RESEARCH ARTICLE

Patterns of Attention and Quality in English-Chinese Simultaneous Interpreting with Text

Received: 16 October 2022; Revised: 16 December 2022; Published: 21 December 2022

Longhui Zou

Kent State University, USA
Email: lzou4@kent.edu
https://orcid.org/0000-0002-0175-4029

Michael Carl

Kent State University, USA
Email: mcarl6@kent.edu
https://orcid.org/0000-0002-2815-0292

Jia Feng

Renmin University of China, China
Email: fengjia_ruc@ruc.edu.cn

Abstract:

Simultaneous interpreting with text (SIMTXT) is generally seen as being cognitively more demanding than either simultaneous interpreting (SI) or sight translation (STT). However, little research has been done on how interpreters allocate their attention proportionally during SIMTXT between their auditory and visual attention. This study examines attention patterns during SIMTXT and investigates the relationship between the interpreters' attention patterns and the quality of the interpreting output. Nine professional interpreters were recruited for SIMTXT, interpreting six English STs into Chinese. The interpreters listened to the audio input while the transcribed text of the audio input was displayed on the screen by Translog-II. The eye movements were recorded by an eye-tracker (Tobii TX300) while their interpretations were recorded by Audacity. The gaze and the spoken data (source and target) were synchronized on a word-level and aligned with the final STs and TTs. We explored and categorized the visual and auditory attention patterns based on their ear-voice span (EVS), eye-voice span (IVS), and ear-eye span (EIS) and found three types of attention patterns in our data, namely, ear-dominant (ED), eye-dominant (ID), and ear-eye-balanced (EIB). In addition, the interpreting output was annotated by three professional interpreters based on an error taxonomy adapted from Multidimensional Quality Metrics (MQM). We found that the EIB interpreters produced the lowest translation quality in terms of total error numbers and accuracy, followed by ED, and then ID interpreters. ED interpreters produced the highest translation quality in terms of fluency, followed by EIB, and then ID interpreters.

Keywords: simultaneous interpreting with text, quality assessment, attention allocation, translation process research, language specificity

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

1. Introduction

Simultaneous interpreting with text (SIMTXT) has long been an essential component of conference interpreters' professional endeavors, whether they are working for international organizations or private companies (Setton & Motta, 2007; Cammoun et al., 2009; Seeber & Delgado, 2020). This translation modality has become even more common during the global pandemic and is now a staple in many interpreting training programs, since many conferences are held online where interpreters must always be attentive to both visual and auditory inputs.

1.1 Cognitive load in SIMTXT

SIMTXT refers to the process of simultaneous interpretation of a speech whose manuscript has been made available to the interpreter before or during the delivery of the speech. Although, theoretically, SIMTXT is a combination of simultaneous interpreting (SI) and sight translation (STT), existing research suggests that it should be categorized as SI because the auditory input takes precedence, and the interpreter is externally paced by the delivery of the auditory input (Lambert, 2004; Setton, 2015; Pöchhacker, 2016; Chmiel & Lijewska, 2019).

As a special mode of SI, SIMTXT incorporates not only the task of simultaneously interpreting the auditory input from the source language (SL) into its oral equivalent in the target language (TL), but also an additional visual processing task to relate the written input to the auditory input. The dual input (auditory and visual) makes SIMTXT cognitively more demanding than SI with impromptu speech as well as STT with written text (Chmiel & Mazur, 2013; Čeňková, 2015). According to Gile's effort models for interpreting, the effort in SIMTXT is made up of the reading effort (R), listening and analysis effort (L), memory effort (M), speech production effort (P), and a coordination effort (C) that links to the resources needed to coordinate the former four efforts (SIMTXT = R + L + M + P + C) (Gile, 2009). Because of this, SIMTXT entails an extra R and a more complex C when compared to SI (SI = L + P + M + C), and an extra L and a more complex C when compared to STT (STT = R + M + P + C). Therefore, there is an additional cognitive load in SIMTXT due to the requirement to follow both the oral text and the written text.

Setton & Dawrant (2016) compare SIMTXT to SI with spontaneous speech and STT via two phases, i.e., first pass and second pass. Whereas STT interpreters may or may not have the text delivered in advance or read out loud in SL during the first pass (phase 1), SIMTXT interpreters have the text while SI interpreters only have speculative preparation. In the second pass (phase 2), the resources available to SI interpreters are speech and memory, while the resources available to STT interpreters are text, notes, and memory. The resources available to SIMTXT interpreters in this phase are speech, text, notes, and memory, giving them a total of four types of resources. In this sense, SIMTXT is more sophisticated than other modes of interpreting because the interpreters have more resources to choose from and attend to in both phases.

Drawing on a multimodal processing perspective, Seeber (2017) compares the cognitive resource footprints of SI and SIMTXT and finds that SIMTXT has an apparent added visual-verbal component that influences both the perception and cognition stages of processing. SIMTXT is expected to have a much higher overall interference score than SI, taking into account both the demand vectors and

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

conflict coefficients. This reveals that adding a written text to the SI process increases the degree of task interference and consequently, the overall cognitive load.

1.2 Attention allocation in SIMTXT

Under the extreme cognitive load in SIMTXT, it is anticipated that interpreters choose strategically how to divide their attention between the spoken and the written inputs in order to reduce the information load and maximize their performance (Ivanov et al. 2014). Although theoretical and professional standards advise that interpreters should pay primary attention to the auditory channel and that the written channel should only serve to support the auditory input (Jiménez Ivars, 1999; Pöchhacker, 2004; Gile, 2009; Setton & Dawrant, 2016), recent empirical research reveals varied findings on this topic.

Chmiel et al. (2020) hired 24 conference interpreters with Polish as their L1 and English as their L2 to carry out an SIMTXT experiment. An English audio speech was manipulated to include 60 items of stimuli (i.e., 20 proper names, 20 numbers, and 20 content terms) that were delivered in an Irish accent to lead interpreters to pay attention to the written text. There were two versions of the written text, each of which contained 10 items that were consistent with the audio text and 10 that were not for each type of stimulus. Each participant had access to the same audio text and one version of the written text during the experiment. The accuracy for each of the controlled stimuli was used to evaluate the quality of the interpreting output. The findings demonstrate that interpreters depend more on visual than on auditory input when coping with challenges in processing both inputs. Particularly in instances where the two stimuli are in conflict, this would lead to a detrimental effect on accuracy. They argue that these results are either because of the visual dominance effect (Spence, 2009) or due to the scenarios in which participants purposefully decided to base their interpretation on the visual modality for risk avoidance.

Seeber et al. (2020) compared the temporal dynamics of visual and auditory attention between SIMTXT and reading while listening (RWL). 15 conference interpreters were recruited. Four were L1 speakers of German, five were L1 speakers of Italian and six were L1 speakers of French. All were L2 speakers of English and interpreted from their L2 into L1 in the experiment. Each participant was requested to complete a SIMTXT task and RWL task with the same audio recording and its written transcript. The results indicate that during SIMTXT, the prior sentence is attended to preferentially by the interpreters, who clearly favor a visual lag. They believe this visual lag could mean that by accessing the visual input, interpreters supplement or alleviate short-term memory for auditory input rather than taking precedence over it.

Both above-mentioned studies did not find any significant effect of language specificity on either interpreter's attention allocation or interpreting quality in SIMTXT. Ma & Cheung (2020) compared SIMTXT and SI with a more genetically distinct language pair, English-to-Chinese (Xiao, 2010). They investigated seven SIMTXT sessions and eight SI sessions in regard to linguistic features such as lexical density, high-frequency words, passive constructions, and attributive clauses. It is reported that the output of SIMTXT is closer to the features of written language than that of SI, and interpreters typically employ more structural reformulation strategies while performing SIMTXT. They contend

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

that the written text frees up more attentional resources for TL generation in the bimodal setting of SIMTXT, as demonstrated by reformulation planning and online decision-making.

However, little research has been done as to how visual and auditory attention is allocated proportionally by interpreters during SIMTXT. The current study examines the allocation of the interpreter's attention when reading the written ST while listening to the spoken version of the text. We investigate whether the quality of SIMTXT is influenced by the interpreter's attention pattern to one or the other mode of input. We address the following research questions in this study:

1. How do interpreters distribute their attention between the auditory input and the visual input, and how do interpreters vary in their attention patterns?
2. How do different attention patterns affect the quality of SIMTXT?

2. Experimental Design

We use the IMBst18 and IMBi18[1] datasets for this research, which were collected in April 2018 at Renmin University in China (Feng et al., 2020), to examine how visual and auditory information is processed by interpreters.

2.1 Materials

The STs for this study are six consecutive sections from the first part of a live audio recording of a political speech. This speech was initially given by the Australian Minister for Foreign Affairs during her 2014 visit to Japan for the Fifth Japan-Australia Joint Foreign and Defense Ministerial Consultations (also known as "2+2"). It was delivered at the National Press Club in Tokyo, Japan, which is regarded as one of the highest-ranking political addresses. Each section of the audio recording lasts for roughly one minute, and the corresponding transcription is approximately 150 words. Overall, the STs in this study have a college-level reading grade level, which is considered to be difficult to read. The six STs have a combined word count of 871; Table 1 shows their readability index scores (Flesch-Kincaid Grade Level), which indicates that they are reasonably comparable (Kincaid et al., 1975).

[1] These datasets can be downloaded from the CRITT TPR-DB (https://sites.google.com/site/centretranslationinnovation/tpr-db/public-studies?authuser=0#h.p_psB1dhTX_VFL)

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

Table 1. General Descriptions of the Six STs

Text	#Words	#Segments	Readability Index
1	156	6	15.8
2	152	5	17.5
3	139	7	11.1
4	146	4	16.3
5	136	6	13.3
6	142	5	15.5
Average	145.2	5.5	14.9

2.2 Participants

Nine participants (two males and seven females) were recruited for the experiment, interpreting the six STs from English into Chinese. All participants were professional interpreters, who graduated from a postgraduate professional interpreter training program, majoring in interpreting and doing interpreting regularly. The average length of their professional interpreting practice was 6.4 years. Their first language was Chinese, and second language English.

2.3 Procedure

Each participant simultaneously interpreted the six texts by hearing the audio English ST with the written English ST shown on the screen and interpreting into Chinese. Prior to the experiment, the participants received a translation brief (in English) about their interpreting task, in which they were given time to acquire background knowledge about the speech to be interpreted. This background knowledge included profile information about the speaker, location of the speech, the audience, and the speech's major points. Besides, a glossary was provided for each text before interpreting started. The glossary contained the difficult words and expressions used in the speech along with their equivalent Chinese translations, most of which were proper nouns and noun phrases (e.g., Manila, ASEAN, the Solomon Islands). In addition, a warm-up session was conducted before the participant started interpreting the six texts. The ST for the warm-up session was extracted from the introductory part of the political speech that the participants were supposed to work with. In this way, the participants were better equipped to interpret the six texts that came after.

The experiment was conducted with each participant separately, following the same procedures. No time restrictions were imposed. The experimental design of the SIMTXT task is shown in Figure 1. Each participant received input from two sources simultaneously: the auditory input (i.e., six brief audio clips) and the written input, the transcriptions of the speech, shown on the Translog-II screen. The audio input of the speech was replayed through a headset, and the audio recordings of the participant's SIMTXT output (i.e., their spoken translations into Chinese) were gathered along with the original speech by Audacity. Translog-II was used to show the transcriptions of the audio recordings, and a Tobii TX300 eye-tracker was used to record the participant's eye movement data (Carl, 2012). From the 33 source segments, the nine interpreters produced a total of 297 target segments. The study also utilized the Adapted NASA Task Load Index to measure the participants'

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

self-reported translation difficulty for each text. The Adapted NASA Task Load Index was given to the participants to complete right after they finished interpreting each text (Sun & Shreve, 2014).

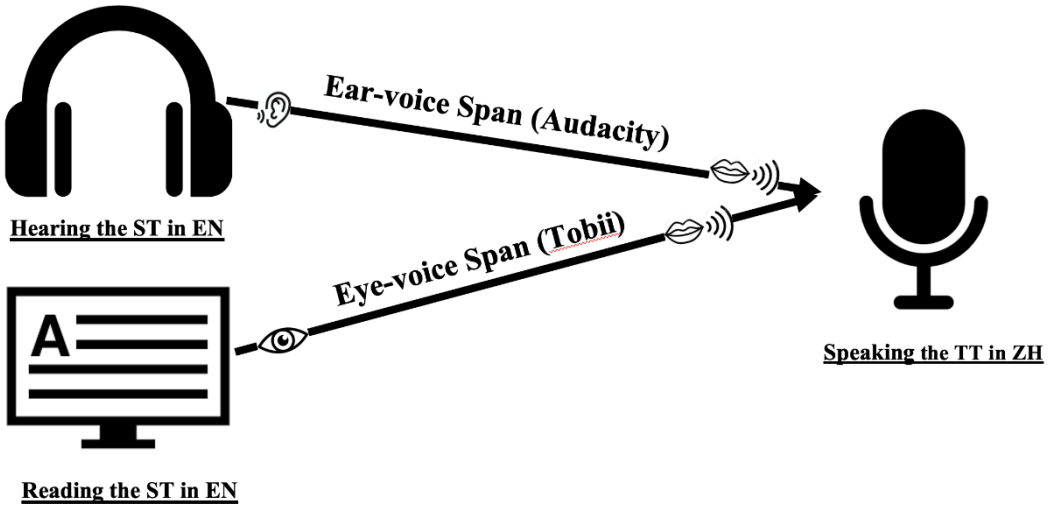

Figure 1. The Experimental Design of the SIMTXT Task

2.4 Data collection and processing

The collected datasets comprise of the input audio data, the interpreting audio output data, and the participants' reading data (i.e., eye movements during reading the English transcripts of the speech, displayed in Translog-II). The Chinese interpretation output was later transcribed, using Watson ASR and manually checked in Elan (Brugman, Russel, and Nijmegen, 2004), using a procedure described in Carl & Yamada (2017). This transcription procedure makes sure that the time stamps are preserved for each spoken word, in the ST and the TT. The data was then uploaded to the TPR-DB, and the English source and transcribed Chinese translations were aligned on the segment and word level. This somewhat laborious procedure allows us to determine the lag of time between the perception (reading/comprehension) of the English source words and the production of their Chinese translation.

3. Analysis of the Attention Patterns

We employ three time-span measures to investigate the SIMTXT process: ear-voice span (EVS), eye-voice span (IVS)[2] and ear-eye span (EIS). The time lag between the source speech input and the interpreting output, or EVS, is a measurement that is frequently used in SI. It reflects the temporal synchronization between hearing the ST and speaking the TT as well as the cognitive load during the interpreting process (Timarová, Dragsted, and Hansen, 2011; Pöchhacker, 2016). IVS, which stands

[2] The acronym IVS for eye-voice span has recently been suggested by Chmiel & Lijewska (2022). For convenience we adopt this here.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

for the time lag between the source written text input and the interpreting output, is the equivalent of EVS. IVS is commonly used in STT studies and serves as an indicator of processing effort (Inhoff et al., 2011; Zheng & Zhou, 2018; Chmiel, Janikowski, and Cieślewicz, 2020; Chmiel & Lijewska, 2022). Since SIMTXT is a hybrid of SI and STT, we can also use a measure of EIS, which is the temporal difference between the source speech input and the source written text input. We compute the EVS and IVS to assess auditory and visual attention patterns, respectively, and the EIS to measure the relationship between them.

3.1 Ear-voice span (EVS) and eye-voice span (IVS)

In this study, EVS refers to the interval between the starting time of a ST token being spoken and the beginning time of its produced TT equivalence. After eliminating the outliers that fall outside the range of 0.05 quantile and 0.95 quantile, we find that the mean value of EVS in our datasets is 6002 ms. This indicates that there was generally six seconds' lag before the participants produced the interpretation of an ST token. The EVS value ranges from 871 ms to 20692 ms, with a standard deviation of 5161 ms. The histogram on the upper portion of Figure 2 demonstrates that EVS (on the x-axis) is mostly relatively short, with a noticeable gap and an unbalanced distribution. When aggregating the mean value of EVS for each participant, we find a significant variation across participants. Participants on the lower end of the spectrum spent on average 3 seconds before producing the interpretation, whilst participants on the upper end spent about 11 seconds doing so. This substantial inter-subject heterogeneity is consistent with earlier studies on EVS in SI (Lamberger-Felber, 2001; Timarová, Dragsted, and Hansen, 2011), which demonstrates that individual characteristics are one factor affecting EVS.

Additionally, we measure IVS by the interval between the time a source token is first fixed and the beginning time of its corresponding TT is produced. We find that the mean value of IVS in our datasets is 2187 ms. This indicates that participants generally attend to the visual input about 2 seconds before producing their interpretation. This reading-ahead tendency corroborates the existing research on IVS in STT (Agrifoglio, 2004; Huang, 2011; Chen, 2015). The IVS value varies from -14623 ms to 21143 ms, with a standard deviation of 8411 ms. The histogram on the right side of Figure 2 shows that IVS (on the y-axis) is generally balanced and rather concentrated. When we examine the average IVS for each participant, we observe that some interpreters read the source tokens about 14 seconds before producing the oral translation, whereas others read the source tokens around a second after making the oral translation.

How do visual attention and auditory attention relate to one another? According to the regression plot in the center of Figure 2, there is no correlation between EVS and IVS across the entire datasets ($r=-0.05$, $p<.01$). In the context of multimodal processing, specifically, does visual attention predict or verify the oral signal, or do they function as independent inputs?

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

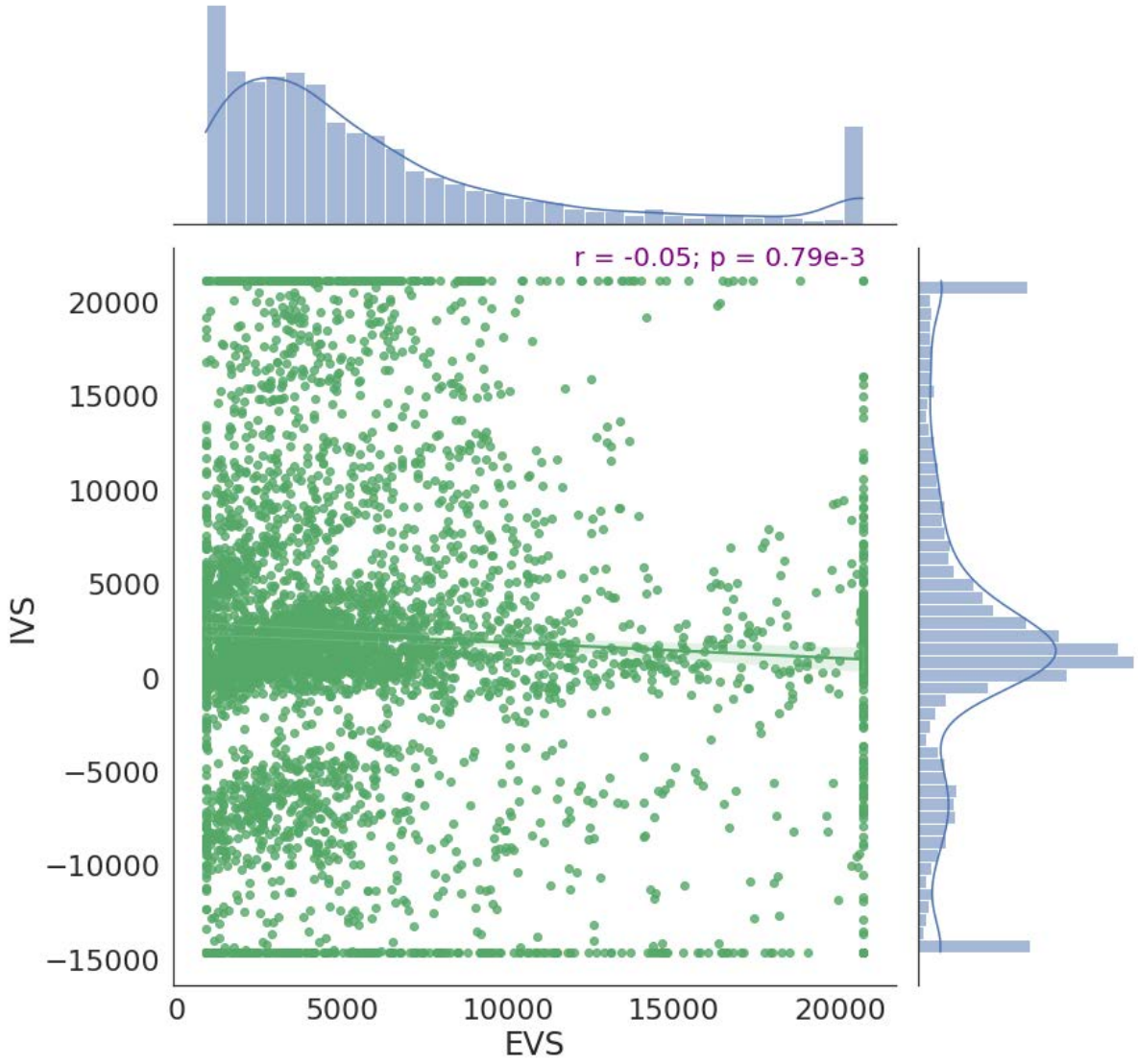

Figure 2. Joint Plot of Ear-Voice Span and Eye-Voice Span

3.2 Ear-eye span (EIS)

We compute the ear-eye span (EIS) to investigate the relationship between auditory and visual inputs. That is, EIS indicates the general tendency of interpreters' attention toward the ear or the eye. For each source text token (SToken), we establish the EIS as the temporal difference between the time stamp at which this SToken was spoken (STime) and the time stamp at which the first fixation was recorded on this SToken (FFTime). As shown in the following Figure 3, the average EIS for all the STokens in our datasets is μ = -4400 ms. This indicates that in general, the interpreters attend first to the auditory input and their earliest attention to the visual input comes on average 4 seconds after hearing the speech. Therefore, we see an overall ear-lead-eye pattern for the interpreters in our experiment, which is consistent with the conclusions from Seeber et al. (2020) and Ma & Cheung (2020).

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

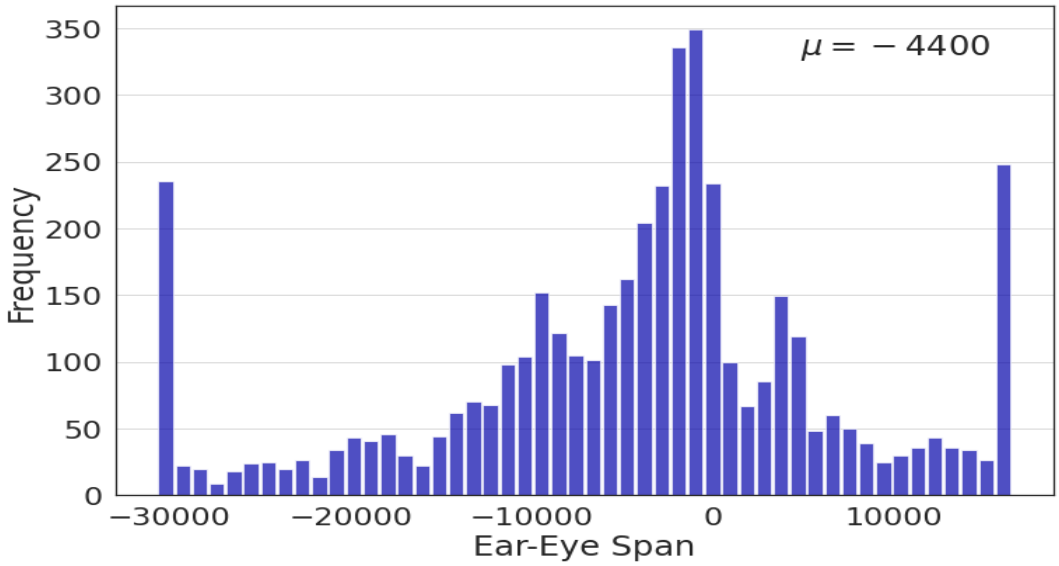

Figure 3. Distribution of Ear-Eye Span

3.3 Attention patterns of interpreters in SIMTXT

We look further into the ratio of interpreters' EVS to IVS to determine how the allocation of attention between auditory and visual inputs relates to their interpreting output. We calculate the interpreters' ear-eye ratio (EIR) by dividing the absolute value of EVS by the absolute value of IVS. We then aggregate the average EIR for each participant and map them to the entire distribution of the EIR. According to our results, there are three types of attention patterns: the ear-dominant interpreter (ED), the eye-dominant interpreter (ID), and the ear-eye-balanced interpreter (EIB).

3.3.1 Ear-dominant interpreter (ED)

Four out of the nine participants fall into the ED category. There is a tendency for the ED-type interpreters to produce a translation much closer to the time at which the SToken is heard than at the time when the SToken was first fixated. Therefore, ED's average EIR would be smaller than the mean EIR of the entire datasets. Figure 4 visualizes one ED's translation process for one segment of a SIMTXT session. The translation progression graph shows how auditory input, visual input, and spoken output interact (Carl & Jakobsen, 2009). The left vertical axis lists the STokens by their index numbers in the source text while the vertical axis on the right shows the transcripts of their spoken Chinese translations. The horizontal axis indicates the timeline in which the translations emerge. The orange circles represent the temporal progression of the spoken ST, blue dots represent the temporal progression of the participant's eye movement on the ST, and the Chinese characters in black represent the temporal progression of the participant's spoken output. All points on the same horizontal line show the temporal dynamics of the listening, reading, and interpreting activities for the translation of the SToken at that line. We can see that for ED, the stream of spoken output (in the black Chinese characters) is frequently closer to the stream of auditory input (in the orange circles) than the stream of visual input from the ST (in the blue dots).

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

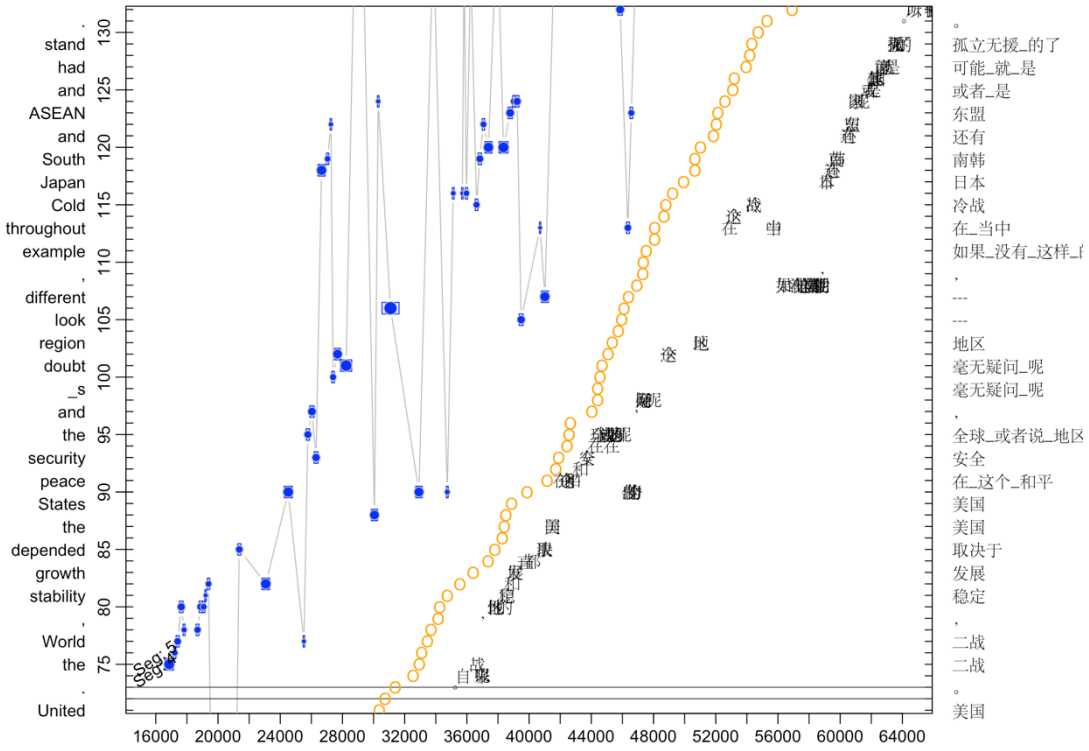

Figure 4. An Example of Ear-Dominant Interpreter's Progression Graph

The ST of the interpreted segment in Figure 4 is:

> "Since the Second World War, regional stability and growth have depended on the United States guaranteeing peace and security in the region and there's no doubt that region would look very different if, for example, throughout the Cold War Japan, South Korea and the ASEAN countries and others had to stand alone."

This segment comprises 58 tokens in total and makes up 42% of the session's ST. According to Figure 4, this segment's auditory input begins at approximately 31000 ms and ends at approximately 56000 ms, with a spoken ST length of about 25 seconds. It takes about 30 seconds for the interpreter to produce the spoken TT. The interpretation of this segment starts at 35000 ms and ends at 65000 ms, which are between 4 and 9 seconds behind the auditory input, respectively. As time evolves, we observe that the delay between auditory input and output gradually increases towards the end, which indicates an excessive working memory load to the interpreter.

Moreover, all fixations on this segment take place prior to the corresponding auditory input and spoken output. The first fixation occurs at 16000 ms, and the last occurs at 48000 ms. This indicates that the participant has a reading-ahead tendency for contextual planning. For instance, when the audio of this segment starts, the participant is reading around "the **United** States" (about 14 tokens ahead of the auditory input), and when the spoken output of this segment starts, the participant is reading around "the Cold **War** Japan" (about 42 tokens ahead of the spoken output). From approximately 46000 ms

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

onwards, when the gap between auditory input and spoken output becomes larger, there are no fixations around the source words for which the translations have just been spoken. This suggests that the participant does not read the written ST to confirm the information, and that the auditory input typically has a greater direct impact on the participant's spoken output than the visual input.

3.3.2 Eye-dominant interpreter (ID)

Two participants are classified as ID. In contrast to ED, an ID-type interpreter speaks the translation of an SToken significantly closer to the time at which the gaze initially fixates on the SToken than at the time at which the SToken is heard. Therefore, ID's average EIR would be smaller than the total mean EIR. The translation process for one ID is shown in Figure 5. We can see that for ED, the spoken output (in the black Chinese characters) and the visual input (in the blue dots) are commonly relatively close to one another.

For the same ST segment as in Figure 4, it takes about 27 seconds for this participant in Figure 5 to produce the spoken TT. The interpretation of this segment starts at 36000 ms and ends at 63000 ms, which are 5 and 7 seconds behind the auditory input, respectively. We notice that the gap between auditory input and output gradually widens with time, though it is not as big as the ED's. Furthermore, all fixations on this segment appear after the auditory input. The first fixation occurs at 38000 ms, and the last occurs at 63000 ms. Some of the fixations are in sync with the spoken output, while others lag behind. This suggests that the participant reads the written ST to gather information or to verify the information.

Take the purple box in Figure 5 as an example. After the segment's auditory input has ended, the fixations almost match or slightly lag behind the source words for which the translations have just been spoken. This suggests that the written ST serves as a supplemental source for the interpreter to assist information processing and prevent working memory overload. There is a gaze-to-word mis-mapping in the red box in Figure 5. The actual gazes in this period are fixated on the line below, which are only a few pixels away, but appear here as a mapping far off in another line. Therefore, we see that the visual input overall has a stronger direct impact on the participant's spoken output than the auditory input.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

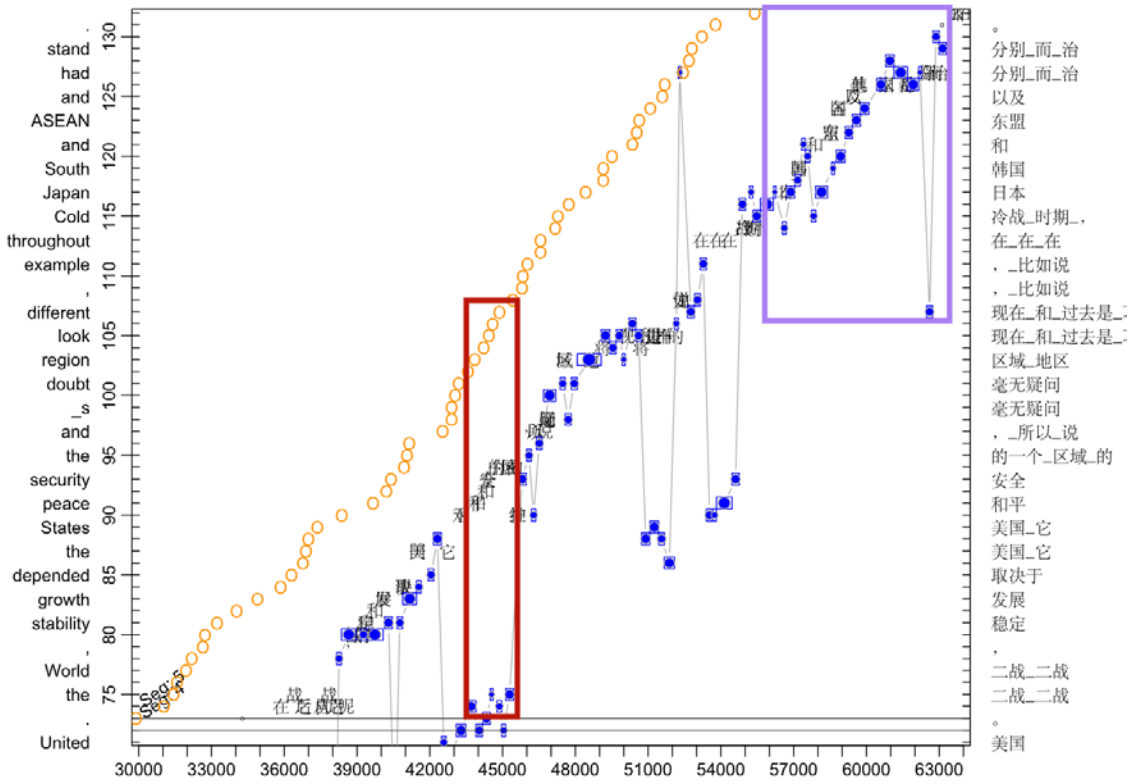

Figure 5. An Example of Eye-Dominant Interpreter's Progression Graph

3.3.3 Ear-eye-balanced interpreter (EIB)

In addition, three participants are classified as EIB. Different from the aforementioned patterns, the time at which an EIB's interpretation of a certain SToken is produced is relatively concomitant with both the time at which the SToken is first fixated and heard. Therefore, the average EIR of EIB would be close to the mean EIR of the entire datasets. The translation process for one EIB is shown in Figure 6 for the same segment as above. We can see that for EIB, the auditory input (in the orange circles) and the visual input (in the blue dots) are often adjacent to each other, suggesting that they have a similar effect on the spoken output.

For the same ST segment as in Figure 4 and Figure 5, the interpretation in Figure 6 starts at 38000 ms and ends at 58000 ms. The participant generates the spoken TT in around 20 seconds, which is the fastest of the three types of interpreters, but the alignment groups on the vertical axis suggest that there are some omissions in the TT. The progression graph for the EIB shows that, for the most part, the difference between auditory input and output does not change over time.

When we look closer into the process data, we see that for the period in the red box (31000 ms – 40000 ms), which is at the beginning of the segment, the interpreting output has about 10 tokens' omission. The participant scans through the written ST until the second comma of the segment, which is prior to both the auditory input and spoken output. This indicates that the participant attends the

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

written ST to anticipate the auditory input during this period in order to keep up with the pace of the audio. Most of the fixations on this segment (40000 ms – 55000 ms) appear in between the associated auditory input and the spoken output. This suggests that the participant generally reads the written ST to verify the information gathered from the auditory input during this period. The gaze data in the period in the purple box (55000 ms – 58000 ms) might be a drift in the word-to-gaze mapping, which should have been in line with the previous period that the participant has a linear pattern in both hearing and reading behaviors. Overall, Figure 6 shows that the participant generally pays attention to auditory and visual inputs similarly when interpreting this segment.

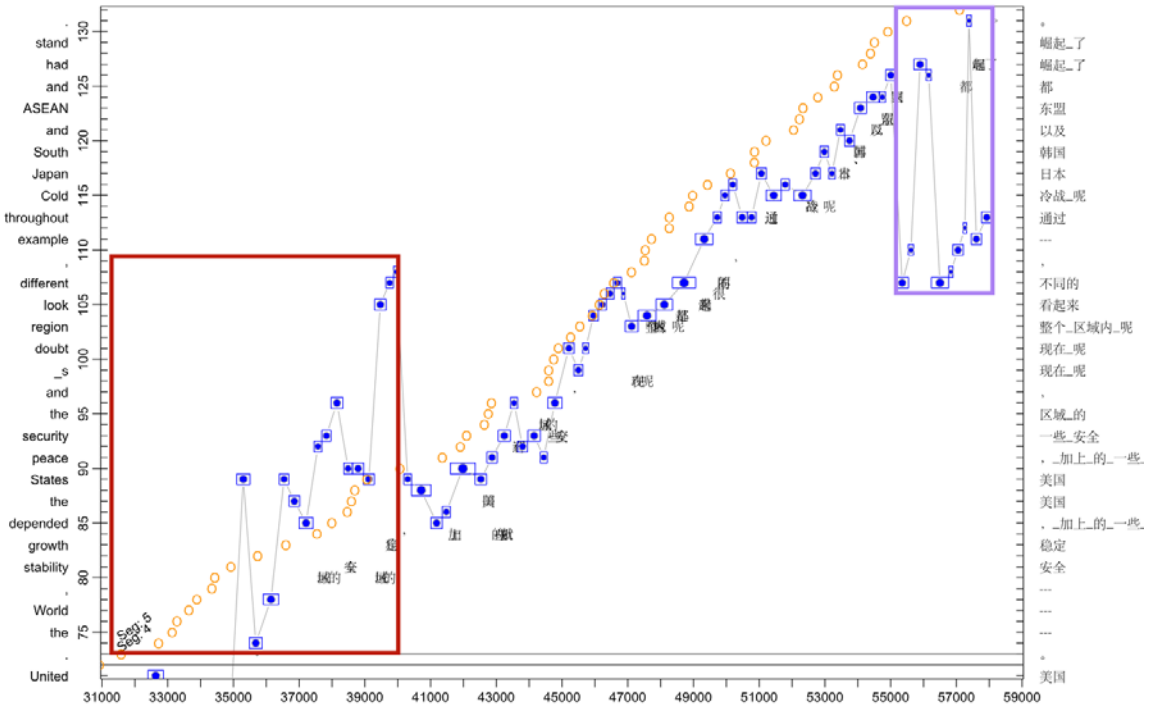

Figure 6. An Example of Ear-Eye-Balanced Interpreter's Progression Graph

4. Analysis of Quality Assessment Data and its Relationship with Attention Patterns

The taxonomy of attention patterns reflects interpreting strategies to address and avoid cognitive overload. Specifically, the ED relies on auditory input, whilst the ID focuses more on visual input, and the EIB attends to both input channels during their interpretations. Here, we investigate how these strategic decisions affect translation quality. We conducted manual assessment to measure the translation quality and distributed the segments to be annotated in such a way that we have three annotations per segment. Since we had limited resources for the annotations, only a subset of 195 segments were annotated.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

4.1 Manual assessment of SIMTXT output

Three professional translators were recruited to annotate translation errors in the English to Chinese translations. Guidelines for translation error typology were provided to the annotators, based on the harmonized Multidimensional Quality Metrics (MQM)-Dynamic Quality Framework (DFQ) of the Translation Automation User Society (TAUS) (Lommel, Uszkoreit, and Burchardt, 2014). Three categories were used to classify errors:

- Accuracy (which includes "Addition", "Omission", "Mistranslation", "Over-translation", "Under-translation", and "Untranslated text"),
- "Fluency" (which includes "Punctuation", "Spelling", "Grammar", "Grammatical register", and "Inconsistency"),
- "Style" (which includes "Awkward", "Inconsistent Style", and "Unidiomatic").

The annotators further classified each error as "Major" or "Minor", depending on its seriousness. The error annotation interface for the annotators in this experiment is described in Zou et al. (2021). Once an error was discovered, annotators were requested to perform word-level alignment between the erroneous TT and their associated ST before assigning the error to the alignment group (AG). They were not, however, requested to perform an alignment whenever an error was found and regarded as an addition or omission because addition only happened in the TT and omission only happened in the ST.

Before engaging in the actual annotation process, the annotators evaluated a mock session with a brief text that contained three intentionally introduced errors to ensure the reliability of the annotations. All annotators spotted the introduced errors, but the labels they assigned to the errors varied. Similar results were found in the actual annotations as well. We take a segment from this mock session to illustrate our error annotation and assessment schema.

There are 16 words in the English ST and 12 words in the Chinese TT. Our three annotators used different labels and colors to mark the errors (we separate the Chinese characters by whitespace to indicate the tokenization adopted by the Translation Process Research Database of Center for Research and Innovation in Translation and Translation Technology (CRITT-TPRDB) hereafter). Accuracy errors are shown in purple, fluency errors in green, and style errors in pink. When using the same color code, critical errors are denoted by a darker shade and minor errors by a lighter shade. Unaligned accuracy errors, such as omissions and additions, are highlighted in pink for critical errors and green for minor ones.

As illustrated in Figure 7a, Annotator A considered the rendition of "our" into "我们" (we) as a minor style error, and the rendition of "make us natural partners" into "一起 来到 这个 会议 当中" (come to this meeting together) as a minor accuracy error. Annotator A also identified the omission of translation for "in our region and beyond" a critical omission error. Therefore, Annotator A's manual assessment of this segment involves overall 16 erroneous words on both ST and TT sides, including 14 words with accuracy errors, 2 words with style errors, 5 words with critical errors, and 11 words with minor errors.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

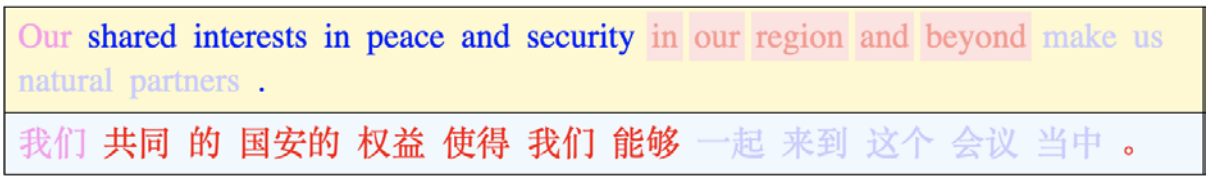

Figure 7a. Manual assessment by Annotator A

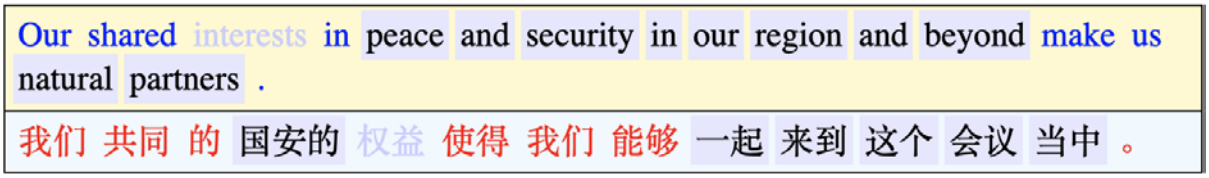

Figure 7b. Manual assessment by Annotator B

Figure 7c. Manual assessment by Annotator C

As shown in Figure 7b, for Annotator B, the rendition of "interests" into "权益" (rights and interests) was regarded a minor accuracy error, "peace and security in our region and beyond" into "国安的" (national security) a major accuracy error, and "natural partners" into "一起 来到 这个 会议 当中" (come to this meeting together) a major accuracy error. Therefore, Annotator B's manual assessment of this segment involves overall 18 erroneous words on both ST and TT sides, all of which are accuracy errors, with 16 words of critical errors, and two words of minor errors.

However, Annotator C (Figure 7c) saw the rendition from "peace and security" into "国安的" (national security) as a minor accuracy error, whereas the rendition from "make us natural partners" into "能够 一起 来到 这个 会议 当中" (come to this meeting together) as a major accuracy error. Additionally, "in our region and beyond" was labeled as a major omission as the interpreter failed to render it in the TT, which was annotated identically with Annotator A. Therefore, Annotator C's manual assessment of this segment involves overall 19 erroneous words on both ST and TT sides, all of which are accuracy errors, with 15 words of critical errors, and four words of minor errors.

Although the three annotators made somewhat different decisions regarding alignment and error identification, they reached a similar number of errors under most error labels, as shown in Table 2. From this example, we might infer that annotators do not always agree on error identification in terms of each word and alignment, but a similarity of total errors in terms of the occurrences under each label may suggest rather good inter-rater agreement.

To account for such classification discrepancies, we aggregate the annotated errors in different ways. All translation errors, regardless of category, were combined under the label of "Any" error for the

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

purposes of this study. As a result, there are six different labels of errors altogether in this study and we count both source and target words in an AG that involve each of the six labels of errors. The frequency of each label of error established the following numerical connection:

$$\text{Any} = \text{Accuracy} + \text{Fluency} + \text{Style} = \text{Critical} + \text{Minor} \quad (1)$$

Table 2. Number of Errors under each Error Label for Example (1) by Three Annotators

ST			TT			
Our shared interests in peace and security in our region and beyond make us natural partners. **(16 words)**			我们 共同的 国安的 权益 使得 我们 能够 一起 来到 这个 会议 当中。 **(12 words)** Gloss: Our shared rights and interests of national security allow us to come to this meeting together.			
Annotator	**Any**	**Accuracy**	**Fluency**	**Style**	**Critical**	**Minor**
A	16	14	0	2	5	11
B	18	18	0	0	16	2
C	19	19	0	0	15	4

4.2 Inter-rater agreement

We examined the inter-rater agreement of the three annotators in this study using weighted Fleiss' Kappa. The weighted Fleiss' Kappa scores (Fleiss & Cohen, 1973) in Table 3 show that, for all the segments in these datasets, the three annotators highly agree on practically all error labels with the exception of style errors. Moreover, annotators tend to agree more on the accuracy and critical errors than other categories of errors (Zou, Saeedi, and Carl, 2022).

Table 3. Scores of Weighted Fleiss' Kappa and SEG-EA for all the Error Labels

Error Label	Weighted Fleiss' Kappa	SEG-EA
Any	0.853	0.279
Accuracy	0.915	0.201
Fluency	0.815	0.068
Style	0.739	0.008
Critical	0.910	0.194
Minor	0.866	0.084

We further measured the Error Average (SEG-EA) based on the total errors annotated by all the annotators on both ST and TT on the segment level. Since the lengths of the ST and TT tokens vary depending on the segment, we normalize the translation error counts by the tokens on both the ST side and the TT side using the equation below:

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

$$\text{SEG-EA} = \frac{Total\ ST\ errors + Total\ TT\ errors}{Total\ ST\ tokens + Total\ TT\ tokens} \quad (2)$$

The SEG-EA over all error labels in the datasets is 0.279, which indicates that about 28% of the words contain an error annotation. Approximately 20% of the words in the datasets have accuracy and critical errors, according to the SEG-EA scores for accuracy (0.201) and critical errors (0.194). These two errors are substantially more common or evident than the fluency (0.068), style (0.008), and minor (0.084) errors.

The spearman's correlation coefficients between SEG-EA scores of different error labels indicate the relationship between the frequency of various categories of translation errors. The results in Table 4 show that accuracy errors very strongly and positively correlate with both the total number of errors ("any") (rs=0.9, p<.01) and critical errors (rs=0.937, p<.01), but they moderately and negatively correlate with fluency (rs=-0.38, p.<01) and minor errors (rs=-0.394, p.<01). Fluency errors strongly and positively correlate with minor errors (rs=0.548, p<.01), while style errors correlate weakly with minor errors (rs=0.211, p<.01) (Dancey & Reidy, 2017). In line with the findings in previous studies (Carl & Báez, 2019; Zou et al., 2022), our results suggest that accuracy errors are more likely critical errors, and fluency errors are more frequently minor errors in our datasets.

Table 4. Spearman's Correlation between SEG-EA Scores

	Any	**Accuracy**	**Fluency**	**Style**	**Critical**	**Minor**
Any	1	0.9**	0.05	-0.073	0.946**	-0.151
Accuracy	0.9**	1	-0.38**	-0.112	0.937**	-0.394**
Fluency	0.05	-0.38**	1	-0.095	-0.135**	0.548**
Style	-0.073	-0.112	-0.095	1	-0.132*	0.211**
Critical	0.946**	0.937**	-0.135**	-0.132*	1	-0.462**
Minor	-0.151	-0.394**	0.548**	0.211**	-0.462**	1

**. Correlation at 0.01 *. Correlation at 0.05

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

4.3 Relationship between attention patterns and SIMTXT quality

In this section we address the relationship between different types of attention patterns (ED, ID, EIB) and the translation quality of SIMTXT. By examining the distribution of each interpreter's SEG-EA score for the six error labels (i.e., Any, Fluency, Accuracy, Style, Critical, Minor), we are able to relate the translation quality of various interpreters to attention patterns (i.e., ED, ID, and EIB). Figure 8 illustrates our findings that, regardless of error type, EIB creates the most errors, followed by ED and ID ($p<.01$). Similar trend is observed about accuracy error, with ID interpreters producing the fewest accuracy errors, followed by ED and EIB interpreters. We suspect that in the case of bimodal processing (EIB), the interpreter's decision to allocate an equal amount of attention to both the audio and visual inputs could have a detrimental influence on the accuracy of the translation. Extreme cognitive load on reading, listening and coordination effort may be the cause of this. When the interpreter focuses largely on the visual input (ID), they will achieve an accuracy facilitation effect. Although the information provided in both modalities is identical, the written text can be less taxing on working memory and more reliable than the verbal text since it does not disappear with time. This may be a reason why we see higher accuracy for interpreters who choose to focus more on the visual input.

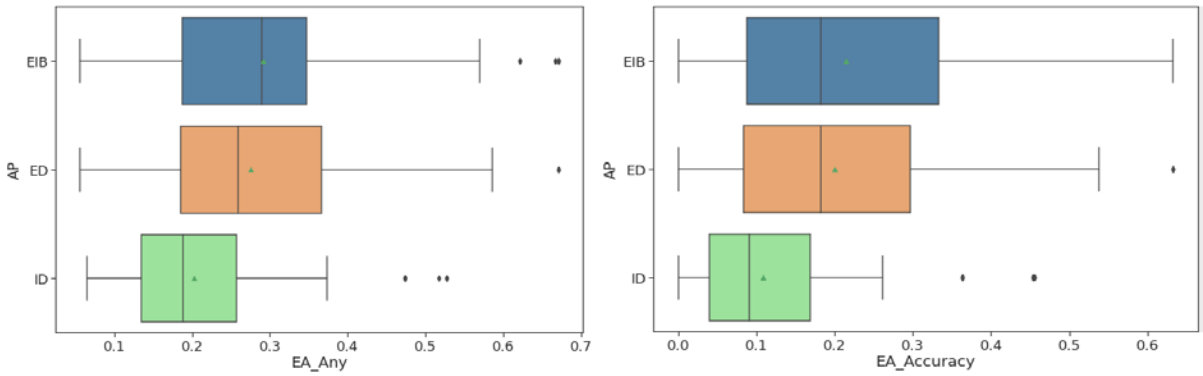

Figure 8. SEG-EA Scores for Any Error and Accuracy Error across Different Attention Patterns

In terms of fluency error, we see that ED-type interpreters produce the least errors, followed by EIB and ID ($p<.01$), as shown in Figure 9. One possible explanation for this result would be the differences in the grammatical structure between English and Chinese (Wang & Gu, 2016). ED-type interpreters follow the flow of the source speech more closely, which usually involves a higher degree of deverbalization, and thus may contribute to their higher level of fluency. ID-type interpreters, in contrast, would experience the highest cognitive load as a result of the greatest visual interference from the written text. They would need a significantly higher amount of reformulation, planning and problem-solving strategies, thus result in less fluent performance than both EIB and ED.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

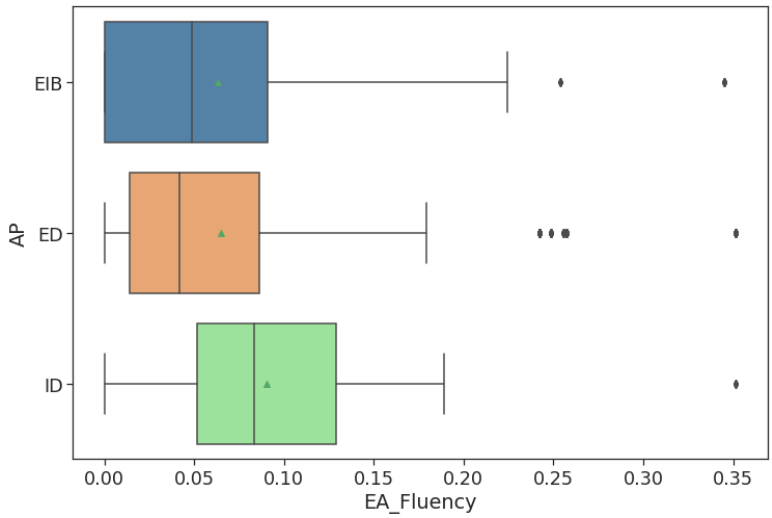

Figure 9. SEG-EA Scores for Fluency Error across Different Attention Patterns

5. Conclusion

SIMTXT is a hybrid form of spoken translation between simultaneous interpreting (SI) and sight translation (STT). In this mode, an interpreter converts an oral ST into an oral TT in real time with the presence of a written ST. This interpreting modality is generally considered to be cognitively more demanding than either SI or STT since the interpreter needs to process the ST both visually and auditorily at the same time. In this study, we investigate how visual and auditory information is processed concurrently (EIB) or independently (ED, ID) and how visual and auditory attention is allocated proportionally by interpreters during SIMTXT. We examine the relationship between different interpreters' attention patterns and translation quality using the IMBst18 and IMBi18 datasets that are accessible at CRITT TPR-DB.

Nine professional interpreters were hired for SIMTXT from English into Chinese. Every participant interpreted six texts with the written English ST provided. The nine interpreters produced 297 segments altogether. These six texts were successive sections from the beginning part of (the same) live audio recording of a political speech (about one minute long for each section) and the transcriptions of these six sections of audio recording (about 150 words long for each transcription) with a total of 871 words. During the experiment, each interpreter had two input sources in parallel, visual and auditory, and produced one spoken translation. The transcriptions of the audio recordings were displayed in Translog-II, and the interpreters' eye movement data were recorded with a Tobii TX300 eye-tracker, whereas the audio input of the speech was replayed via a headset, and the audio recordings of the interpreter's SIMTXT output (i.e., their spoken translations in Chinese) were collected with the original speech by Audacity.

To operationalize patterns of visual or auditory attention for different interpreters, we compute the eye-voice span, ear-voice span and ear-eye span, and classify interpreters as ear-dominant (ED), eye-dominant (ID), or ear-eye-balanced (EIB). We then assess the impact of dominance patterns on translation quality. We hired three professional translators to manually annotate translation errors in

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

the English-Chinese translations based on an MQM-derived error taxonomy at the segment level. The results show that annotators do not always agree on error identification for each word but arrive at a good inter-rater agreement on a segment level. We find higher inter-rater agreement on accuracy and critical errors, as compared to fluency and minor errors.

Moreover, we see that eye-dominant (ID) interpreters tend to produce less accuracy errors, while ear-dominant (ED) interpreters tend to have less fluency issues. However, the interpreter's decision to allocate an equal amount of attention to both the audio and visual inputs (EIB) could have a detrimental influence on the overall quality of the translation. These findings shed light on translation strategies for optimizing the utilization of available resources to prevent cognitive overload, leading to possibly better interpreting performance.

Understanding allocation patterns of the interpreters can provide more insight into decision-making during the SIMTXT process and the interaction between translation process and translation product. This research has conducted an experiment only on a limited sample size, professional interpreters, and one language combination of interpreting (i.e., English into Chinese). Future research can analyze additional language pairs, text types, with more participants, or replicate the experiment by comparing the translation process of novice interpreters to experienced interpreters. Such studies might yield intriguing results and offer recommendations for translation and interpreting pedagogy.

Declarations and Acknowledgement:

The authors declare there is no conflict of interest.

The manual assessment in this study was organized by Dr. Lucas Nunes Vieira and funded by the IMPETUS project (https://gtr.ukri.org/projects?ref=ES%2FS014446%2F1).

References

Agrifoglio, Marjorie. 2004. "Sight translation and interpreting: A comparative analysis of constraints and failures." Interpreting 6 (1): 43–67. https://doi.org/10.1075/intp.6.1.05agr

Brugman, H., Russel, A., & Nijmegen, X. (2004, May). Annotating Multi-media/Multi-modal Resources with ELAN. In *LREC* (pp. 2065-2068).

Čeňková, I. (2015). Sight interpreting/translation. In F. Pöchhacker, *Encyclopedia of Interpreting* (pp. 374-375). London: Routledge.

Cammoun, R., Davies, C., Ivanov, K., & Boris, N. (2009). *Simultaneous Interpretation with Text: Is the Text «Friend» or «Foe»? Laying Foundations for a Teaching Module*. Masters' dissertation. FTI, University of Geneva.

Carl, M., Dragsted, B., & Jakobsen, A. L. (2011). A taxonomy of human translation styles. Translation journal, 16(2), 155-168.

Carl, M., & Jakobsen, A. L. (2009). Towards statistical modelling of translators' activity data. *International Journal of Speech Technology*, *12*(4), 125-138.

Carl, M. (2012). Translog-II: A program for recording user activity data for empirical reading and writing research. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)* (pp. 4108-4112).

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

Carl, M & Yamada, M. (2017). Tutorial on Multimodal Integration of Written and Spoken Translation Production. Presented at the 4th International Conference on Cognitive Research on Translation and Interpreting. Beijing.

Carl, M., & Báez, M. C. T. (2019). Machine translation errors and the translation process: a study across different languages. *Journal of Specialised Translation*, 107-132.

Chen, Jui-Ching. 2015. "Sight translation." In The Routledge Handbook of Interpreting, ed. by Holly Mikkelson and Renee Jourdenais, 144–53. New York: Routledge.

Chmiel, A., & Mazur, I. (2013). Eye tracking sight translation performed by trainee interpreters. In C. Way, S. Vandepitte, & R. M. Bartłomiejczyk.

Chmiel, A., & Lijewska, A. (2019). Syntactic processing in sight translation by professional and trainee interpreters: Professionals are more time-efficient while trainees view the source text less. *Target. International Journal of Translation Studies*, *31*(3), 378-397.

Chmiel, A., Janikowski, P., & Cieślewicz, A. (2020). The eye or the ear?: Source language interference in sight translation and simultaneous interpreting. *Interpreting*, *22*(2), 187-210.

Chmiel, A., & Lijewska, A. (2022). Reading patterns, reformulation and eye-voice span (IVS) in sight translation. Translation and Interpreting Studies.

Feng, J., Carl, M., Zhu, Y., Chen, S., & Chen, J. (2020). Eye-Voice Span in Sight Interpreting: Evidence from Both Process and Product. *Translation in Transition*, 25.

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

Huang, Chih-Chieh. 2011. "Tracking eye movements in sight translation." Unpublished M.A. thesis, Taiwan: National Taiwan Normal University. http://rportal.lib.ntnu.edu.tw/

Inhoff, A. W., Solomon, M., Radach, R., & Seymour, B. A. (2011). Temporal dynamics of the eye–voice span and eye movement control during oral reading. *Journal of Cognitive Psychology*, *23*(5), 543-558.

Ivanov, K., Davies, K., & Naimushin, B. (2014). Teaching simultaneous interpreting with text. *Fighting the Fog of Multiculturalis. A Festschrift in Honour of Irina S. Alekseeva*, 48-61.

Jiménez Ivars, M. A. (1999). *La traducción a la vista. Un análisis descriptivo* (Doctoral dissertation, Universitat Jaume I).

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel*. Naval Technical Training Command Millington TN Research Branch.

Lamberger-Felber, Heike. 2001. "Text-oriented research into interpreting – Examples from a case-study." HERMES – Journal of Language and Communication in Business 14 (26): 39–63. https://doi.org/10.7146/hjlcb.v14i26.25638

Lambert, S. (2004). Shared attention during sight translation, sight interpretation and simultaneous interpretation. *Meta: Journal des traducteurs/Meta: Translators' Journal*, 294-306.

Lambert, S. (2004). Shared attention during sight translation, sight interpretation and simultaneous interpretation. *Meta: Journal des traducteurs/Meta: Translators' Journal*, 294-306.

Lommel, A., Uszkoreit, H., & Burchardt, A. (2014). Multidimensional quality metrics (MQM): A framework for declaring and describing translation quality metrics. Revista Tradumàtica: tecnologies de la traducció, (12), 455-463.

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

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

Pöchhacker, F. (2004). I in TS: On partnership in translation studies. In *Translation Research and Interpreting Research* (pp. 104-115). Multilingual Matters.

Pöchhacker, F. (2016). *Introducing Interpreting Studies.* London: Routledge.

Savenkov, K., & Lopez, M. (2022). The State of the Machine Translation 2022. (pp. 32-49). Proceedings of the 15th Biennial Conference of the Association for Machine Translation in the Americas (Volume 2: Users and Providers Track and Government Track).

Seeber, K. G. (2017). Multimodal processing in simultaneous interpreting. *The handbook of translation and cognition*, 461-475.

Seeber, K. G. & Delgado Luchner, C. M. (2020). Simulating simultaneous interpreting with text: From training model to training module. In I. H. M. D. Rodríguez Melchor, *The Role of Technology in Conference Interpreter Training* (pp. 70–86). Peter Lang.

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., & Motta, M. (2007). Syntacrobatics: Quality and reformulation in simultaneous-with-text. *Interpreting*, *9*(2), 199-230.

Setton, R. & Dawrant, A. (2016). *Conference Interpreting – A Complete Course.* Amsterdam: Benjamins Translation Library.

Setton, R. (2015). Simultaneous with Text. In F. Pöchhacker, *Routledge encyclopedia of interpreting studies* (pp. 385-386). London: Routledge.

Spence, C. (2009). Explaining the Colavita visual dominance effect. *Progress in brain research*, *176*, 245-258.

Sun, S., & Shreve, G. M. (2014). Measuring translation difficulty: An empirical study. *Target. International journal of translation studies*, *26*(1), 98-127.

Timarová, Šárka, Barbara Dragsted, and Inge Gorm Hansen. 2011. "Time lag in translation and interpreting: A methodological exploration." In Methods and Strategies of Process Research: Integrative approaches in Translation Studies, ed. by Cecilia Alvstad, Adelina Hild, and Elisabet Tiselius: 121–46. Amsterdam: John Benjamins. https://doi.org/10.1075/btl.94.10tim

Wang, B., & Gu, Y. (2016). An evidence-based exploration into the effect of language-pair specificity in English-Chinese simultaneous interpreting. Asia Pacific Translation and Intercultural Studies, 3(2), 146-160.

Xiao, R. (2010). How different is translated Chinese from native Chinese?: A corpus-based study of translation universals. *International Journal of Corpus Linguistics*, *15*(1), 5-35.

Zheng, B., & Zhou, H. (2018). Revisiting processing time for metaphorical expressions: an eye-tracking study on eye-voice span during sight translation. Foreign language teaching and research., 50(5), 738-753.

Zou, L., Carl, M., Mirzapour, M., Jacquenet, H., & Vieira, L. N. (2021). AI-Based Syntactic Complexity Metrics and Sight Interpreting Performance. In *International Conference on Intelligent Human Computer Interaction* (pp. 534-547). Springer, Cham.

Zou, L., Carl, M., Yamada, M., & Mizowaki, T. (2022). Proficiency and External Aides: Impact of Translation Brief and Search Conditions on Post-editing Quality. (pp. 60-74). Proceedings of the 15th biennial conference of the Association for Machine Translation in the Americas (Workshop 1: Empirical Translation Process Research).

Zou, L., Saeedi, A., & Carl, M. (2022). Investigating the Impact of Different Pivot Languages on Translation Quality. (pp. 15-28). Proceedings of the 15th biennial conference of the Association for Machine Translation in the Americas (Workshop 1: Empirical Translation Process Research).

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing

About the Authors:

Longhui Zou completed her M.A. in Interpreting and Translation Studies from Wake Forest University in the US. She is currently a doctoral fellow at Kent State University. Longhui focuses on Chinese and English translation and interpreting, as well as machine translation and post-editing, translation and interpreting processes, and translation technologies. Her most recent effort intends to collect new and analyze legacy keylogging and eye tracking data to investigate translation processes and identify higher-order cognition based on behavioral patterns of monitoring activity observed in logged translation sessions.

Dr. Michael Carl is a Distinguished Professor at Kent State University, USA and Director of the Center for Research and Innovation in Translation and Translation Technology (CRITT). He has studied Computational Linguistics and Communication Sciences in Berlin, Paris and Hong Kong and obtained his PhD degree in Computer Sciences from the Saarland University, Germany. He has worked and published for more than 25 years in the fields of translation studies, machine translation and natural language processing. His current research interest is related to the investigation of human translation processes and interactive machine translation.

Dr. Jia Feng currently works as an associate professor of Translation Studies at the School of Foreign Languages, Renmin University of China. Jia has been conducting research on cognitive translation studies and translation processes since 2012, when she started her PhD at Beijing Foreign Studies University. Her current research interests are translation directionality and translation styles.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
© 2022 All Terrain Publishing