	.	.	.
[SpeechRate=1] * [Difficulty_nb_RP=1]	-11,187	3,737	58,453	-2,994	,004	-18,667	-3,708
[SpeechRate=1] * [Difficulty_nb_RP=2]	0[b]	0	.	.	.	.	.
[SpeechRate=2] * [Difficulty_nb_RP=1]	-12,354	3,737	58,453	-3,306	,002	-19,834	-4,875
[SpeechRate=2] * [Difficulty_nb_RP=2]	0[b]	0	.	.	.	.	.
[SpeechRate=3] * [Difficulty_nb_RP=1]	0[b]	0	.	.	.	.	.
[SpeechRate=3] * [Difficulty_nb_RP=2]	0[b]	0	.	.	.	.	.
Visit_AverageDuration_Transcript	-,157	,076	43,559	-2,073	,044	-,310	-,004
a. Dependent Variable: OQI.							
b. This parameter is set to zero because it is redundant.							

Table 4: Estimates of fixed effects

Estimates of Covariance Parameters[a]							
						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Residual		41,900	7,750	5,406	<,001	29,159	60,210
Intercept [subject = ID_Participant]	Variance	33,566	17,733	1,893	,058	11,918	94,533
a. Dependent Variable: OQI.							

Table 5: estimate of covariance parameters

Model C2 – PTAR, SIMTXT only

Model Dimension[a]					
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	SpeechRate	3		2	
	Difficulty_nb_RP	2		1	
Random Effects	Intercept	1	Variance Components	1	ID_Participant
Residual				1	
Total		7		6	
a. Dependent Variable: PTAR.					

Table 1: Model dimension

Information Criteria[a]	
-2 Log Likelihood	588,21734463
Akaike's Information Criterion (AIC)	600,21734463
Hurvich and Tsai's Criterion (AICC)	601,50965232
Bozdogan's Criterion (CAIC)	619,87734134
Schwarz's Bayesian Criterion (BIC)	613,87734134
The information criteria are displayed in smaller-is-better form.	
a. Dependent Variable: PTAR.	

Table 2: Information criteria

Type III Tests of Fixed Effects[a]				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	12	930,820	<,001
SpeechRate	2	60	4,238	,019
Difficulty_nb_RP	1	60	5,674	,020
a. Dependent Variable: PTAR.				

Table 3: Tests of fixed effects

Estimates of Fixed Effects[a]							
						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	76,641	3,725	42,880	20,574	<,001	69,128	84,154
[SpeechRate=1]	2,055	3,822	60	,538	,593	-5,589	9,700
[SpeechRate=2]	-8,442	3,822	60	-2,209	,031	-16,086	-,798
[SpeechRate=3]	0[b]	0	.	.	.	.	.
[Difficulty_nb_RP=1]	7,433	3,120	60	2,382	,020	1,191	13,674
[Difficulty_nb_RP=2]	0[b]	0	.	.	.	.	.
a. Dependent Variable: PTAR.							
b. This parameter is set to zero because it is redundant.							

Table 4: Estimates of fixed effects

Estimates of Covariance Parameters[a]							
						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Residual		175,247	31,996	5,477	<,001	122,530	250,645
Intercept [subject = ID_Participant]	Variance	49,686	32,647	1,522	,128	13,707	180,105
a. Dependent Variable: PTAR.							

Table 5: Estimates of covariance parameters

Appendix 3

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Fixation_TotalDuration_Transcript	,196	72	<,001	,909	72	<,001
Fixation_AverageDuration_Transcript	,127	72	,006	,955	72	,012
Fixation_Count_Transcript	,118	72	,014	,937	72	,001
Visit_TotalDuration_Transcript	,122	72	,010	,917	72	<,001
Visit_AverageDuration_Transcript	,307	72	<,001	,765	72	<,001
Visit_Count_Transcript	,168	72	<,001	,868	72	<,001

a. Lilliefors Significance Correction

Table 1. Tests of normality (Section 3.2.3)