**QUEEN'S
UNIVERSITY
BELFAST**

What does professional experience have to offer?. An eye tracking study of sight interpreting/translation behaviour

Ho, C-E. (2021). What does professional experience have to offer?. An eye tracking study of sight interpreting/translation behaviour. *Translation, Cognition & Behavior*, 4(1), 47-73.
<https://doi.org/10.1075/tcb.00047.ho>

Published in:
Translation, Cognition & Behavior

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights
Copyright 2021, John Benjamins Publishing Company.
This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

What does professional experience have to offer?

An eyetracking study of sight interpreting/translation behaviour

Ho Chen-En
Queen's University Belfast

This study investigated the impact of professional experience on the process and product of sight interpreting/translation (SiT). Seventeen experienced interpreters, with at least 150 days' professional experience, and 18 interpreting students were recruited to conduct three tasks: silent reading, reading aloud, and SiT. All participants had similar interpreter training backgrounds. The data of the SiT task have been reported here, with two experienced interpreters (both AIIC members) assessing the participants' interpretations on accuracy and style, which includes fluency and other paralinguistic performance. The findings show that professional experience contributed to higher accuracy, although there was no between-group difference in the mean score on style, overall task time, length of the SiT output, and mean fixation duration of each stage of reading. The experienced practitioners exhibited more varied approaches at the beginning of the SiT task, with some biding their time longer than the others before oral production started, but quality was not affected. Moving along, the practitioners showed better language flexibility in that their renditions were faster, steadier, and less disrupted by pauses and the need to read further to maintain the flow of interpretation.

Keywords: sight interpreting/translation; cognitive process; reading ahead; chunking/segmentation; pausing behaviour; interpreting experience

1. Introduction

Expertise, i.e., "characteristics, skills, and knowledge that distinguish experts from novices and less experienced people" (Ericsson 2018, 3), has long been of interest to researchers in various disciplines, especially for such skill-based fields as sport, music, chess and, of course, translation and interpreting (T&I). T&I studies on expertise, mostly conducted via expert-novice comparisons, have greatly contributed to our understanding of expert performance on multiple fronts. For instance, the experts in Barik's (1975) study provided more accurate and idiomatic renditions in *simultaneous interpreting* (SI) than the amateurs, and included more harmless additions without compromising quality—this "may be attributable to their very expertise, which frees them from following the text too closely, in an almost literal fashion as is the case with amateurs" (284). Liu, Schallert & Carroll (2004) also reported that professional interpreters were not only more accurate but better at correctly identifying primary versus secondary information, whilst having a *working memory* (WM) of comparable size to two groups of students. This finding about WM is similarly substantiated in several studies, including Hild's (2011) experiment, in which experts were less disrupted by syntactic complexity due to their purpose-driven approach to comprehending complex structures, and Köpke & Nespoulous (2006), which found no clear expert advantage in simple and complex span tasks; the authors further assert that it is the experts' processing mechanism, at least semantics-wise, that is "fundamentally different" (2006, 15). However, Mellinger & Hanson (2019), with a meta-analysis of previous research on the relationship between SI and WM, reported on numerous studies with contradictory findings and confirmed that professional interpreters do possess larger WM capacity than other comparison groups and that WM is positively correlated with SI performance. Similarly, Díaz-Galaz, Padilla & Bajo (2015) recorded more accurate interpretation by professionals, who also exhibited better WM capacity and L2 (English) proficiency than trainees. Nonetheless, the experts still interpreted more accurately even when WM and language abilities were controlled for; hence the authors' suggestion

that “greater accuracy is in part a function of experience and expertise, allowing faster access to (prior) knowledge and efficient use of cognitive resources for better performance in the limited time available” (18). Injoque-Ricle et al (2015) also found that SI performance was positively correlated with amount of professional experience, and interpreters with high WM performed the SI task significantly better than those assigned to a low WM group.¹ Despite this fact, amount of professional experience was not significantly associated with WM tasks without articulatory suppression (AS), but it was with the same WM tasks that had AS, implicating that the “experience in simultaneous interpreting has no effect on storage and processing capacity of WM” (60) and, more importantly, that “SI ability is supported not only by the capacity of WM to store or process information, but also by the ability of the interpreters to cope with the articulatory suppression effect” (ibid., 61).

The mixed findings presented above demand a careful examination of WM, which has been studied using different tasks. Nour et al (2020) looked at three fundamental executive subprocesses—*Response-Distractor Inhibition*, *Shifting* and *Updating*—to identify whether there are professional interpreter advantages. The authors performed a meta-analysis of relevant studies and found little advantage afforded by professional experience. Specifically, there was no advantage on updating or shifting, and the results were inconsistent on response-distractor inhibition. Yudes, Macizo & Bajo (2011) also found that professional interpreters had advantages in cognitive flexibility but not in inhibition and reversed the finding by Timarová et al (2014).

With a slightly different focus, Hild (2014, 137–142) reported that interpreting experts differed from novices in that they frequently used metacognitive strategies and concentrated on the task instead of their mistakes and the negative emotions that followed. Experts were also better able to foster the necessary positive self-appraisal to remain task-focused, and to manage stress or anxiety (see also Kurz 2003); furthermore, they assessed themselves more accurately, attributed the task outcomes to such controllable factors as inadequate strategies, and encouraged self-improvement, rather than blaming external conditions including the speaker’s speech rate and deficiencies. Tiselius & Jenset (2011) compared the process and the product of SI using three groups: interpreters with no, short, and long professional experience. Each participant was rated on intelligibility and informativeness.² The results show that intelligibility did differ between participants with and without professional experience, but informativeness was significantly different in all combinations of the three groups in pairs. In addition, interpreters with extensive experience were distinguished from the no and short experience groups by a clearer preference to monitor the accuracy of their interpretation, “which might very well mirror Vik-Tuovinen’s experienced interpreters, who focused more on situational factors than less experienced interpreters” (Tiselius & Jenset 2011, 290).

In general, expert interpreters produce more accurate output, and converging evidence points to experts’ strategic competences and efficient allocation of attention resources to, and effective switching between, different sub-tasks during interpreting (see also Korpál & Stachowiak-Szymczak 2018 and Stachowiak-Szymczak & Korpál 2019 on shorter mean fixation durations by professional interpreters). By contrast, research on experts’ cognitive advantages is inconsistent, with some studies rejecting their superiority in language proficiency (Moser-Mercer et al 2000) or WM capacity (Köpke & Nespoulous 2006; Liu et al 2004) and others confirming the lead (for more discussion see Liu 2009; Sunnari & Hild 2010; Tiselius & Hild 2017), while still others present partial advantages (see Nour et al 2020).

The above discussion mostly centres on SI. Expertise research on *sight interpreting/translation* (SiT) largely remains an uncharted territory. The term SiT—also called *sight translation*, or *sight interpreting*, as advocated by some scholars (see Čeňková 2015)—is adopted here as it mostly appears in interpreting contexts under time pressure, traditionally in public service interpreting (National Association of Judiciary Interpreters & Translators 2006; Spitzl & Hlavac 2017; Txabarriaga 2009) and but also in conference interpreting (Li 2014; Setton & Dawrant 2016), but is also applicable in translation to boost productivity (Dragsted & Hansen 2009; Dragsted, Mees & Hansen 2011). SiT as a hybrid form of T&I presents challenges of its own. Gile (2009) illustrates its intricate nature, with four non-automatic *efforts* in the Effort Model for SiT—reading, memory, speech production, and

¹ In this study, professional experience is reflected via *years of experience* and *worked days per month*, but only the latter showed significant correlations with other variables.

² Intelligibility considers only the interpretation and includes indices such as fluency and clarity; informativeness, on the other hand, denotes the fidelity and accuracy of the output in relation to the original speech (Tiselius & Jenset 2011, 285).

coordination. Indeed, the requirement of a certain level of cognitive effort to coordinate reception, memory, and production simultaneously attests to the kind of extreme language use, similar to SI. The same concept of simultaneity is further elucidated by Pöckhacker (2016) that “the interpreter’s target-text production is simultaneous not with the delivery of the source text but with the interpreter’s real-time (visual) reception of the written source text” (20). Bóna & Bakti (2020) documented the high cognitive load of SiT, reflected in more pauses, disfluencies, and slower speech and articulation rates than consecutive interpreting, which is often considered harder than SiT. This feature of high cognitive load and the need for multitasking is all the more evident when inadequate reading proficiency lays extra burden on the coordination of various sub-tasks and negatively affects the SiT output (Nilsen & Monsrud 2015). Needless to say, the challenges become even greater when the pair of languages are structurally distinctive from each other—Chinese and English being a good example (Ho, Chen & Tsai 2020).

Although few SiT studies focus on expert performance, the findings so far suffice to justify a deeper exploration of how SiT is proficiently executed. Among these studies, Agrifoglio (2004) found strong visual interference, which plagued even professional interpreters, manifested via a high percentage of expression failures in their output. Also attesting to the effects of visual interference is Viezzi (1989), which investigated the recall of the content in several tasks by the participants and reported their lower retention rate after SiT than SI. The asymmetry between languages formed an additional challenge. Visual interference, together with its knock-on effect, namely shallow information processing caused by the ever-present source text, is also well documented in later studies (Lee 2012; Shreve, Lacruz & Angelone 2010, 2011; also see Chmiel, Janikowski & Cieślewicz 2020 for the same effect in SiT under extreme conditions, i.e., SI with text). After comparing professionals with trainees in *prepared* SiT, in which participants had time to preview the text to be interpreted, Lee (2012) concluded that trainees lack sufficient reading and time management skills to keep up with professionals during preparation, who read the whole text and left time for execution planning and quickly identifying primary versus secondary concepts. This also reflects the difference between the predominantly local problem-driven, bottom-up processing by trainees, and the top-down, holistic processing or problem-solving leveraged by experts (Koda 2005). As a result, the experts’ renditions in Lee’s study were faster and more natural and accurate with fewer major errors. Similar conclusions were reached by Akbari (2017), which ascertained that the experts generally read faster and finished the SiT task quicker and more accurately, with target language expressions of finer quality. This may possibly be due to their internalisation of a “lexical and syntactic compressions strategy” (p. 40). This method of pruning and refining messages helped the practitioners distance from the source text, whereas the trainees were restricted by the surface form and mainly engaged in literal translation. Chmiel & Lijewska (2019) also depicted a similar image in their novice-expert comparison. Although the two groups achieved the same level of accuracy and had comparable WM capacity, the experts completed the SiT task more efficiently, as shown in shorter viewing and translation time than the trainees. The authors argue that, by looking away from the text, the trainees tried to lessen the burden of parallel processing to focus on output production (393) and paid for accurate output by taking more time; the experts were more resistant to visual interference, reflected via a higher percentage of dwell time on the text. On the other hand, Su (2020) invited professional and novice interpreters for a prepared SiT task and found that the behaviours of both groups changed in reaction to the nature of different stages of the task (preparation and the actual SiT); the latter stage saw fewer fixation counts, shorter saccades, and longer fixation durations, indicating closer tracking of what in the text was being translated. Interestingly, the two groups also shared similar behaviour during the actual SiT stage, including fixation count and duration, saccade length, pupil dilation, and task time. However, in addition to providing more accurate interpretations, the experts tended to read further ahead than fixate on what would immediately be translated, while the trainees kept more closely to the region about to be rendered. This behavioural discrepancy reflects the experts’ utilisation of WM (see also McDonald & Carpenter 1981, in which the practitioners and amateurs finished SiT using similar task time, but on average the former group processed larger chunks of information and were more accurate).

In considering SiT as a cross-modal task (Havnen 2019; Vargas-Urpi 2019), two concepts deserving of particular attention are the role played by memory (specifically WM) and the coordination between receiving input via the visual channel and producing and monitoring output via the auditory channel (or kinetic changes as in signed language interpreting (Spitzl & Hlavac 2017)).

However, previous research mainly relies on the output and global indicators such as task time or reading time to make inferences. Despite a few eyetracking studies that look into the moment-to-moment changes, how SiT is done by experts mostly remains unclear and WM is largely missing in the picture. This study aims to fill this gap and include WM to understand its role in SiT.

While the aforementioned studies on expertise all recruit experts, the definition of *expert* varies due to a lack of publicly recognised standard (Tiselius & Hild 2017). Therefore, researchers mostly rely on social or secondary indicators, such as length of experience, certification, or membership of professional associations, which do not necessarily correlate strongly with “objective measures” of performance (Ericsson 2000, 190). Conflicting results from different studies and noticeable within-group differences (such as Felberg & Nilsen 2017 and Nilsen & Monsrud 2015) could plausibly come from the use of these indirect proxies. It is therefore advised to cautiously interpret results from different studies.

2. The present study

This study reports on the influence of professional interpreting experience on SiT and is an integral part of a larger project investigating silent reading (SR), reading aloud (RA), and SiT to understand the nature of each task and how experience and training affect the cognitive processes and the output. The discussion below centres on the (dis)similarities between experienced interpreters and interpreting students in SiT.

2.1 Participants

The current study gathered data from 18 postgraduate interpreting students and 18 experienced interpreters with at least 150 days’ professional experience. AIIC (2020) requires this level of experience for its membership; the same standard is adopted here to distinguish experienced interpreters from trainees. One of the professionals was later excluded from analysis as she had not been working as an interpreter for over a year by the time of the experiment, and her eye movement trajectories, SiT behaviour, and output quality were drastically different from the rest of her group. This is likely because interpreting is a highly specialised trade that requires constant exposure. All participants’ native language was Chinese, and they had normal or corrected-to-normal vision. Their informed consent was obtained prior to participation.

A background questionnaire was filled out by each person and the results are partly reported in Table 1 to provide an overview, including age, English (L2) proficiency, and WM capacity on a scale of 1–5, identified by an English complex reading-span task (details see Ho, Chen & Tsai 2020).

Table 1. Background information of participants

	NOV	<i>s.d.</i>	EXP	<i>s.d.</i>
age	25.94	4.93	35.15	3.76
WM (reading span)	2.67	1.08	2.62	0.86
English proficiency (IELTS)	7.58	0.55	8.08	0.47

*NOV, interpreting students (novices); EXP, experienced interpreters

The experienced interpreters ($M=35.15$ years) were significantly older than the trainees ($M=25.94$ years), $t(33)=6.19$, $p<0.001$. This was readily expected, as it takes quite some time to accrue the number of hours required to qualify as an experienced interpreter.³ Experienced interpreters’ mean IELTS score ($M=8.08$) was significantly higher than the other group ($M=7.58$), $t(28)=2.59$, $p=0.015$. On the other hand, WM capacity failed to show a significant gap, with the experienced interpreters scoring a marginally lower mean at 2.62 than the trainees at 2.67, $t(33)=0.15$, $p=0.88$. We can therefore venture to rule out the influence of WM on the results reported below. L2 proficiency, however, remains a factor to consider when interpreting the observed reading and SiT behaviours. While the professionals all had at least 150 days’ professional experience, 14 out of 18 novices (78%) fell in the region of 1–29 days, with another two (11%) having no experience at all, one (5.5%) had

³ AIIC’s criterion of 150 days’ experience approximates 900 hours of professional practice for the participants in this study.

30–59 days and another (5.5%) had 60–89 days.⁴ Lastly, all participants received highly similar, if not identical, interpreter training as they graduated from the same two postgraduate programmes in Taiwan, except one experienced interpreter who received her postgraduate-level interpreter training overseas.

2.2 Materials and experimental design

The three English texts of 175 words used in this study were excerpts from different speeches by the incumbent Director-General of the World Trade Organization. To limit the influence of domain expertise, all materials were carefully selected to cover only introductory information, such as the mission of the WTO and current development in global trade. All technical terms were excluded except two: multilateralism and Doha Development Agenda—for which Chinese translations were provided to the participants during the introduction to the experiment. The content was only modified when necessary to ensure the texts retained a high level of authenticity and style of diplomatic speeches, typically with more lengthy and complex sentence structures. The difficulty and readability of each text, ranked on a scale of 1–7, with readability and difficulty increasing towards the higher end, were examined by indicators in MS Word and human evaluators, who were four bilinguals without any interpreter training but with comparable language competence to the trainees in this study. Both sources reported similarities across all three texts, with the mean score of the three texts ranked at 3 and Cronbach's α reaching .977 among the scores given by human evaluators (for details about the texts, see Appendix 1).

Each participant had to engage with all three texts, each under a different condition: SR, RA, and SiT. For SiT, the participants had to interpret from English (L2) into Chinese (L1). The pairing of texts with tasks were counter-balanced within and across groups, thereby eliminating practice effects on any given material. Each task was preceded by a practice text and followed by two comprehension questions.⁵ Each text was divided into four parts, called trials, and projected onto four screens consecutively, with multiple complete sentences per screen. When moving on to a new screen, the participant had to fixate on the cross that would appear in the upper-left corner of the screen for the next part of the text to appear. The stimuli were displayed on a computer monitor (1024 × 768 pixels) in Courier font (22 point) with a grey background to reduce glare. The experiment used an Eyelink 1000 eyetracker to sample the participants' dominant eye movements at 1000 Hz. The viewing distance from the monitor and the eyetracker was about 70 cm.

2.3 Procedure

Instructions for the experiment were first shown on the screen, along with the two previously mentioned technical terms and their Chinese translations, and the procedure was verbally explained to the participant. A 9-point calibration then followed, and after successful calibration the participant would commence the formal experiment, which required one to engage in three self-paced tasks, SR, RA, and SiT, in turn. Each time a screen—four for each task—was completed, the old text would be replaced with a blank screen, and a cross would then appear in the upper-left corner; fixating on it, the next part of the text would appear. The participant was informed that returning to previous trials would not be possible. Finally and importantly, the participant was advised that there was an audience waiting for their interpretation in the SiT task, making it feel just like a genuine interpreting assignment (Setton & Dawrant 2016) and suggesting they should tackle the task as they would do when taking on a real case.

2.4 Data analysis

This study analysed the data of 18 interpreting students and 17 experienced interpreters. All 35 participants answered at least five out of six comprehension questions correctly. Aside from the total time used to complete each task and two indicators of SiT quality in the form of numerical ratings (on a scale of 0–5) on accuracy and style (for more details, see Ho, Chen & Tsai 2020), global and local eye

⁴ For students from the two institutions where the trainees in this study were recruited, it takes on average 4–6 years to obtain an MA degree in Interpreting. Thus, students do have some opportunities to take on real interpreting assignments even before graduation.

⁵ There were six comprehension questions in total. Answering at least five correctly was required for the participant's data to be included for further analysis.

movement indices were also analysed to paint a detailed picture of the cognitive processes.⁶ Global indicators of eye movement included fixation count and mean fixation duration in each trial; five local indicators were used to shed light on word-based reading features:

1. *first fixation duration* (FFD), defined as “the duration of the first fixation on a word independent of the number of fixations that are made on that word” (Rayner & Liversedge 2004, 61)
2. *gaze duration* (GD) is “the sum of all fixations on the word prior to moving to another word” (ibid.) in the first reading pass
3. *go-past time* (GPT) means “the time from first entering a region until moving past that region forward in the text” (ibid., 62)
4. *re-reading time* (RRT) represents the “second-pass reading time (the amount of time devoted to rereading a target region)” (ibid.)
5. *total viewing time* (TVT) adds up all fixations in a target region (equal to an English word in this study)

Together, FFD and GD comprise earlier stages of processing. Later stages of processing are represented by GPT, RRT, and TVT. The participants’ SiT recordings were also subjected to further analysis. In addition to the aforementioned quality assessment by two AIIC members, six native Chinese speakers helped mark on each participants’ transcript pauses of any kind they felt existing in the SiT recordings (see §3.3.2 below about the different use of *verbal gap* and *pause* in this study). This subjective evaluation may provide insight into the audience’s genuine perception of pauses (Segalowitz 2010) when interpreting service is used. An independent *t*-test was used to examine each index.

3. Results

The sections below will first present the findings from a macro perspective, such as the task times across groups and the quality and relevant features of the renditions. Word-based reading behaviours will then be reported, and then in the third section the focus narrows to participants’ reading and interpreting behaviours.

3.1 Global indices: task- and trial-based performance

The bird’s eye view of the two groups’ reading behaviours is captured in Table 2, shown as trial-based figures. None of the metrics were found to significantly tell the two groups apart, be it the total time used to finish the SiT task ($M=42.11$ s for NOV; $M=36.53$ s for EXP), $t(33)=1.61$, $p=0.12$, the number of fixations ($M=131.76$ for NOV; $M=116.49$ for EXP), $t(33)=1.13$, $p=0.27$, or the mean fixation duration ($M=264.2$ ms for NOV; $M=254.67$ ms for EXP), $t(33)=0.58$, $p=0.56$.

Table 2. Trial-based global measures in SiT

	NOV	s.d.	EXP	s.d.
total time (s)	42.11	11.10	36.53	9.31
number of fixations	131.76	44.39	116.49	34.83
mean fixation duration (ms)	264.20	54.21	254.66	41.09

NOV, interpreting students (novices); EXP, experienced interpreters

When the above results are examined together with the SiT performance reported in Table 3 (original data in Appendix 2), we can start to see some influence, potentially from professional experience or

⁶ *Style* refers to aspects not covered by accuracy in quality assessment, such as language expression, the use of pauses, voice projection, and pace of interpretation—the initial period of silence before rendition begins is also factored in as in real life this may also affect service users’ perception of quality.

language proficiency. Firstly, the overall quality, composed of accuracy and style, showed a significant difference between the interpreting students ($M=3.92$) and the experienced interpreters ($M=4.38$), $t(33)=3.55$, $p=0.001$. That said, the difference came from accuracy ($M=4.03$ for NOV; $M=4.57$ for EXP), $t(33)=5.08$, $p<0.0001$, not style ($M=3.82$ for NOV; $M=4.18$ for EXP), $t(33)=1.82$, $p=0.08$. It means that, after going through one-semester or year-long training, the trainees managed to match their experienced counterparts in paralinguistic and extralinguistic aspects. Moreover, the length of the renditions was equal between the two groups (represented by Chinese character count; $M=350.39$ for NOV; $M=364.12$ for EXP), $t(33)=1.24$, $p=0.22$. The number of Chinese characters were used instead of words to account for the cognitive effort required for speech planning and execution because, given the smallest meaning unit in Chinese can consist of more than one character (better understood as a morpheme, Rayner et al 2012), there could be some additional efforts involved in proceeding from knowing and planning what characters to put together in what sequence to form a complete and meaningful word to successfully saying them aloud with poise.

Table 3. Task-based SiT performance and ratings (on a scale of 0–5)

	NOV	<i>s.d.</i>	EXP	<i>s.d.</i>
overall performance	3.92	0.38	4.38	0.37
accuracy	4.03	0.35	4.57	0.26
style	3.82	0.63	4.18	0.55
character count (Chinese)	350.39	29.41	364.12	35.69

NOV, interpreting students (novices); EXP, experienced interpreters

Judging from the SiT output, the influence of professional experience or, potentially, language proficiency, only affected accuracy. Two other possibilities point either to the potent effects of training in developing interpreting skills or the professionals' strategic decision to prioritise accuracy over other aspects.

3.2 Local reading indices: word-based eyetracking data

Delving into different stages of the cognitive processes of SiT could potentially uncover the mechanisms at work behind the façade of the sameness in behaviour. As shown in Table 4, the trainees did consistently have marginally longer mean FFD than the practitioners, but an independent *t*-test failed to show any significant between-group difference ($M=254.35$ ms for NOV; $M=253.5$ ms for EXP, $t(33)=0.08$, $p=0.94$).

Table 4. Mean duration of word-based reading behaviours in SiT (ms)

	NOV	<i>s.d.</i>	EXP	<i>s.d.</i>
first fixation duration (FFD)	254.35	34.76	253.50	30.83
gaze duration (GD)	287.93	45.66	276.46	38.91
go-past time (GPT)	410.82	155.82	389.60	188.10
re-reading time (RRT)	455.43	173.77	447.02	157.51
total viewing time (TVT)	508.05	202.91	510.78	180.31

NOV, interpreting students (novices); EXP, experienced interpreters

In fact, all the indices in Table 4 present the same result, with no significant difference between the two groups. The trainees spent similar time to the more experienced cohort on GD ($M=287.93$ ms for NOV; $M=276.46$ ms for EXP, $t(33)=0.8$, $p=0.43$), GPT ($M=410.82$ ms for NOV; $M=389.6$ ms for EXP, $t(33)=0.36$, $p=0.72$), and RRT ($M=455.43$ ms for NOV; $M=447.02$ ms for EXP, $t(33)=0.15$, $p=0.88$). A *t*-test again confirmed the comparability of the two groups on TVT ($M=508.05$ ms for NOV; $M=510.78$ ms for EXP, $t(33)=1.35$, $p=0.19$). Up to this point, a same or largely similar reading approach seems to have been adopted by both groups.

3.3 Experienced practitioners' advantages in SiT

The following section focuses on the interpreting behaviours of both groups in SiT. To reiterate, the instructions were used to create a task environment similar to real-world scenarios—participants were told their interpretation was for a genuine audience and that excessive silence could lead to a negative impression. They were asked to stick to their normal work habits.

3.3.1 Reading ahead: flexibility afforded by experience

Reading is required in SiT before reformulation can take place. Here, *reading ahead* does not necessarily suggest previewing new information, but simply a temporal linearity of reading preceding rendition. Two types of metrics on reading ahead are reported below: (1) *initial eye-mouth span* (EMS), which gauges how much time passed from the text first appearing to the participant opening their mouth for the first time to render the text; (2) *subsequent EMS*, equal to the time gap between the utterances of every two Chinese characters, or words if the participant produced inseparable, continuous sounds.

The two groups did not significantly differ from each other on the initial EMS and more experienced participants spent marginally longer time during the initial preparation stage ($M=6.85$ s for EXP; $M=5.99$ s for NOV, $t(33)=0.33$, $p=0.74$).

Figure 1. Time in silence before oral rendition began; *diamond* = group mean; *dot* = data points that are outliers

The distribution plotted in Figure 1 better shows the variability of both groups. While half of the participants of each group had a comparable initial EMS ($Mdn=3.03$ s for NOV; $Mdn=2.84$ s for EXP)—which is also shown by the comparison of the group means—the third quartile value reveals an interesting insight: 75% of the interpreting students began rendering the text in 4.94 seconds, while the more seasoned counterparts went on reading for up to 8.46 seconds, indicating greater individual difference in the approach taken to ensure high-quality output. Considering this from another perspective, namely how much information was collated during the initial period of silence, the trainees on average had 22.57 fixations ($SD=28.98$), similar to the experienced practitioners' 25.88 fixations ($SD=27.58$) with no significant difference, $t(33)=0.35$, $p=0.73$.

Figure 2. Reading and collating information during the initial EMS; *diamond* = group mean; *dot* = data points that are outliers

Again, the boxplot for each group in Figure 2 shows rather homogeneous behaviour among most trainees and more varied approaches among the practitioners, although as above half of all participants were similar in this regard ($Mdn=12.88$ for NOV; $Mdn=11.75$ for EXP). The third quartile value again provides the insight: 75% of the trainees used up to 17 fixations before oral production commenced, while the practitioners used up to 37 fixations. Each text was divided as evenly as possible into four trials, giving an average of 43.75 words per trial. That means some practitioners read the majority of the text before their rendition started.

Once the oral production began, what ensued started to set the two groups apart. Figure 3 illustrates the reading behaviour during each *subsequent* EMS, i.e., how much was read before each Chinese character/word was uttered. The two groups were close, as all participants relied on anything between one to just above two fixations to main the flow of rendition ($M=1.81$ for NOV; $M=1.5$ for EXP), but this difference was significant, $t(33)=3.81, p<0.001$.

Figure 3. Information retrieval during each subsequent EMS; *diamond* = group mean

The distribution patterns in Figure 3 shows the opposite trend of Figure 2. Despite exhibiting more distinctive styles within the group at the outset, the practitioners' on-going SiT behaviour was rather consistent; by contrast, the trainees manifested a highly similar inclination to begin interpreting as soon as possible but afterwards they were more variable, and some started to waver and subsequently required more fixations along the way.

3.3.2 Interpreting behaviour in the process

Interpreting behaviour is the most readily available, and easily perceptible, index that leaves an impression on the users of interpreting services. Below the results of three different indicators are reported, among which verbal gap is the first point attended to. A verbal gap tells us how much time has elapsed between two utterances and therefore gives us some idea about one's speech rate. Verbal gap is an important factor that affects the users' perception, as too long or unnatural a gap or too much variance in the durations of gaps brings little confidence in the interpreter's ability to handle the task successfully. The literature on interpreter training repeatedly mentions that when conducting SiT interpreters should sound like they are simply reading the source text out loud and not performing a cognitively demanding task (Angelelli 1999; Chen 2015). Although a verbal gap seems equal to a pause, I use them differently here: a verbal gap is established by calculating the lapse of time between two utterances, whereas a pause is subjectively perceived by the listener and could be affected by many external factors, such as the speech context, overall speech rate and personal style of the speaker/interpreter. Therefore, verbal gaps that are not detected by the listener cannot be classified as pauses. Boomer & Dittmann (1962, 219) also suggests that "pauses cannot be abstracted from their speech context and studied in isolation". Previous studies have adopted various minimum lengths to define a pause, from no cut-off point (defined only by judgement), through 250 ms, 300 ms (Wang & Li 2015), and one second, to three seconds (overview in Hieke, Kowal, & O'Connell 1983). In this study, verbal gaps of 600 ms and longer were consistently detected by six native Chinese speakers. Thus, 600 ms was adopted as the minimum length for a pause.

As mentioned above, a verbal gap is the timespan between two utterances. The length of each verbal gap was thus calculated by identifying the gap between the completion of one utterance and the onset of the following utterance. The trainees had a mean verbal gap of 258.95 ms ($SD=33.23$ ms), against the more experienced group's 231.68 ms ($SD=24.5$ ms). This difference was small but statistically significant, $t(33)=3.66$, $p<0.001$. The verbal gaps that were 600 ms or longer were further labelled as pauses. The number of pauses were then divided by the number of verbal gaps to arrive at the percentage of verbal gaps being detected by listeners as the "interruptions" in each participant's SiT output. Similarly, the percentage of pauses made by the trainees ($M=13.9\%$, $SD=3.67\%$) was significantly higher than their counterparts' ($M=9.77\%$, $SD=2.95\%$), $t(33)=3.66$, $p<0.001$. The results on verbal gaps and pauses highlight an outstanding achievement of the practitioners: they had a much quicker pace and the audience sensed much fewer interruptions or "rests" along the way. What further adds to the gap between the two groups is the significantly different mean pause length, which was 1.31 s ($SD=0.28$ s) for the trainees and 1 s ($SD=0.2$ s) for the practitioners, $t(33)=3.68$, $p<0.001$. The finding suggests that the professionals not only were faster and stopped fewer times, but that when they did stop it took them much less time to resume, without sacrificing accuracy or paralinguistic and extralinguistic performance, as shown in their mean score on style.

Did the professional interpreters use their comparatively meagre number of pauses differently than their counterparts? The number of fixations in each pause is a proxy for the amount of information gathered in each pause. Fewer fixations suggest less information is gathered in each one. The practitioner group on average utilised fewer fixations ($M=3.15$, $SD=0.65$) than the trainees ($M=3.99$, $SD=1.11$), with a small but statistically significant difference, $t(33)=2.7$, $p=0.01$ (see Table 5).

Table 5. Number of fixations in pauses and their distribution

	NOV	EXP
median	1	1
3 rd quartile	1	1
max.	16	5
mean (<i>s.d.</i>)	3.99 (1.11)	3.15 (0.65)

The source data for this indicator were then analysed to obtain the values of different quartiles (Table 5). That is, each valid data point was treated as equally unique without averaging them to arrive at an individual mean and then a group mean. The finding suggests that both groups had only one fixation in 75% of the pauses; in the remaining 25%, the trainees used a maximum of 16 fixations, but the practitioners were able to maintain only five fixations maximum in all circumstances.

Finally, a multiple regression analysis was performed to understand which background factor or aforesaid index can explain the difference in accuracy. Part of the reason for conducting this analysis is to ascertain the influence of L2 proficiency (reported earlier to have reached a between-group significant difference). Variance inflation factor (VIF) of each factor/index was checked to avoid multicollinearity, and those with a score of four or above were re-examined to see if they could theoretically explain the variation of accuracy; if not, they were deleted (O'Brien 2007). The remaining factors entered into the analysis were *professional experience*, *IELTS score*, *reading span*, *character count*, *verbal gap*, *initial EMS*, and *average pause length*. The analysis shows statistical significance, $F(7, 27)=4.83$, $p=0.001$, $R^2=.56$, $R^2_{Adjusted}=.44$, but only professional experience could explain the variance of accuracy, $Beta=.07$, $t(34)=2.36$, $p=0.026$ (details in Appendix 3). However, bear in mind that the model only moderately explains the result of accuracy.

4. Discussion

The interpreting students kept up with the experienced interpreters in many aspects, including three global indices (task time, number of fixations, and mean fixation duration), paralinguistic and extralinguistic performance, fixation durations in all reading stages, and the time taken at the very beginning of the SiT task to plan for interpretation. That said, professional experience still had obvious benefits once the rendition began. Practitioners were significantly more accurate, with fewer pauses, fewer fixations to address all challenges they faced, and a faster, steadier speech rate. These findings are compatible with previous studies that it is the accuracy, not the overall speed or fluency, that tells professionals and trainees apart (McDonald & Carpenter 1981; Tiselius & Jensen 2011)—especially Su (2020), which investigates Chinese-to-English SiT, the same language pair as this study but with opposite directionality. As for WM capacity (represented by one's reading span here), the gap between the two groups did not reach significance (Table 1); in addition, the same indicator was not found by the multiple regression analysis to have contributed to the between-group difference in accuracy. Therefore, this study fails to corroborate Mellinger & Hanson's (2019) conclusion that professional interpreters possess larger WM capacity, which is positively correlated with interpreting performance⁷. The result of this study does confirm Nour et al's (2020) finding regarding no interpreter advantage on the updating ability (measured in this study by the complex reading-span task). Even so, professionals still outperform trainees in many ways.

Looking at this mixed picture, several points are worthy of discussion. Firstly, the number of words used may not be a top priority for professionals. Trainees' use of (marginally) fewer Chinese characters indicates their endeavour to avoid unnecessary verbosity; in contrast, the practitioners may have prioritised accurate rendition over extreme concision and thus they managed to free up some mental capacity to concentrate on the content, in turn leading to higher accuracy. On the other hand, because the difference in brevity between the groups was not significant, an alternative possibility is that the trainees had to devote extra cognitive resources to achieve the same level of conciseness as the professionals due to insufficient language command, thereby affecting other parts of their performance.

Secondly, although the multiple regression analysis pinpointed professional experience as the sole factor explaining the inter-group difference in accuracy, I would argue that language proficiency (or, more specifically, language flexibility) is still key. For one, this study used general texts that were

⁷ This does not negate the validity of Mellinger & Hanson's (2019) findings, but rather proves the point made earlier that WM capacity has been measured by various tasks in different studies, and thus a finer classification is required when making comparisons. A caveat is that SI performance was the topic in Mellinger & Hanson's discussion, whereas SiT is the target of analysis in this study, so the output quality and its relationship with WM are perhaps incomparable.

almost devoid of technical terms. Therefore, familiarity with the domain due to prolonged experience is unlikely to have overwhelming effects. Experience did buy some professionals more time to plan ahead without jeopardising the impression of their performance, but at least half still began the SiT task almost immediately, thus were constantly juggling reading and reformulating new information just like the trainees. Without finishing reformulation in advance, plus fewer fixations in each subsequent EMS (Figure 3), the practitioners would require not just professional experience but sufficient command of the language(s) involved to solve the challenges arising along the way. From Table 4 and Table 5, we know the experienced interpreters did not adopt a fundamentally different approach when reading and preparing their reformulation, such as speeding up information retrieval to read ahead as much as possible. Here, I argue that no participants scanned back and forth or read further ahead to plan their next move but that, regardless of group, they all rendered the text close to where they were reading. The between-group mismatch in 25% of the pauses possibly reveals the professionals' skilful chunking behaviour and language flexibility in rising to the challenges fast and decisively, echoing Tiselius & Hild's (2017, 429) claim that experts are "less likely to run into processing problems than their inexperienced counterparts, and when they do they have a wider variety of problem-solving strategies to choose from". I would thus contend that professional experience helps sharpen this skill set of chunking plus reformulating, supported by language flexibility. That is to say, through experience, professionals have built up sufficient knowledge of optimally chunking a sentence but also know how the same sentence can be segmented differently; accordingly, they can flexibly reformulate the message in response to whatever segment they have at hand. Therefore, language proficiency is the necessary, but not sufficient, condition for successful interpretation.

The variability of the practitioners' behaviour during the initial EMS warrants another mention. While showing the same overall tendency to initiate the rendition sooner rather than later, possibly due to the similar interpreter training background of the two groups, the practitioners showed more personal styles in their approaches, some of which are more strategic. Previous studies mention experts' situational awareness leading to strategic decisions and, in turn, higher accuracy (Su 2020; Tiselius & Jenset 2011). I would argue the prolonged start adopted by some practitioners in this study substantiates the use of situational awareness to some extent. Sufficient experience gave the practitioners more confidence, knowing that they could afford to dwell longer on initial preparation in exchange for rapid and accurate rendition later, without losing favour with the audience. While the trainees diligently applied the principle of "the sooner the better" learned in training, the practitioners had a better "grasp or reading of the atmosphere"—a kind of prior knowledge, not of the text but the interpreting scenario—to anticipate the unfolding of eventualities, enabling optimised decision-making (Endsley 2018).⁸

5. Conclusion

The findings of this study show little difference between the experienced interpreters and the interpreting students on global indices and local reading indices in all stages, which—together with high resemblance of the reading behaviour in pauses—indicate that the two groups adopted fundamentally similar reading approaches. Performance-wise, participants with extensive experience provided more accurate interpretations, but the mean score on style and the length of rendition (in Chinese characters) were on par. The two groups reacted similarly after the text first appeared on the screen, with at least half of each group trying to initiate the rendition almost immediately, which is probably the result of largely overlapping interpreter training background. On the surface, it seems prolonged professional experience mostly benefits accuracy; however, its influence had already emerged early in the SiT task, with the practitioners adopting more varied approaches in using the initial period of silence. I believe it is the years of immersion in real-case scenarios that gave the practitioners better ability to gauge the (imagined) audience's expectations, so some were confident

⁸ This is a prevalent principle adhered to by many trainers, at least for the English-Chinese language pair. With English being right-branching and Chinese left-branching, the adjectival and adverbial components of more structurally difficult sentences develop in opposite directions in the two languages (Ho, Chen & Tsai 2020). To avoid prolonged waiting time for the audience when the interpreter reads through the sentence, trainees are trained to begin whenever they have one or two translation units—the shorter, the better—they can confidently render.

enough to remain silent slightly longer to plan their production. More obvious inter-group dissimilarities also appeared after the onset of rendition: the professionals were faster, stumbled fewer times, and regained their feet much quicker than the trainees. With the similar reading approaches adopted by both groups in mind, I argue that language proficiency (not only in the source language, which clearly set the two groups apart here, but also the target language), and the flexibility that comes with it, is one of the major factors in the practitioners' success, while the situational awareness developed via extensive experience is also crucial.

In short, professional experience translates into higher accuracy and situational awareness that affords more strategic SiT planning and execution, which are reflected in the process once oral rendition begins. The implications of the current study are two-fold. Firstly, instead of training students to begin rendition as soon as possible indiscriminately, we should teach them to strike a balance between a speedy start and thorough preparation. Developing situational awareness is essential, which entails the sensitivity to all contributing factors in a particular setting or context to the success of the task. Secondly, (flexibility) training in both languages is equally important, even for postgraduate-level programmes, as it is the foundation of successful application of interpreting skills or strategies.

Without examining in detail what each group of participants actually looked at, the present study only provides an overall picture of the reading and interpreting behaviour in SiT. Thus, quantitative analysis to understand the nature of pre-reading and chunking and how the output is affected, coupled with interviews to triangulate the findings, should be conducted to further establish disparities and similarities between professional interpreters and trainees. The length of the material used is also shorter than what interpreters usually encounter. Longer texts could possibly reveal experience's true influence, as stamina, the mastery of the required skill set, and the ability to manage the interpreter-mediated event would all be crucial in keeping one's cognitive load under saturation level and finish the SiT task successfully.

References

- Agrioglio, 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>
- AIIC. 2020. "How to join AIIC." <https://aiic.ch/how-to-join-aiic/>. Accessed 23 August 2020.
- Akbari, Alireza. 2017. "Needed Skills and Strategies to Improve the Efficiency of Sight Translation in Classroom Context." *Trans-kom* 10 (1): 22–44.
- Angelelli, Claudia V. 1999. "The Role of Reading in Sight Translation: Implications for teaching." *The ATA Chronicle* 28 (5): 27–30.
- Barik, Henri C. 1975. "Simultaneous Interpretation: Qualitative and linguistic data." *Language and Speech* 18 (3): 272–297. <https://doi.org/10.1177/002383097501800310>
- Bóna, Judit, and Mária Bakti. 2020. "The Effect of Cognitive Load on Temporal and Disfluency Patterns of Speech: Evidence from consecutive interpreting and sight translation." *Target* 32 (3): 482–506. <https://doi.org/10.1075/target.19041.bon>
- Boomer, Donald S., and Allen T. Dittmann. 1962. "Hesitation Pauses and Juncture Pauses in Speech." *Language and Speech* 5 (4): 215–220. <https://doi.org/10.1177/002383096200500404>
- Čeňková, Ivana. 2015. "Sight translation/interpreting." In *Routledge Encyclopedia of Interpreting Studies*. Edited by F. Pöchhacker, 374–375. London: Routledge.
- Chen, Wallace. 2015. "Sight translation." In *The Routledge Handbook of Interpreting*. Edited by H. Mikkelsen and Renée Jourdenais, 144–153. London: Routledge.
- Chmiel, Agnieszka, Przemysław Janikowski, and Anna Cieśliewicz. 2020. "The Eye or the Ear? Source language interference in sight translation and simultaneous interpreting." *Interpreting* 22 (2): 187–210. <https://doi.org/10.1075/intp.00043.chm>
- Chmiel, Agnieszka, and Agnieszka Lijewska. 2019. "Syntactic Processing in Sight Translation by Professional and Trainee Interpreters: Professionals are more time-efficient while trainees view the source text less." *Target* 31 (3): 378–397. <https://doi.org/10.1075/target.18091.chm>
- Díaz-Galaz, Stephanie, Presentacion Padilla, and María Teresa Bajo. 2015. "The Role of Advance Preparation in Simultaneous Interpreting: A comparison of professional interpreters and interpreting students." *Interpreting* 17 (1): 1–25. <https://doi.org/10.1075/intp.17.1.01dia>
- Dragsted, Barbara, and Inge Gorm Hansen. 2009. "Exploring Translation and Interpreting Hybrids. The case of sight translation." *Meta* 54 (3): 588–604. <https://doi.org/10.7202/038317ar>

- 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.
- Endsley, Mica R. 2018. "Expertise and Situation Awareness." In *The Cambridge Handbook of Expertise and Expert Performance*. 2nd ed. Edited by K. A. Ericsson, R. R. Hoffman, A. Kozbelt and A. M. Williams, 1907–1978. Cambridge: Cambridge University Press.
- Ericsson, K. Anders. 2000. "Expertise in Interpreting: An expert-performance perspective." *Interpreting* 5 (2): 187–220. <https://doi.org/10.1075/intp.5.2.08eri>
- Ericsson, K. Anders. 2018. "An Introduction to the Second Edition of The Cambridge Handbook of Expertise and Expert Performance: Its development, organization, and content." In *The Cambridge Handbook of Expertise and Expert Performance*. 2nd ed. Edited by K. A. Ericsson, R. R. Hoffman, A. Kozbelt and A. M. Williams, 3–20. Cambridge: Cambridge University Press.
- Felberg, Tatjana R, and Anne Birgitta Nilsen. 2017. "Exploring Semiotic Resources in Sight Translation." *The Journal of Specialised Translation* 28: 230–249.
- Gile, Daniel. 2009. *Basic Concepts and Models for Interpreter and Translator Training: Revised Edition*. Amsterdam: John Benjamins.
- Havnen, Randi. 2019. "Multimodal and Interactional Aspects of Sight Translation – A critical review." *FITISPos International Journal* 6 (1): 91–106.
- Hieke, Adolf E., Sabine Kowal, and Daniel C. O'Connell. 1983. "The Trouble with 'Articulatory' Pauses." *Language and Speech* 26 (3): 203–214. <https://doi.org/10.1177/002383098302600302>
- Hild, Adelina. 2011. "Effects of Linguistic Complexity on Expert Processing during Simultaneous Interpreting." In *Methods and Strategies of Process Research*. Edited by C. Alvstad, A. Hild and E. Tiselius, 249–267. Amsterdam: John Benjamins.
- Hild, Adelina. 2014. "The Role of Self-Regulatory Processes in the Development of Interpreting Expertise." *Translation and Interpreting Studies* 9 (1): 128–149. <https://doi.org/10.1075/tis.9.1.07hil>
- Ho, Chen-En, Tze-Wei Chen, and Jie-Li Tsai. 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>
- Injoque-Ricle, Irene, Juan Pablo Barreyro, Jesica Formoso, and Virginia I. Jaichenco. 2015. "Expertise, Working Memory and Articulatory Suppression Effect: Their relation with simultaneous interpreting performance." *Advances in Cognitive Psychology* 11 (2): 56–63. <https://doi.org/10.5709/acp-0171-1>
- Koda, Keiko. 2005. *Insights into Second Language Reading: A Cross-Linguistic Approach*. Cambridge: Cambridge University Press.
- Köpke, Barbara, and Jean-Luc Nespoulous. 2006. "Working Memory Performance in Expert and Novice Interpreters." *Interpreting* 8 (1): 1–23. <https://doi.org/10.1075/intp.8.1.02kop>
- Korpal, Paweł, and Katarzyna Stachowiak-Szymczak. 2018. "The whole picture: Processing of numbers and their context in simultaneous interpreting." *Poznan Studies in Contemporary Linguistics* 54 (3): 335–354. <https://doi.org/10.1515/psicl-2018-0013>
- Kurz, Ingrid. 2003. "Physiological Stress during Simultaneous Interpreting: A comparison of experts and novices." *The Interpreters' Newsletter* 12: 51–67.
- Lee, Jieun. 2012. "What Skills Do Student Interpreters Need to Learn in Sight Translation Training?" *Meta* 57 (3): 694–714. <https://doi.org/10.7202/1017087ar>
- Li, Xiangdong. 2014. "Sight Translation as a Topic in Interpreting Research: Progress, problems, and prospects." *Across Languages and Cultures* 15 (1): 67–89. <https://doi.org/10.1556/Acr.15.2014.1.4>
- Liu, Minhua. 2009. "How do Experts Interpret? Implications from research in Interpreting Studies and cognitive science." In *Efforts and Models in Interpreting and Translation Research. A tribute to Daniel Gile*. Edited by G. Hansen, A. Chesterman, and H. Gerzymisch-Arbogast, 159–177. Amsterdam: John Benjamins.
- Liu, Minhua, Diane L. Schallert, and Patrick J. Carroll. 2004. "Working Memory and Expertise in Simultaneous Interpreting." *Interpreting* 6 (1): 19–42. <https://doi.org/10.1075/intp.6.1.04liu>
- McDonald, Janet L., and Patricia A. Carpenter. 1981. "Simultaneous Translation: Idiom interpretation and parsing heuristics." *Journal of Verbal Learning and Verbal Behavior* 20 (2): 231–247. [https://doi.org/10.1016/S0022-5371\(81\)90397-2](https://doi.org/10.1016/S0022-5371(81)90397-2)
- Mellinger, Christopher D., and Thomas A. Hanson. 2019. "Meta-Analyses of Simultaneous Interpreting and Working Memory." *Interpreting* 21 (2): 165–195. <https://doi.org/10.1075/intp.00026.mel>
- Moser-Mercer, Barbara, Ulrich H. Frauenfelder, Beatriz Casado, and Alexander Künzli. 2000. "Searching to Define Expertise in Interpreting." In *Language Processing and Simultaneous Interpreting*. Edited by B. E. Dimitrova and K. Hyltenstam, 107–131. Amsterdam: John Benjamins.

- 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 10th August 2018.
- Nilsen, Anne Birgitta, and May-Britt Monsrud. 2015. "Reading Skills for Sight Translation in Public Sector Services." *Translation & Interpreting* 7 (3): 10–20.
- Nour, Soudabeh, Esli Struys, Evy Woumans, Ily Hollebeke, and Hélène Stengers. 2020. "An Interpreter Advantage in Executive Functions?: A systematic review." *Interpreting* 22 (2): 163–186. <https://doi.org/10.1075/intp.00045.nou>.
- O'Brien, Robert M. 2007. "A Caution Regarding Rules of Thumb for Variance Inflation Factors." *Quality & Quantity* 41 (5): 673–690. <https://doi.org/10.1007/s11135-006-9018-6>
- Pöckhacker, Franz. 2016. *Introducing Interpreting Studies*. 2nd ed. London: Routledge.
- Rayner, Keith, and Simon P. Liversedge. 2004. "Visual and Linguistic Processing during Eye Fixations in Reading." In *The Interface of Language, Vision, and Action: Eye movements and the visual world*. Edited by J. M. Henderson and F. Ferreira, 59–104. New York: Psychology Press.
- Rayner, Keith, Alexander Pollatsek, Jane Ashby, and Charles Clifton Jr. 2012. *Psychology of Reading*. 2nd ed. New York: Psychology Press.
- Segalowitz, Norman. 2010. *Cognitive Bases of Second Language Fluency*. New York: Routledge.
- Setton, Robin, and Andrew Dawrant. 2016. *Conference Interpreting. A complete course*. Amsterdam: John Benjamins.
- 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.
- Shreve, Gregory M., Isabel Lacruz, and Erik Angelone. 2011. "Sight Translation and Speech Disfluency: Performance analysis as a window to cognitive translation processes." In *Methods and Strategies of Process Research*. John Benjamins.
- Spitzl, Karlheinz, and Jim Hlavac. 2017. "Sight Translating Interview Transcripts." In *UNHCR: Handbook for Interpreters in Asylum Procedures*. Edited by S. Pöllabauer and I. Topolovec, 121–136. Vienna: Frank & Timme.
- Stachowiak-Szymczak, Katarzyna, and Paweł Korpala. 2019. "Interpreting Accuracy and Visual Processing of Numbers in Professional and Student Interpreters: An eye-tracking study." *Across Languages and Cultures* 20 (2): 235–251. <https://doi.org/10.1556/084.2019.20.2.5>
- Su, Wenchao. 2020. *Eye-Tracking Processes and Styles in Sight Translation*. Singapore: Springer.
- Sunnari, Marianna and Adelina Hild. 2010. "A Multi-Factorial Approach to the Development and Analysis of Professional Expertise in SI." *The Interpreters' Newsletter* 15: 33–49.
- Timarová, Šárka, Ivana Čeňková, Reine Meylaerts, Erik Hertog, Arnaud Szmalec, and Wouter Duyck. 2014. "Simultaneous Interpreting and Working Memory Executive Control." *Interpreting* 16 (2): 139–168. <https://doi.org/10.1075/intp.16.2.01tim>
- Tiselius, Elisabet, and Adelina Hild. 2017. "Expertise and Competence in Translation and Interpreting." In *The Handbook of Translation and Cognition*. Edited by J. W. Schwieter and A. Ferreira, 423–444. New Jersey: John Wiley & Sons, Inc.
- Tiselius, Elisabet, and Gard B. Jensen. 2011. "Process and Product in Simultaneous Interpreting: What they tell us about experience and expertise." In *Methods and Strategies of Process Research: Integrative Approaches in Translation Studies*. Edited by C. Alvstad, A. Hild and E. Tiselius, 269–300. Amsterdam: John Benjamins.
- Txabarriaga, Rocío. 2009. "IMIA Guide on Medical Translation." https://www.imiaweb.org/uploads/pages/380_2.pdf. Accessed 15th March 2020.
- Vargas-Urpi, Mireia. 2019. "Sight Translation in Public Service Interpreting: A dyadic or triadic exchange?" *The Interpreter and Translator Trainer* 13 (1): 1–17. <https://doi.org/10.1080/1750399X.2018.1503834>
- 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, Binhua, and Tao Li. 2015. "An Empirical Study of Pauses in Chinese-English Simultaneous Interpreting." *Perspectives* 23 (1): 124–142. <https://doi.org/10.1080/0907676X.2014.948885>
- Yudes Gómez, Carolina, Pedro Macizo Soria, and María Teresa Bajo Molina. 2011. "The Influence of Expertise in Simultaneous Interpreting on Non-Verbal Executive Processes." *Frontiers in Psychology* 2 (309). <https://doi.org/10.3389/fpsyg.2011.00309>

Appendix 1. Features of the three English texts used in this study

	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 human evaluators	3	2.75	3.25

Appendix 2. SiT performance: Scores of accuracy and style for each trial (on a scale of 0-5)

rater		Rater 1								Rater 2								
construct		accuracy				style				accuracy				style				mean
trial		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	
participant																		
NOV-01		4	3	3	2	4	4	4	3	5	5	3	4	4	4	5	5	3.875
NOV-02		3	4	2	1	4	3	3	3	5	5	4	3	4	4	3	4	3.4375
NOV-03		4	4	4	5	4	4	4	5	5	5	4	5	4	5	5	5	4.5
NOV-04		5	3	4	5	4	4	5	5	4	4	3	4	4	4	3	5	4.125
NOV-05		3	4	2	5	4	4	3	4	3	3	3	5	4	3	4	5	3.6875
NOV-06		2	5	5	4	3	4	4	4	3	5	5	5	3	5	4	4	4.0625
NOV-07		4	4	5	3	3	4	4	3	5	5	5	3	5	5	4	3	4.0625
NOV-08		3	4	2	5	3	3	3	4	4	5	3	5	3	3	3	5	3.625
NOV-09		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
EXP-01		4	5	4	4	3	4	4	3	5	5	5	5	5	5	5	5	4.4375
EXP-02		5	5	4	5	4	4	4	5	5	5	5	5	5	5	5	5	4.75
EXP-03		5	5	5	5	4	5	5	4	5	5	5	5	5	5	5	5	4.875
EXP-04		4	4	4	5	4	4	4	4	5	4	3	5	3	5	3	4	4.0625
EXP-05		4	5	5	5	4	4	4	4	5	5	5	5	4	5	5	5	4.625
EXP-06		4	4	4	3	4	5	3	4	5	5	5	5	5	5	5	4	4.375
EXP-07		4	4	3	4	4	3	2	3	5	5	4	5	4	5	4	5	4
EXP-08		5	4	4	5	5	5	5	5	5	5	5	5	3	3	5	3	4.5
EXP-09		5	5	4	5	5	5	5	5	5	5	4	5	5	5	5	5	4.875
EXP-10		5	4	5	4	5	3	4	4	5	5	5	3	5	5	5	4	4.4375
EXP-11		5	5	5	5	4	5	5	5	5	4	5	5	5	5	5	5	4.875
EXP-12		4	4	4	5	4	4	4	4	5	5	3	5	5	5	4	5	4.375
EXP-13		4	4	5	5	4	4	4	4	5	5	5	5	4	3	4	5	4.375
EXP-14		3	4	4	5	4	4	4	4	5	5	4	5	5	5	4	5	4.375
EXP-15		4	4	4	4	4	3	3	3	5	5	4	4	3	3	3	3	3.6875
EXP-16		3	5	4	5	4	4	4	4	5	5	3	4	3	3	3	5	4
EXP-17		5	4	5	5	4	4	4	4	5	3	5	4	0	4	4	0	3.75

NOV, interpreting students (novices); EXP, experienced interpreters

Appendix 3. Results of multiple regression analysis on inter-group difference in accuracy

Regression statistics							
multiple R		0.745757214					
R square		0.556153822					
adjusted R square		0.44108259					
standard error		0.308089416					
observations		35					

ANOVA						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>significance F</i>	
regression	7	3.21129	0.458755965	4.833126534	0.001248313	
residual	27	2.56282	0.094919088			
total	34	5.77411				

	<i>coefficients</i>	<i>standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>lower 95%</i>	<i>upper 95%</i>	<i>VIF</i>
intercept	3.103840104	1.35728	2.286805246	0.030271246	0.318927023	5.888753185	679.2893511
professional experience	0.066626636	0.02828	2.35583749	0.025996759	0.008597776	0.124655497	1.859626741
IELTS	0.131631604	0.11979	1.09882588	0.281546118	-0.114163246	0.377426455	1.469346525
reading span	0.08170088	0.05638	1.449172813	0.158805472	-0.033976393	0.197378152	1.064367302
character count	2.42313E-05	0.00175	0.013850782	0.98905087	-0.003565351	0.003613814	1.184518708
verbal gap	0.000365851	0.00214	0.170950085	0.865537696	-0.004025276	0.004756977	1.719852975
initial EMS	0.00226468	0.00784	0.288834664	0.774913949	-0.013823207	0.018352567	1.26644436
pause length	-0.405188408	0.25619	-1.581581974	0.125390075	-0.930850659	0.120473843	1.942870757

Address for correspondence

Ho Chen-En
School of Arts, English, and Languages
Queen's University Belfast
34 Kingscourt Avenue, Belfast,
BT6 9AN, Northern Ireland, UK

c.ho@qub.ac.uk

Biographical note

Ho Chen-En (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 span from cognitive translation and interpreting studies, through T&I education and industry, to public service interpreting, with focus currently landing on the cognitive aspect of T&I and students' learning motivation and employability.

<https://orcid.org/0000-0001-7321-2029>