How does training shape English-Chinese sight translation behaviour?: An eyetracking study

Ho, C.-E., Chen, T.-W., & Tsai, J.-L. (2020). How does training shape English-Chinese sight translation behaviour?: An eyetracking study. *Translation, Cognition & Behavior*, *3*(1), 1-24.
https://doi.org/10.1075/tcb.00032.ho

Published in:
Translation, Cognition & Behavior

Document Version:
Peer reviewed version

How does Training Shape English-Chinese Sight Translation Behaviour?

An eyetracking study

Chen-En Ho[1], Tze-Wei Chen[2] & Jie-Li Tsai[3]

This study aims to investigate the cognitive nature of sight translation by analysing the reading behaviour in the process and the output. Three tasks, silent reading, reading aloud, and sight translation, were included in the experiment, in which two groups of participants—interpreting students and untrained bilinguals—took part. The results show that the two groups were almost identical in the first two tasks, further substantiating the similarity of their language command, but were drastically different in how they tackle sight translation. Interpreting students provided much more accurate, fluent, and adequate renditions with much less time and fewer fixations. However, their efficiency at information retrieval was statistically similar to that of the untrained bilinguals', showing that the students were more efficient by being more "economical" during reading rather than by reading ahead faster, as some would intuitively expect. Chunking skills seem to have been at play behind their remarkable performance.

Keywords: sight translation; cognitive process; reading ahead; chunking/segmentation; pausing behaviour; interpreter training

1. Introduction

Sight translation (SiT, henceforth) has been readily embraced in the interpreting classroom because it overlaps with other modes of interpreting (Li 2014). SiT also has potential to help reduce the cognitive load of the interpreter by helping them chunk longer sentences into smaller meaningful units, or by facilitating preparation for interpreting assignments (Weber 1990; Viaggio 1995; Agrifoglio 2004). SiT does play an important role in professional, public service interpreting (ISO 2014), where people's basic rights and well-being depend heavily on the quality of interpretation. Interpreters are often prompted to tackle written exhibits in court or technical reports during medical consultations without preparation. Proper SiT skills are thus required for the interpreter to adequately render all messages (Sampaio 2007). Accordingly, the importance of SiT has been highlighted by practitioners and professional organisations, such as the National Association of Judiciary Interpreters & Translators (2006), the National Council on Interpreting in Health Care (2009), the International Medical Interpreters Association (2009) and the International Organization for Standardization (2014).

Awareness of the SiT benefits has been recently on the rise. Teamed up with voice recognition technology, SiT has been shown to boost translators' productivity and meet clients' case-by-case requirements while maintaining acceptable quality (Biela-Wolonciej 2007; Dragsted & Hansen 2009; Dragsted, Mees & Hansen 2011; also see Ciobanu 2016 but with caveats and mixed feelings of some practitioners). In addition, nowadays more conference speakers tend to hand scripts or slides to the interpreters, thereby prompting them to resort to SiT from time to time. In brief, SiT has every reason to be explored from different angles.

Earlier endeavours proved that SiT is just as demanding as other modes of interpreting and unique in its own way, particularly for language pairs that are syntactically more dissimilar, such as English and Chinese. The differences in the two language systems, even in the lexicon (Huang 2009; Frost 2012), lead to divergent characteristics of language use. For example, word order plays a more crucial role in Chinese (Jiang 2009), which "must express various syntactic and semantic relations

[1] Queen's University Belfast (ORCID: 0000-0001-7321-2029)
[2] National Taiwan Normal University
[3] National Chengchi University (ORCID: 0000-0001-5413-4889)

through the use of function words and functional manipulation of word order" (Huang 2009, 3); on the other hand, English allows for a more flexible organisation of meaning units. At the level of phrases, clauses, or sentences, the concept of *principal branching direction* (PBD) becomes a thorny issue—describing the relative position of a modifier, e.g., a relative clause or an adverbial subordinate clause, to its head, or modified unit, e.g., the antecedent or main clause (Diessel 2004; Chen 2006). PBD poses a formidable challenge for interpreters because, as sentence complexity increases, the position of the modifier is more likely to differ in the two languages. Modifiers in Chinese are mostly left-branching (Huang 2009; Yang, Perfetti & Liu 2010), but English gravitates towards the right-branching principle on phrase and sentence levels (Chen 2006).

The above-mentioned language specificity has been reflected in several studies. Viezzi (1989) registered a lower retention rate in SiT than in simultaneous interpreting (SI), especially when the syntactic rules of the involved language pair varied substantially. Her (1997) and Chang (2008) repeatedly found that the differences in syntactic structures between languages were dragging their test participants down. Aiming to understand the difficulty of SiT, Agrifoglio's (2004) informants had a much more serious problem with language expression than with accuracy. Chiang, Kuo & Chen (2009) focused on inadequate pauses, again attesting to the woes faced by sight translators. Output analysis in previous research has evidenced the troubles of SiT, in which language use of the source text (ST) tends to be more convoluted than oral speech (Stubbs 1983; Chafe & Danielewicz 1987; Biber 1991).

Nonetheless, some additional questions seem in order. What does the interpreter gaze at in order to collate information in SiT? How can she ensure high-quality interpretation with poised looks and a steady voice? What leads to a faltering, staccato rendition, and how? These questions can only be answered when the above findings are complemented with data on the SiT process, namely, how and when information is retrieved, reformulated, and delivered. New findings from a cognitive perspective may shed light on the nature of SiT and how training can be designed to facilitate skill development.

The first step towards this goal is an examination of reading because, in SiT, input comes via the visual channel. Reading in one language is complex enough, since it is an artificial activity designed by humans for documenting and communicating thoughts across time and space, and involves elaboration of phonological and morphological information, meaning construction, and constant verification and integration of meaning by simultaneously contemplating the relationships between meaningful units conveyed through syntax (Rayner et al 2012; Willingham 2017). Bilingual reading further complicates the process as language command, the environment, the distance between the reader's language repertoire—even when the reader starts learning each language—all have a role to play.

One of the most relevant factors may be syntactic differences between languages that exert apparent influence on bilingual reading (Fernández 2002; Frenck-Mestre 2002) and on SiT (Viezzi 1989; He 1997; Chang 2008). The latter requires the interpreter to consciously and concurrently strengthen the activation of one linguistic system and inhibit the other and then reverse activation and inhibition at different stages of processing (discussion in Grosjean 2001). Therefore, cognitive studies of SiT should include reading and its interaction with reformulation and speech production. However, for now the two areas seem to have been flourishing independently and without sufficient interchange, except for a few papers (e.g., Lykke Jakobsen & Jensen 2008; McDonald & Carpenter 1981; Shreve, Lacruz & Angelone 2010; Shreve, Schäffner, Danks & Griffin 1993; Chmiel & Mazur 2013).

This study aims to outline interpreters' cognitive processes by tracking how the text to be translated is read and how training influences reading behaviour in SiT, with output quality serving as a reference point for comparison.

2. The present study

SiT involves ST reading and oral production. The first part of this task highly resembles *silent reading* (SR) for information retrieval, while the second part is similar to *reading aloud* (RA) to some extent, except that RA does not require reformulation in a different language. Therefore, SR and RA have also been included for comparisons.[4] The data used in this study comes from a larger database, and details will be explained below.

2.1. *Participants*

The data of two groups of 18 participants was included. One group consisted of interpreting students having received postgraduate SiT training—lasting between one semester (dedicated course) and one year (intermittent guided practice)—within three years by the time of the experiment. The other group was late bilinguals with comparable language proficiency who had not received any interpreting training. All participants held a language certificate equivalent to or above TOEFL 90 (or IELTS 6.5), and everyone had normal or corrected-to-normal vision and signed an informed consent form in advance.

A questionnaire was used to record participants' basic information, language proficiency and learning experience, educational background, work experience in T&I, and language use in their daily lives, but results will only be reported selectively below. In addition, an English reading span test, designed on the basis of Daneman & Carpenter (1980), was administered to determine each participant's working memory (WM) because we only used English materials and previous studies have found that interpreting performance could be moderated by one's WM capacity (Christoffels, de Groot & Waldorp 2003; Christoffels, de Groot & Kroll 2006). A total of 42 sentences were used in the reading span test. The last word of each sentence ranged between 5 and 9 letters, with an average word frequency of 83.03. A final score on a scale of 1-5 would represent the size of one's reading span, or WM.

INSERT TABLE 1 ABOUT HERE

Table 1. Selective background information of participants

	NOV	**UNT**
Age	25.94 *(4.93)*	25.23 *(5.95)*
Language proficiency (IELTS)	7.58 *(0.55)*	7.27 *(0.46)*
Reading span	2.67 *(1.08)*	2.11 *(0.7)*

*NOV, interpreting students (novices); UNT, untrained bilinguals
**Figures in the parentheses indicate group standard deviation

Table 1 presents participants' average age, language proficiency, and reading span size. These indicators have been subjected to an independent *t*-test respectively. The average age of interpreting students (M = 25.94) was slightly higher than that of untrained bilinguals (M = 25.23) and did not show any significant difference, p = 0.71, $t(31)$ = 0.38. The same held true for language proficiency, for which interpreting students scored better (M = 7.58) than untrained bilinguals (M = 7.27), p = 0.09, $t(31)$ = 1.78. WM capacity, as manifest through reading span, again copied the same trend, with interpreting students scoring 2.67 and untrained bilinguals 2.11 on average, p = 0.08, $t(34)$ = 1.83.

2.2. *Materials*

Three different speeches on the same topic—the current status of the WTO and the development of global trade—from the Director-General of the World Trade Organization were reduced to three 175-word texts. Rewriting was kept to a minimum to maintain the authenticity and style of formal speeches in a diplomatic setting, but domain-specific terms or particulars were avoided, so as to prevent domain knowledge from interfering with the results. No specific *Areas of Interest* (AOIs) were clearly chosen for the eye tracker, as the focus of this study was to provide a broad view of the reading and SiT processes.

[4] On the other hand, comparing SR and RA should allow for us to single out special effects derived from oral production, which, however, is not the focus of this study.

Table 2 provides a breakdown of text descriptors. Microsoft Word was used to rate sentence difficulty and readability of each text. The percentage of passive sentences was almost the same; the other two indices also indicated similar readability. Linguistic units that might pose challenges for reformulation were equally distributed among the texts. Four additional untrained bilinguals with comparable language command to that of test informants were recruited to evaluate the three texts and confirmed that they were easily comprehensible without background knowledge. Limited by the size of the computer monitor, each text was divided into four paragraphs (hereafter called *trials*) and projected on the screen consecutively. Each trial received a score from 1–7 (number increasing with difficulty). A score representing each text was obtained by averaging four sub-scores from each evaluator. The mean score is 3 for Text A, 2.75 for Text B, and 3.25 for Text C, with Cronbach's α reaching .977, indicating high reliability and a similar difficulty level between texts.

INSERT TABLE 2 ABOUT HERE

Table 2. Text features of formal trials

	Text A	**Text B**	**Text C**
Word count	175	175	175
Adjectival words & phrases	11	10	10
Adjectival clauses	2	3	2
Adverbial words & phrases	8	8	9
Adverbial clauses	2	2	2
Passive sentences (in %)	11	12	10
Flesch Reading Ease	44.0	53.1	39.6
Flesch-Kincaid Grade level	11.9	11.2	12.0
Difficulty rating by raters	3	2.75	3.25

2.3. *Design*

The current study adopts a 2 × 3 design: group (interpreting students and untrained bilinguals) and condition/task (SR, RA, and SiT). Each participant had to engage with all the texts and tasks, with the presentation sequence and combination of texts and tasks counter-balanced respectively across participants. Through the counter-balancing of the stimuli, the results may avoid the influence of the characteristics of each text and practice effects. Two comprehension questions also came at the end of each text to ensure the results reflect sufficient understanding.

All stimuli were presented on a desktop computer screen (1024 × 768 pixels). Movements of the dominant eye were recorded with an Eyelink 1000, with the sampling rate set at 1000 Hz. The stimuli used Courier (22 pixels) and were shown in the middle of the screen, one trial at a time, on a grey background. The distance between the monitor and the participant was approximately 70 cm.

2.4. *Procedure*

First, the experiment was introduced, followed by instructions on how to proceed. The only two technical words, along with their translation, were provided and each participant was given sufficient time after the introduction to familiarise themselves with the terms before the experiment started. Successful 9-point calibration would bring the participant to the next phase, and everyone would take turns engaging with all three tasks. Before each task formally began, one paragraph (2–3 sentences) appeared for the participant to practice the assigned task, and two comprehension questions followed at the end of the task. After answering the questions, the participant went through calibration again before taking on the next task. Every time a page was turned, participants had to fixate on a cross located in the upper-left corner of a blank screen, where the new paragraph would begin, for the text to appear. Participants were advised to proceed at their own pace and adhere to the ways they would normally use to tackle each task, and their interpretation was for a group of audience genuinely relying on the service. They were also reminded that the sentences in each trial were complete, and it would not be possible to return to the previous trial once they move on to the next one.

2.5. *Data analysis*

Previous studies have mainly focused on production, which tell us about the end result but do not provide insights into how interpreters actually address the task along the way or identify how we may actually facilitate the development of interpreting skills. On the other hand, an increasing number of studies focus on uncovering the underlying mechanism of information processing (Wang & He 2018; Su & Li 2019). However, conclusions could be difficult to reach, since we are not in a position to judge how efficient or effective their strategies were without assessing the outcome. A study examining both comprehension and production provides a window into distinct cognitive processes or strategies. Participants may finish a task within similar task time with utterly disparate results or they may achieve a similar level of quality with the final output produced in different reading passes.

To provide a bird's eye view of the cognitive phenomena surrounding SiT, this study turned to the final product and the indices that could reflect the mental processes of this mode of interpreting. Answering five out of six comprehension questions correctly was required for the data from any participant to be included. Local, word-based indices that reveal reading behaviour were categorised into first-pass and non-first-pass indicators and reported in this study. For first-pass indicators—*first fixation duration* (FFD) and *gaze duration* (GD)—only fixations between 80–800 ms were selected for analysis to ensure that the participant was focused and competent enough to at least comprehend individual words in a reasonable time. On the other hand, all data points of the non-first-pass indicators—*go-past time* (GPT), *re-reading time* (RRT), and *total viewing time* (TVT)—were examined for a clearer picture of later-stage processing, such as information synthesis, reformulation, or even production. As SiT invokes the knowledge of two language systems and demands conscious inhibition of one and activation of the other at different stages, non-first-pass reading measures might reflect higher cognitive demands on the part of the interpreter.

In addition to the above reading indices, the participant's recorded interpretation was sent to two senior AIIC members for quality assessment to support our analysis with additional evidence. The two raters used the assessment criteria adopted by the English-Chinese Interpreting Exam hosted by the Language Training and Testing Center in Taiwan.[5] SiT performance was marked on two components: accuracy and style—the latter entailing fluency, language expression, pace control, and other paralinguistic features. The two components were marked independently on a scale of 0–5 for each trial under the SiT task (see Appendix 1).

Six additional Chinese native speakers were invited to act as interpreting service users and informed of the context of the interpretation. They were then requested to listen to each SiT recording and mark on the transcript when they sensed a pause in the output, so we can understand the pausing behaviour from the perspective of real audience.[6]

3. Results

The results presented below are divided into three major categories: (1) global data, i.e., total time spent on each task and quality of the SiT performance; (2) local data, covering each group's reading performance, informed by five eyetracking indices; and (3) the reading and output behaviour in the SiT process. An independent two-tailed *t*-test was used as the statistical method for each index

[5] This is the only standardised interpreting exam used in Taiwan, with clear guidelines and assessment criteria agreed by T&I scholars.

[6] We want to use listeners' perception as the definition of a pause—i.e. how long for them means a pause—instead of imposing a pre-determined threshold because data analysis and interpretation could be affected.

below until specified otherwise, with a 95% confidence interval, and alpha was set at 0.05.

3.1. *Global indices: task- or trial-based performance*

Global data gives us an overview of each group's performance in different tasks. We mentioned only participants with comparable linguistic competence were recruited for this study, and a brief look at the task of SR and RA in Table 3 along with their language certificates suffices to support this claim. For SR, no significant difference was found between interpreting students and untrained bilinguals on all three indices listed in Table 3, including trial-based total time spent on the task (M = 21.47s for NOV and M = 25s for UNT, $t(34) = 1.47$, $p = 0.15$), trial-based number of fixations (M = 75.86 for NOV and M = 88.97 for UNT, $t(34) = 1.56$, $p = 0.13$), and mean fixation duration in each trial (M = 235.25ms for NOV and M = 231.97ms for UNT, $t(34) = 0.41$, $p = 0.68$). The above findings indicate that both groups were able to complete the task with similar cognitive resources and amount of time used and at sufficient level of comprehension.

The results for RA send a generally resembling message. No significant difference was found between the two groups on either the total time spent (M = 23.93s for NOV and M = 26.92s for UNT, $t(29) = 1.73$, $p = 0.1$), or the mean fixation duration on words (M = 270.8ms for NOV and M = 245.5ms for UNT, $t(29) = 1.81$, $p = 0.08$). However, this time interpreting students did rely on significantly fewer fixations (M = 76.5) than untrained bilinguals (M = 93.13), $t(28) = 2.43$, $p = 0.02$. The two groups might have adopted different strategies, but they still reached the finish line at roughly the same time. Interpreting students appear to have read fewer times in total but generally spent more time per fixation for comprehension, and the strategy or inclination did not lead to a significantly less total time for the task.

ST results display a completely different picture. The average fixation duration was the only index for which a statistically significant difference was absent (M = 264.2ms for NOV and M = 255.06ms for UNT, $t(25) = 0.64$, $p = 0.53$). Interpreting students (M = 42.11s) spent only a little more than half of the time than untrained bilinguals (M = 76.07s) on SiT, $t(21) = 4.17$, $p < 0.001$, and relied on significantly fewer fixations to complete the task (M = 131.76 for NOV and M = 239.65 for UNT, $t(23) = 3.91$, $p < 0.001$).

INSERT TABLE 3 ABOUT HERE

Table 3. Trial-based global measures for three tasks

		SR		RA		SiT	
		NOV	**UNT**	**NOV**	**UNT**	**NOV**	**UNT**
Total time (s)		21.47	25.00	23.93	26.92	42.11	76.07
	s.d.	*7.22*	*7.22*	*3.95*	*6.20*	*11.10*	*32.73*
Number of fixations		75.86	88.97	76.50	93.13	131.76	239.65
	s.d.	*24.91*	*25.59*	*14.90*	*24.96*	*44.39*	*108.37*
Mean fixation duration (ms)		235.25	231.97	270.80	245.50	264.20	255.06
	s.d.	*27.40*	*19.53*	*49.90*	*31.81*	*54.21*	*27.04*

SR, silent reading; RA, reading aloud; SiT, sight translating; NOV, interpreting students (novices); UNT, untrained bilinguals

Table 3 shows only the reading behaviour. When examined together with Table 4, which focuses on the output of SiT, we may have a more holistic understanding of participants' performance as a group. The number of Chinese characters[7] used by interpreting students (M = 350.39) and untrained bilinguals (M = 369.94) respectively was the only index out of four in Table 4 that did not generate a significant difference after inferential statistical analysis was applied, $t(21) = 0.91$, $p = 0.38$. This shows that training at this stage does not seem to have enough influence to elicit consistent brevity across the group of interpreting students.

On the other hand, trainees' overall SiT quality (M = 3.92) was significantly better than that of untrained bilinguals (M = 2.87), $t(25) = 5.42$, $p < 0.0001$, as was accuracy (M = 4.03 for NOV and M = 3.15 for UNT, $t(23) = 4.05$, $p < 0.001$) and style (M = 3.82 for NOV and M = 2.58 for UNT, $t(34) = 5.49$, $p < 0.0001$).

INSERT TABLE 4 ABOUT HERE

Table 4. ST performance and ratings

	NOV	*s.d.*	UNT	*s.d.*
Overall performance	3.92	*0.38*	2.87	*0.73*
Accuracy	4.03	*0.35*	3.15	*0.84*
Style	3.82	*0.63*	2.58	*0.72*
Character count (Chinese)	350.39	*29.41*	369.94	*86.58*

NOV, interpreting students (novices); UNT, untrained bilinguals

With a short period of SiT training between one semester and a year, trainees already exhibited a potent enough change in the outcome to differentiate themselves from average untrained bilinguals in terms of accuracy and style.

3.2. *Local reading indices: word-based eyetracking data*

Superior SiT performance delivered by interpreting students made us wonder about the source from which this difference came. Could a refined classification of fixation duration give us more insight into when and how training exerts its impact? After independent *t*-tests were applied to each stage of cognitive processing on the reception side in all three different tasks, an unexpected finding emerged: interpreting students and untrained bilinguals were strikingly similar in a statistical sense along the way of each task. In other words, both groups had indistinguishable efficiency of meaning construction, as far as all five indices listed in Table 5 are concerned.

INSERT TABLE 5 ABOUT HERE

Table 5. Mean duration of local, word-based reading indices in ms

		Silent reading		Reading aloud		Sight translation	
		NOV	UNT	NOV	UNT	NOV	UNT
First fixation duration (FFD)		262.11	261.15	307.54	284.54	254.35	258.32
	s.d.	*32.76*	*20.42*	*45.75*	*31.12*	*34.76*	*42.29*
Gaze duration (GD)		321.26	316.01	370.02	360.29	287.93	298.76
	s.d.	*38.61*	*18.70*	*48.44*	*37.29*	*45.66*	*53.16*
Go-past time (GPT)		402.65	413.66	470.02	475.20	410.82	431.38

[7] A Chinese character represents a morpheme instead of a word, which sometimes requires more than one Chinese character to constitute a meaning unit (Rayner et al 2012).

s.d.	*80.20*	*66.31*	*82.39*	*87.37*	*155.82*	*178.51*
Re-reading time (RRT)	367.67	343.75	375.72	368.26	455.43	525.53
s.d.	*86.57*	*54.30*	*89.54*	*64.32*	*173.77*	*163.28*
Total viewing time (TVT)	438.67	426.73	461.84	473.81	508.05	603.21
s.d.	*113.33*	*71.78*	*94.38*	*79.42*	*202.91*	*221.94*

NOV, interpreting students (novices); UNT, untrained bilinguals

As mentioned, FFD and GD mainly stand for the cognitive effort needed for stimuli identification and initial information retrieval, while GPT, RRT, and TVT more adequately explain the process of validating, integrating, constructing and interpreting—in a hermeneutic sense—meaning. In this regard, SR and RA, the two tasks we used in this study to substantiate our claim of homogeneity of participants' language abilities, again showed no significant difference between groups on any of the indices in Table 5.

Starting with SR, FFD (*M* = 262.11ms for NOV and *M* = 261.15ms for UNT, $t(28) = 0.11$, $p = 0.92$) and GD (*M* = 321.26ms for NOV and *M* = 316.01ms for UNT, $t(25) = 0.52$, $p = 0.61$) both failed to differ. Later stages of reading process, including GPT (*M* = 402.65ms for NOV and *M* = 413.66ms for UNT, $t(34) = 0.45$, $p = 0.66$), RRT (*M* = 367.67ms for NOV and *M* = 343.75ms for UNT, $t(29) = 0.99$, $p = 0.33$), and TVT (*M* = 438.67ms for NOV and *M* = 426.73ms for UNT, $t(29) = 0.38$, $p = 0.71$), also consistently pointed out that meaning access was equally easy or laborious for both groups.

RA conveyed the same message: we could not find any significant difference between the two groups, be it FFD (*M* = 307.54ms for NOV and *M* = 284.54ms for UNT, $t(34) = 1.76$, $p = 0.09$), GD (*M* = 370.02ms for NOV and *M* = 360.29ms for UNT, $t(34) = 0.68$, $p = 0.5$), GPT (*M* = 470.02ms for NOV and *M* = 475.20ms for UNT, $t(34) = 0.18$, $p = 0.86$), RRT (*M* = 375.72ms for NOV and *M* = 368.26ms for UNT, $t(34) = 0.29$, $p = 0.78$), or even TVT (*M* = 461.84ms for NOV and *M* = 473.81ms for UNT, $t(34) = 0.41$, $p = 0.68$).

The above results had been readily expected because participants' language command was controlled for. We only assumed that the impact of training would manifest itself in later stages of reading during SiT, since these stages are, intuitively speaking, the time when reformulation might reasonably take place. However, the results in Table 5 failed to show any significant difference in this regard between groups either. While FFD (*M* = 254.35ms for NOV and *M* = 258.32ms for UNT, $t(34) = 0.31$, $p = 0.76$) and GD (*M* = 287.93ms for NOV and *M* = 298.76ms for UNT, $t(34) = 0.66$, $p = 0.52$) were statistically similar between interpreting students and untrained bilinguals, the same trend continued all the way through for GPT (*M* = 410.82ms for NOV and *M* = 431.38ms for UNT, $t(34) = 0.37$, $p = 0.72$), RRT (*M* = 455.43ms for NOV and *M* = 525.53ms for UNT, $t(34) = 1.25$, $p = 0.22$), and TVT (*M* = 508.05ms for NOV and *M* = 603.21ms for UNT, $t(34) = 1.34$, $p = 0.19$).

3.3. *Sight translation in action: behaviour in the process*

The process of SiT is the next place we turned to, after failing to see statistically significant impacts of training on word-based local indices. As mentioned, participants were asked not to rush to begin but to proceed when they felt more comfortable and to pretend they were presented with a speech script and requested to perform onsite oral translation for a group of audience that would genuinely require interpreting service. By giving these instructions, we tried to observe each person's behaviour under more natural circumstances, without forcing them to start sight translating the text as soon as possible as we would ask students to do in training. It turns out that trainees' behaviour patterns are highly distinct from their counterparts—and much more effective, considering the scores of their accuracy and style.

3.3.1. *Reading ahead: how much pre-reading does one need?*

Reading is the only input channel in SiT, while *reading ahead* is required to "ensure a smooth delivery of the translation without causing unnecessary interruption in the reception of the message for the audience" (Chen 2015, 144). With the need to safeguard the accuracy of the interpretation and avoid staccato performance at the same time, it's reasonable to assume that the interpreter would need to read through or glance over the text before SiT even begins. This was indeed the case for untrained bilinguals, who on average took significantly longer time (*M* = 17.27s; *SD* = 12.83s) than interpreting students (*M* = 5.99s; *SD* = 7.8s) before they felt ready to start producing output, $t(28) = 3.19$, p = 0.004; hence a larger eye-mouth span (EMS)—the lapse of time between the moment the eyes start gathering information and the onset of oral rendition—ahead of the initial utterance, which is equal to one Chinese character throughout this study or occasionally two if they are continuous, inseparable sounds.

INSERT FIGURE 1 ABOUT HERE

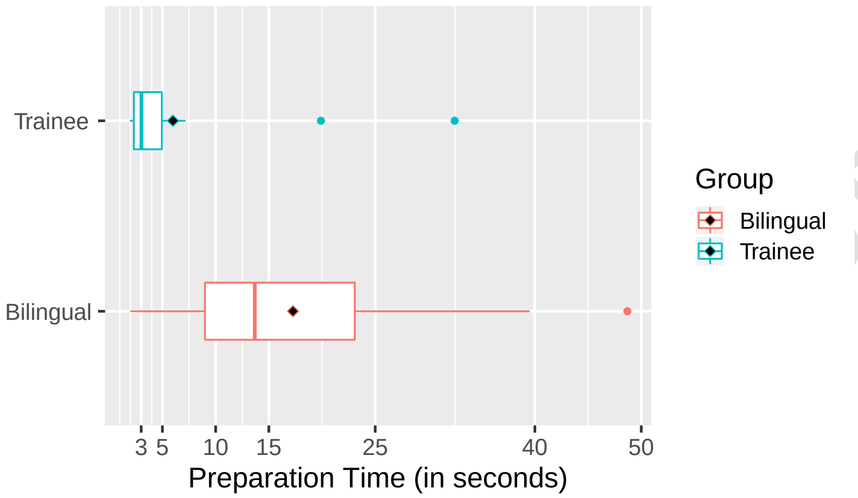

Figure 1. Period of silence before SiT begins

A look at the distribution of each group in Figure 1 offers a clearer picture. The box plot indicates that half of the interpreting students had an EMS for about three seconds (*Mdn* = 3.03s)—and 75% within five seconds—while half of the untrained bilinguals spent almost five times longer perusing the text (*Mdn* = 13.68s), with 75% slightly under 25 seconds. However, how much information did the participants from each group gather during the *initial* EMS—i.e., the EMS before oral rendition begins in each trial? Calculating the number of fixations incurred in the preparation stage could potentially offer more insights, as it is a more accurate measure of cognitive activities. Untrained bilinguals on average had 59.56 fixations (*SD* = 41.95) before oral output began; on the other hand, the mean value for interpreting students was significantly lower at 22.57 fixations (*SD* = 28.98), $t(34) = 3.08$, $p = 0.004$.

INSERT FIGURE 2 ABOUT HERE

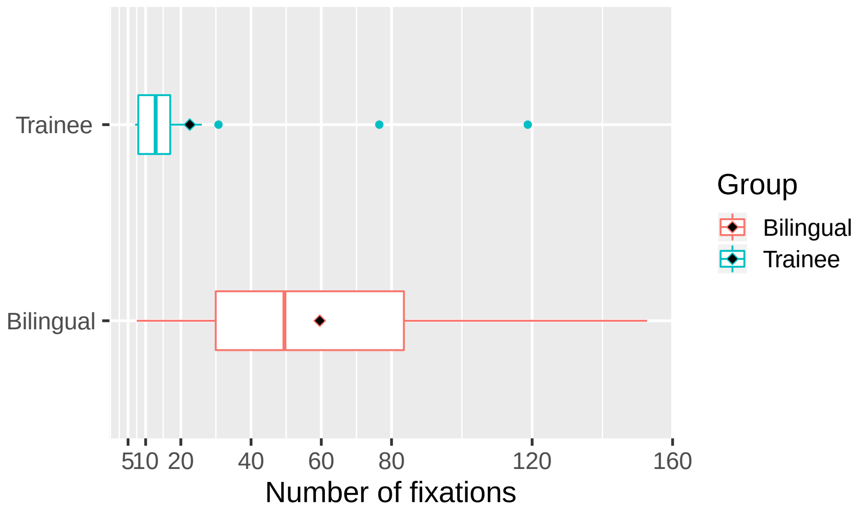

Figure 2. Information retrieval during the initial EMS

Again, the distribution of each group in Figure 2 tells a more compelling story (*Mdn* = 12.88 for trainees; *Mdn* = 49.5 for untrained bilinguals). The distribution of the trainees' group is squeezed tightly towards the left-hand side of the figure, while that of their counterparts is stretched across, showing rather diverse behaviours among untrained bilinguals. Theoretically speaking, each fixation would most likely land on only one single word. Nonetheless, parafoveal vision could potentially cover some part of the previous/next word and provide some extra hints (Radach & Kennedy 2004; Tsai et al 2004; Clifton et al 2016). Thus, it would be safer at this stage to use the number of fixations as the unit of measurement. The aforementioned disparity between the two groups continued throughout after SiT began. Figure 3 shows the mean number of fixations ahead of each non-initial utterance, or in each *subsequent* EMS, for both groups. Untrained bilinguals on average required significantly more fixations (M = 2.74; SD = 1.1; *Mdn* = 2.54) than interpreting students (M = 1.81; SD = 0.31; *Mdn* = 1.83) to maintain the flow of interpretation, $t(20) = 3.46$, $p = 0.002$.

INSERT FIGURE 3 ABOUT HERE

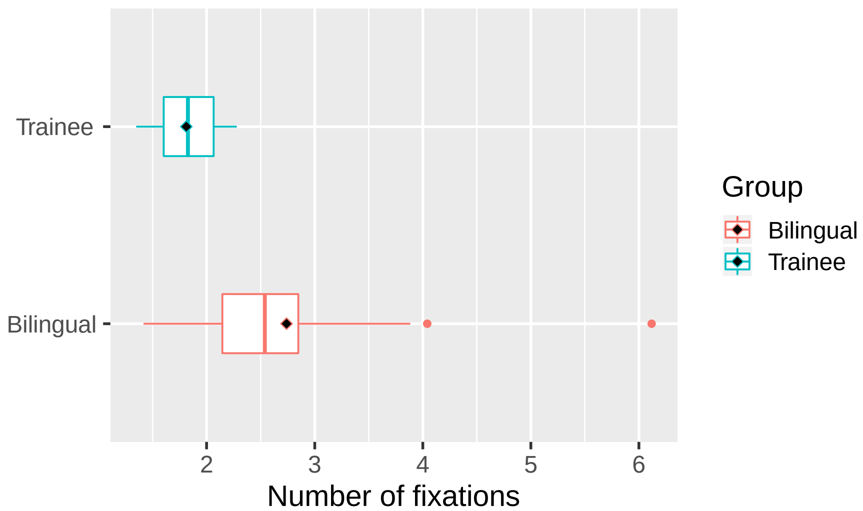

Figure 3. Number of fixations during each subsequent EMS

At this point, it is quite obvious that interpreting students, after receiving training, were able to make a head start much sooner than untrained bilinguals and managed to rely on significantly less information to maintain the flow of their output all the way until they finished the task of SiT.

3.3.2. *Observable "on-line" measures: more signposts for the effects of training*
In addition to eye movements, some other more readily observable behaviours are also telltale signs of how training has exerted its influence on the interpreting students. Verbal gap is the first and foremost index, which denotes the gap in time between every two utterances. The mean duration of a verbal gap was 345.64ms (SD = 103.23ms) for untrained bilinguals and 258.95ms (SD = 33.9ms) for trainees, with the difference reaching statistical significance, $t(21) = 3.39$, p = 0.003. The pace of interpretation provided by the interpreting students was clearly faster than that of untrained bilinguals.

Longer-than-necessary verbal gaps would give the listener the impression that the speaker (or the interpreter in this case) is "dragging on" and therefore sounds disfluent or unnatural. This can be further corroborated by a rule of thumb in the interpreting profession that SiT output should sound like reading the source text aloud instead of performing oral translation on the spot (Angelelli 1999; Chen 2015). Research on spontaneous speech found that *verbal gaps*—or *pauses,* as they were called in these studies—longer than 200ms could be detected (Boomer & Dittmann 1962); in this study, however, only gaps longer than 600ms were consistently identified by six native speakers of the target language, so "pauses cannot be abstracted from their speech context and studied in isolation" (219). As SiT and spontaneous speech are two entirely dissimilar tasks, for which the listener could also have different expectations, a pause has hence been defined accordingly as verbal gaps lasting more than 600ms and classified as a separate index.

Among all the verbal gaps in the interpreting students' output, only an average of 13.9% (SD = 3.67%) were marked as pauses, regardless of their types, positions in or between sentences, or functions; on the other hand, untrained bilinguals had a significantly higher percentage of pauses (M = 19.44%; SD = 5.17%), $t(34) = 3.71$, $p = 0.001$. This means the interpreting students were not only producing oral output at a faster pace, but they also were much more fluent, so that the audience sensed "breaks" in their performance less frequently. Moreover, the average length of a pause produced by interpreting students (M = 1.31s, SD = 0.28s) was significantly shorter than that of untrained bilinguals (M = 2.06s, SD = 1s), $t(20) = 3.07$, $p = 0.006$; hence a smoother rendition—which is duly reflected in their score on style.

The last index we intend to cover here is the amount of information each group of participants tried to collect during a pause—be it a hesitant one that requires problem-solving clues or a proper one that gives the interpreter some break time and a chance to read ahead to plan the next step. Interpreting students, in this regard, consistently relied on significantly fewer fixations (M = 3.99, SD = 1.11) than untrained bilinguals (M = 6.36, SD = 3.23) during pauses, $t(21) = 2.95$, $p = 0.008$, denoting a much lower kick-start threshold every time when they had to stop and plan their next move to ensure successful SiT delivery (see Table 6).

INSERT TABLE 6 ABOUT HERE

Table 6. Number of fixations in pauses and an overview of the distribution

	NOV	**UNT**
Median	1	1
3rd Quartile	1	1
Max.	16	215
Mean *(s.d.)*	3.99 *(1.11)*	6.36 *(3.23)*

NOV, interpreting students (novices); UNT, untrained bilinguals

The raw data of the number of fixations in pauses by group in Table 6 perhaps tells a more intriguing story. To calculate the mean number of fixations for a group, we first obtained a mean value for each participant and then averaged these numbers to reach a group mean; however, here we decided to treat each pause as an independent point of observation, regardless of whether the pauses came from the same person, and we found that a mere 25% of the pauses were the deciding factor that told the two groups apart. Every participant—be it a trainee or an untrained bilingual—had only one fixation in 75% of the pauses. When stopped, trainees moved on before long (16 fixations maximum), while untrained bilinguals seem to have buried themselves in their own world (215 fixations maximum).

4. Discussion

With age, language proficiency, and WM capacity maintained at a comparable level, we still see obvious effects of training in the process and the product of SiT. Interpreting students managed to provide significantly more accurate, adequate, and fluent interpretation at higher efficiency in that they started sight-translating the text much sooner and finished the task in almost half of the time required by their counterparts. However, we acknowledge that the term *efficiency* used here is vague and less than ideal because the interpreting students' efficiency of information retrieval during each stage of reading was not different from that of the untrained bilinguals, nor was the number of Chinese characters used in their output. We initially assumed that training would speed up processing in later stages of reading when meaning integration and reformulation are involved; yet, statistical analysis failed to show any significant difference in this study.[8] Considering the significantly fewer fixations needed to complete SiT by interpreting students, we can see that the efficiency actually came from more economical ways of reading rather than quicker processing each time when the trainees' eyes fixated on the text.

When further examining other aspects of the SiT process, fundamental changes induced by training became more readily observable. The reading-ahead behaviour, including the initial and subsequent EMS, was perhaps the most distinctive feature that made interpreting students stand out. The trainees, apparently aware that longer silence might leave a negative impression on the audience, almost consistently chose to start sight translating the text in about six seconds; in stark contrast, untrained bilinguals had a mean initial EMS of 17.27 seconds, which was almost three times longer. While some might expect that the benefits coming with a rushed start could be offset by slower progress afterwards, the interpreting students somehow managed to finish SiT at a faster average pace throughout. What was more, a target-language audience sensed significantly fewer and shorter pauses amid this group's interpretation.

In order to pull of the aforementioned feats, interpreting students would need to have gleaned sufficient information that guarantees appropriate reformulation choices—and even have already decided on the words to use—before opening their mouth every time. The results would not come as a surprise if the text had been perused thoroughly during the initial EMS. Untrained bilinguals generally read through the text before SiT began (59.56 fixations for an average of 43.75 words in each trial). However, the distinctive behaviour of the interpreting students presented them with an additional challenge not faced by their counterparts.

This means we can safely assume that, when entering the production stage, untrained bilinguals were tackling old information; on the contrary, interpreting students were constantly dealing with new stimuli they saw the first time because they made a quick start for oral rendition. A rapid head start plus a faster average pace of SiT show that these students were effective in reading ahead, collating information required for reformulation, quickly (if not almost immediately) engaging in oral production, and consistently relying on less perceptual input throughout to finish SiT—and ending with significantly better quality.

However, English and Chinese are not similar in their syntactic structures—Chinese is consistently left-branching and English is mostly right-branching, especially for more complex and difficult sentence structures, which are not rare in the genre of the text adopted in this study. Without peeking far ahead at new information in each EMS, the interpreting students would have had to rely on other strategies. At this point, we would make a bold assumption that chunking, or segmentation skills normally taught in the SiT classroom would be the major contributor to the interpreting students' success, as can be indirectly corroborated by the significantly fewer fixations in both initial and subsequent EMS and by the fact that even in the longest pause the number of

[8] These untrained bilinguals were mainly seniors in the university and only one out of 18 have stayed abroad for 5-7 years and the other two less than three years, which makes them average late bilinguals and possibly rules out the impacts of language immersion.

fixations was only 16, compared to 215 for untrained bilinguals—which all seem to convey a rhythmic chunking behaviour, even only vaguely for now.

5. Conclusion

This study set out to investigate the effects of training on the process and product of sight translation. Product-wise, participants with training performed significantly better than untrained bilinguals, while language proficiency, WM capacity, and age did not differ significantly between groups and were therefore excluded from the discussion.

Since local reading indices failed to show any difference between the two groups, we turned to other behavioural indicators. Training fundamentally changed how the participants tackled information. With training, the participants were able to start sight translating the text before delaying for too long and managed to maintain a fairly low kick-start threshold most of the time—thus breaking long strings of information into shorter segments—to keep their rendition flowing. We would go one step farther to claim that the application of chunking skills, which necessarily entails more flexible language use, was a decisive factor in the successful completion of SiT in this study.

However, we only have circumstantial evidence at this stage, and examining the exact locations of the interpreter's eyes during each EMS will be an inevitable step to substantiate our claim and provide a more precise picture of the reading behaviour and strategies used in SiT. Another drawback is that this study favours authenticity of the materials over tight experiment control. More basic features of reading during SiT might need to be identified with stricter manipulation of the materials and more detailed comparisons between different kinds of tasks. In addition, some personal characteristics might still have a role to play. Therefore, it might be worthwhile to break down the boundary between groups to observe the moderating effects of background factors.

Acknowledgements

We would like to thank the two anonymous reviewers for their invaluable feedback and suggestions. Any inadequacy of this paper is inevitably the responsibility of the authors.

References

Agrifoglio, Marjorie. 2004. "Sight Translation and Interpreting: A comparative analysis of constraints and failures." *Interpreting* 6 (1): 43–67.

Angelelli, Claudia V. 1999. "The Role of Reading in Sight Translation: Implications for teaching." *The ATA Chronicle* 28 (5): 27–30.

Biber, Douglas. 1991. *Variation across Speech and Writing*. New York: Cambridge University Press.

Biela-Wolonciej, Aleksandra. 2007. "A-Vista: New challenges for tailor-made translation types on the example of the recorded sight translation." *Kalbotyra* 57: 30–39.

Boomer, Donald S., and Allen T. Dittmann. 1962. "Hesitation Pauses and Juncture Pauses in Speech." *Language and Speech* 5 (4): 215–220.

Chafe, Wallace, and Jane Danielewicz. 1987. "Properties of Spoken and Written Language." https://www.nwp.org/cs/public/download/nwp_file/142/TR05.pdf?x-r=pcfile_d. Accessed 8 September 2016.

Chang, Huei-Fang. 2008. "A Study on GITI Students' ST Difficulty & Performance." *SPECTRUM: NCUE Studies in Language, Literature, Translation, and Interpretation* (2): 61–83.

Chen, Jyun-Gwang. 2006. *Contrative Research & Crosslinguistic Influence*. Taipei: Crane.

Chen, Wallace. 2015. "Sight translation." In *The Routledge Handbook of Interpreting*. Edited by H. Mikkelson and R. Jourdenais, 144–153. Abingdon: Routledge.

Chiang, Heien-Kun, Feng-Lan Kuo, and Wei-Chun Chen. 2009. "Exploring Pausing Patterns in Two-way Sight Translation." *SPECTRUM: NCUE Studies in Language, Literature, Translation, and Interpretation* (5): 155–167.

Chmiel, Agnieszka, and Iwona Mazur. 2013. "Eye Tracking Sight Translation Performed by Trainee Interpreters." In *Tracks and Treks in Translation Studies*. Edited by C. Way, S. Vandepitte, R. Meylaerts and M. Bartłomiejczyk, 189–205. Amsterdam: John Benjamins.

Christoffels, Ingrid K., Annette M. B. de Groot, and Judith F. Kroll. 2006. "Memory and Language Skills in Simultaneous Interpreters: The role of expertise and language proficiency." *Journal of Memory and Language* 54 (3): 324–345.

Christoffels, Ingrid K., Annette M. B. de Groot, and Lourens J. Waldorp. 2003. "Basic Skills in a Complex Task: A graphical model relating memory and lexical retrieval to simultaneous interpreting." *Bilingualism: Language and Cognition* 6 (3): 201–211.

Ciobanu, Dragoș. 2016. "Automatic Speech Recognition in the Professional Translation Process." *Translation Spaces* 5: 124-144.

Clifton, Charles, Fernanda Ferreira, John M. Henderson, Albrecht W. Inhoff, Simon P. Liversedge, Erik D. Reichle, and Elizabeth R. Schotter. 2016. "Eye Movements in Eeading and Information Processing: Keith Rayner's 40year legacy." *Journal of Memory and Language* 86: 1–19.

Daneman, Meredyth, and Patricia A. Carpenter. 1980. "Individual Differences in Working Memory and Reading." *Journal of Verbal Learning and Verbal Behavior* 19 (4): 450–466.

Diessel, Holger. 2004. *The Acquisition of Complex Sentences*. Cambridge: Cambridge University Press.

Dragsted, Barbara, and Inge Gorm Hansen. 2009. "Exploring Translation and Interpreting Hybrids. The case of sight translation." *Meta* 54 (3): 588–604.

Dragsted, Barbara, Inger Margrethe Mees, and Inge Gorm Hansen. 2011. "Speaking Your Translation: Students' first encounter with speech recognition technology." *Translation & Interpreting* 3 (1): 10–43.

Fernández, Eva M. 2002. "Relative Clause Attachment in Bilinguals and Monolinguals." In *Bilingual Sentence Processing*. Edited by R. R. Heredia and J. Altarriba, 187–215. Amsterdam: North-Holland.

Frenck-Mestre, Cheryl. 2002. "An On-line Look at Sentence Processing in the Second Language." In *Bilingual Sentence Processing*. Edited by R. R. Heredia and J. Altarriba, 217–236. Amsterdam: North-Holland.

Frost, Ram. 2012. "Towards a Universal Model of Reading." *Behavioral and Brain Sciences* 35 (5): 263–279.

Grosjean, François. 2001. "The Bilingual's Language Modes." In *One Mind, Two Languages: Bilingual Language Processing.* Edited by J. L. Nicol, 1–22. Malden: Blackwell.

Her, Emily. 1997. "Binary Error Analysis of Sight Interpretation from English into Chinese and its Pedagogical Implications." *Studies of Interpretation and Translation* 2: 111–135.

Huang, Li-Shih. 2009. "The Potential Influence of L1 (Chinese) on L2 (English) Communication." *ELT Journal* 64 (2): 155–164.

International Medical Interpreters Association. 2009. "IMIA Guide on Sight Translation." http://www.imiaweb.org/uploads/pages/380_2.pdf. Accessed 12 October 2016.

International Organization for Standardization. 2014. "Interpreting—Guidelines for Community Interpreting." https://www.iso.org/obp/ui/#iso:std:iso:13611:ed-1:v1:en. Accessed 12 October 2018.
Jiang, Wenying. 2009. *Acquisition of Word Order in Chinese as a Foreign Language*. Berlin: De Gruyter Mouton.
Li, Xiangdong. 2014. "Sight Translation as a Topic in Interpreting Research: Progress, problems, and prospects." *Across Languages and Cultures* 15 (1): 67–89.
Lykke Jakobsen, Arnt and Kristian Tangsgaard Hvelplund Jensen. 2008. "Eye Movement Behaviour Across Four Different Types of Reading Task." *Copenhagen Studies in Language* 36: 103–124.
McDonald, Janet L., and Patricia A. Carpenter. 1981. "Silmultaneous Translation: Idiom interpretation and parsing heuristics." *Journal of Verbal Learning and Verbal Behavior* 20 (2): 231–247.
National Association of Judiciary Interpreters & Translators. 2006. "Modes of Interpreting: Simultaneous, consecutive & sight translation." https://najit.org/wp-content/uploads/2016/09/Modes_of_Interpreting200609.pdf. Accessed 12 October 2016.
National Council on Interpreting in Health Care. 2009. "Sight Translation and Written Translation: Guidelines for healthcare interpreters." https://www.ncihc.org/assets/documents/publications/Translation_Guidelines_for_Interpreters_FINAL042709.pdf. Accessed 12 October 2016.
Radach, Ralph, and Alan Kennedy. 2004. "Theoretical Perspectives on Eye Movements in Reading: Past controversies, current issues, and an agenda for future research." *European Journal of Cognitive Psychology* 16 (1-2): 3–26.
Rayner, Keith, Alexander Pollatsek, Jane Ashby, and Charles Clifton Jr. 2012. *Psychology of Reading*. 2nd ed. New York: Psychology Press.
Sampaio, Glória Regina Loreto. 2007. "Mastering Sight Translation Skills." *Tradução and Comunicação* 16: 63–69.
Shreve, Gregory M., Christina Schäffner, Joseph H. Danks, and Jennifer Griffin. 1993. "Is There a Special Kind of 'Reading' for Translation: An empirical investigation of reading in the translation process." *Target* 5 (1): 21–41.
Shreve, Gregory M., Isabel Lacruz, and Erik Angelone. 2010. "Cognitive Effort, Syntactic Disruption, and Visual Interference in a Sight Translation Task." In *Translation and Cognition*. Edited by G. M. Shreve and E. Angelone, 63–84. Amsterdam: John Benjamins.
Stubbs, Michael. 1983. *Discourse Analysis: The sociolinguistic analysis of natural language*. Chicago: University of Chicago Press.
Su, Wenchao, and Defeng Li. 2019. "Identifying Translation Problems in English-Chinese Sight Translation." *Translation and Interpreting Studies* 14 (1): 110–134.
Tsai, Jie-Li, Chia-Ying Lee, Ovid J. L. Tzeng, Daisy L. Hung, and Nai-Shing Yen. 2004. "Use of Phonological Codes for Chinese Characters: Evidence from processing of parafoveal preview when reading sentences." *Brain and Language* 91 (2): 235–244.
Viaggio, Sergio. 1995. "The Praise of Sight Translation (and squeezing the last drop thereout of)." *The Interpreters' Newsletter* 6: 33–42.
Viezzi, Maurizio. 1989. "Information Retention as a Parameter for the Comparison of Sight Translation and Simultaneous Interpretation: An experimental study." *The Interpreters' Newsletter* 2: 65–69.
Wang, Jia-yi, and Yan He. 2018. "How Effortful Are Interpreters in Translation Related Reading Tasks?: An eye-tracking study." *Journal of Literature and Art Studies* 8 (10): 1497–1508.
Weber, Wilhelm K. 1990. "The Importance of Sight Translation in an Interpreter Training Program." In *Interpreting: Yesterday, Today, and Tomorrow*. Edited by D. Bowen and M. Bowen, 44–52. Amsterdam: John Benjamins.
Willingham, Daniel T. 2017. *The Reading Mind: A cognitive approach to understanding how the mind reads*. San Francisco: Jossey-Bass.
Yang, Chin Lung, Charles A. Perfetti, and Ying Liu. 2010. "Sentence Integration Processes: An ERP study of Chinese sentence comprehension with relative clauses." *Brain and Language* 112 (2): 85–100.

Appendix 1. SiT performance: Scores on accuracy and style in each trial (on a scale of 0-5)

Rater	Rater 1								Rater 2								
Construct	Accuracy				Style				Accuracy				Style				Mean
Trial / Participant	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
NOV-1	4	3	3	2	4	4	4	3	5	5	3	4	4	4	5	5	3.875
NOV-2	3	4	2	1	4	3	3	3	5	5	4	3	4	4	3	4	3.4375
NOV-3	4	4	4	5	4	4	4	5	5	5	4	5	4	5	5	5	4.5
NOV-4	5	3	4	5	4	4	5	5	4	4	3	4	4	4	3	5	4.125
NOV-5	3	4	2	5	4	4	3	4	3	3	3	5	4	3	4	5	3.6875
NOV-6	2	5	5	4	3	4	4	4	3	5	5	5	3	5	4	4	4.0625
NOV-7	4	4	5	3	3	4	4	3	5	5	5	3	5	5	4	3	4.0625
NOV-8	3	4	2	5	3	3	3	4	4	5	3	5	3	3	3	5	3.625
NOV-9	5	4	3	2	4	4	3	3	5	4	3	4	5	4	3	5	3.8125
NOV-10	4	5	4	4	5	4	4	4	5	5	3	5	5	4	5	4	4.375
NOV-11	3	4	3	3	3	3	3	3	5	5	5	3	4	5	4	3	3.6875
NOV-12	3	4	2	5	3	3	3	4	3	4	3	5	3	3	0	3	3.1875
NOV-13	3	3	4	4	4	4	4	4	4	5	4	4	3	5	5	5	4.0625
NOV-14	5	4	4	4	4	4	3	4	5	5	4	4	4	4	4	4	4.125
NOV-15	3	4	5	4	4	4	4	3	4	5	5	5	5	5	4	5	4.3125
NOV-16	4	4	4	3	4	4	4	3	4	5	5	4	5	5	5	4	4.1875
NOV-17	4	3	4	5	4	4	5	4	3	5	5	5	4	4	4	5	4.25
NOV-18	5	5	5	4	4	4	4	4	4	4	4	5	0	0	0	0	3.25
UNT-1	4	0	1	2	4	4	3	3	3	2	3	4	0	0	2	0	2.1875
UNT-2	3	2	2	3	2	1	1	3	5	3	4	4	3	2	2	2	2.625
UNT-3	2	1	4	3	3	3	3	3	4	4	3	3	3	3	3	3	3
UNT-4	3	3	4	4	4	4	4	4	3	4	5	5	0	0	0	3	3.125
UNT-5	2	0	0	1	3	3	3	3	3	3	3	3	3	2	2	2	2.25
UNT-6	1	1	1	2	1	1	1	1	4	4	3	3	3	3	3	3	2.1875
UNT-7	3	3	4	4	3	4	4	4	4	4	4	4	3	3	3	3	3.5625
UNT-8	4	3	5	5	4	2	4	2	3	3	5	4	0	3	3	3	3.3125
UNT-9	3	5	3	5	3	3	3	4	3	5	3	5	2	3	2	3	3.4375
UNT-10	4	5	4	3	3	4	4	3	3	3	2	4	3	2	2	3	3.25
UNT-11	3	4	3	2	3	3	2	2	4	5	3	3	3	3	3	2	3
UNT-12	3	2	2	2	4	4	3	3	4	5	3	5	3	2	2	3	3.125
UNT-13	0	3	2	0	0	3	3	0	0	4	3	3	0	0	3	0	1.5
UNT-14	1	1	2	2	2	2	3	2	0	3	2	2	0	0	0	0	1.375
UNT-15	3	4	3	3	3	3	3	3	4	5	4	3	4	4	4	3	3.5
UNT-16	4	3	3	1	3	3	2	1	4	4	3	3	3	3	3	2	2.8125
UNT-17	3	3	4	4	3	2	4	4	3	3	3	2	3	3	3	3	3.125
UNT-18	4	5	5	5	4	4	4	4	4	5	5	5	4	4	3	3	4.25

NOV, interpreting students (novices); UNT, untrained bilinguals

Corresponding author's address

Chen-En Ho
Centre for Translation and Interpreting
School of Arts, English and Languages
Queen's University Belfast
5 University Square
Belfast
BT7 1PA
United Kingdom
c.ho@qub.ac.uk

Biographical note

Chen-En Ho, as known as Ted Ho, is Lecturer in Translation and Interpreting in the School of Arts, English and Languages, Queen's University Belfast. He is also a practicing interpreter and translator, with experience in conference and court interpreting and a variety of translation projects, such as books, magazines and game localisation. His research interests cover such topics as cognitive studies of translation and interpreting, T&I education and industry, and public service interpreting, with his focus currently landing on the cognitive aspect of T&I and students' learning motivation and employability.

Tze-Wei Chen is Associate professor in the Graduate Institute of Translation and Interpretation, National Taiwan Normal University. He is also an AIIC member. His major research areas include interpreting pedagogy, T&I as an industry, community interpreting and the cognitive process of interpreting.

Jie-Li Tsai is Associate Professor at Department of Psychology of National Chengchi University and serves as the director of Researcher of Center for Mind, Brain, and Learning at NCCU. In his Eye Movement and Reading Laboratory, the research works are mainly on psycholinguistics issues of reading Chinese and eye movement control and language processing. The research topics include lexical and contextual processing, parafoveal processing, semantic ambiguity resolution, reading non-native languages, and neural mechanisms of natural reading.