Eye tracking sight translation performed by trainee interpreters

AGNIESZKA CHMIEL, IWONA MAZUR

Adam Mickiewicz University, Poznań, Poland

Abstract

This paper presents results of an eye-tracking study involving sight translation. It was assumed that interpreting trainees at a more advanced stage of training would display more efficient reading patterns than their less experienced colleagues. Eighteen participants with either one year or two years of interpreting training were asked to sight translate a text from A language (Polish) into B language (English). The text included such independent variables as target sentence type (simple SVO sentences and complex non-SVO sentences) and low frequency lexical items. The dependent variables included measures assumed to indicate lexical access and syntactic processing, such as fixation count, fixation length and observation length. The study found no group effect in total task time and processing of lexical items, which indicates that one year of training might not be sufficient to show differences in the sight translation skill development. The study also revealed that sentence readability could be a better predictor of processing load than syntax and, as expected, that more readable sentences generated less cognitive load than less readable ones.

1.   Introduction

Sight translation seems to be an increasingly useful skill in the professional practice, for example, when interpreting speakers who read out their presentation slides or simply read their speeches. As a result, it is included increasingly in interpreting courses. It is thus worth researching this phenomenon to find out how the sight translation skill develops.

This paper presents results of a pilot eye tracking study involving sight translation performed by interpreting trainees. We assume that trainees at a more advanced stage of training will invest less cognitive effort in the assigned task than their less experienced colleagues. Also, by selecting appropriate variables we expect to discover which features of a written text present the greatest difficulty to the students and require more cognitive load and

whether it is possible to establish at which stage of training particular issues are mastered when students at different stages of training are compared.

2.   Sight translation – practice and research

Sight translation (ST) involves transposing a message in the source language expressed in writing into a message in the target language rendered orally. It therefore engages two types of information processing: while the source text is fed in using the visual channel, the output of the translator is oral. Seen in that way, ST is sometimes treated as a hybrid between translation and interpreting, as what is translated is a written text, which however is rendered orally (as in interpreting) and not in writing (as in translation). Also, in simultaneous interpreting the text is translated in real time, i.e. the translator reads, processes and produces the target text almost at the same time or with a slight time lag. When it comes to the actual application of sight translation in professional settings, it is used, for example, in the courtroom where the interpreter is often asked to sight translate procedural documentation for the judge (cf. Lambert 2004: 298). In cognitive terms, sight translation places a lot of strain on the processing capacity of the interpreter as it involves a number of simultaneous subtasks, such as processing the visual input, producing the oral output, and monitoring one's output aurally, and as such could in fact be compared to simultaneous interpreting (ibid.).

What is more, Lambert (ibid.) differentiates between sight translation and sight interpreting suggesting that in the latter mode the interpreter interprets a speech that is fed in aurally, while at the same time they have to process some visual input in the form of slides or a written speech that is being read out by the speaker (this is referred to simultaneous interpreting with text by Pöchhacker (2004)). This mode of interpreting seems to pose an even greater cognitive challenge to the interpreter, as in addition to the simultaneous subtasks mentioned above, in this case we also have yet another such subtask in the form of the processing of the aural input.

Despite the fact that both sight translation and sight interpreting are used in the professional practice, they seem to be rarely taught on their own to prospective interpreters in courses and programmes known to the authors. Admittedly, more and more interpreting courses incorporate speeches that are accompanied by slides, but when it comes to sight translation, it appears that it is used as a method in interpreter training (usually treated as preparation for simultaneous interpreting), rather than taught as a separate skill that can be usefully applied on its own in the professional world.

Sight translation is generally much less researched than more popular modes of conference interpreting, i.e. consecutive and simultaneous. In an analysis of constraints, obstacles, efforts and failures, Agrifoglio (2004) compared these three activities and found a considerable difference in meaning and expression problems professional interpreters demonstrated in sight translation as opposed to consecutive (CI) and simultaneous interpreting (SI). Although ST seems an easier task due to lower memory saturation, the constant presence of the source text can also be an obstacle to expression in the target language since the interpreter has to suppress the source language interference. The coordination effort is focused on silent reading and expression. According to Agrifoglio (ibid.: 61), "visual interference seems to be stronger than audio interference, and the sight translator may have to devote more effort to resisting this influence in ST than in SI, sometimes at the expense of fluency", which results in a higher number of expression failures in ST (75% of all failures) than in SI (37%) and CI (24%) as reported by the author.

Lambert (2004), on the other hand, embarked on a study to determine the role of visual presentation of the source text in sight translation by comparing it to normal simultaneous interpreting and SI with text. Lambert (ibid.: 302) evaluated the participants' output and found, in line with Agrifoglio's findings, a higher performance score for ST (82%) than for SI (69%). Lambert explains her results by pointing to more attention devoted by subjects to input processing. Unlike SI, in sight translation attention does not have to be shared and less cognitive and memory load is required by this task.

Viezzi (1989) researched information retention after a sight translation task as compared to simultaneous interpreting. Surprisingly, his study participants remembered more after SI than after ST. Viezzi explained the finding in line with Craik and Lockhart's theory of levels of processing (1972), i.e. that retention is higher after SI since it involves deeper processing of incoming information. Another interesting finding of Viezzi's study was the correlation between morphosyntactic differences in the language pair and retention rates (1989: 59). The interpreters examined remembered more after ST from such morphosyntactically closer languages as Spanish into Italian (85%) than after ST from German into Italian, i.e. a language pair with less morphosyntactic similarity (64%). It seems that the interpreters can retain more in the sight translation task if it involves less transformation due to morphosyntactic similarities between the languages.

All in all, the above-mentioned findings suggest that ST does involve different cognitive processes than, for example, simultaneous interpreting, and since ST – unlike SI – involves the visual channel, new possibilities have opened up to probe into the black box of the sight translator with the coming of the eye tracking methodology.

3.　　Skill development

Conference interpreting is considered a complex cognitive phenomenon including many individual skills that have to be mastered to achieve a professional level of expertise (Moser-Mercer et al. 1997, Chmiel 2006). It is assumed here that interpreting (including sight translation as one of its varieties) is developed as a skill in interpreter trainees in the course of their practice through automatisation of various processes (Chmiel 2006), and that as an automatised process, interpreting performed by experts requires fewer attentional resources and cognitive effort as compared to interpreting performed by less skilled trainees. For a skill to develop, it has to be modified in three major stages: declarative and controlled processing that is slow, prone to errors and generates a lot of cognitive load; associative processing that is more procedural but still involves attentional resources for control; automatic processing that generates much less cognitive load and engages fewer attentional resources (for a detailed overview of a skill-based approach to conference interpreting see Chmiel 2006). It is assumed for the purpose of this study that trainees with a year's difference in their exposure to training and skill development will differ in the amount of cognitive load identified via eye tracking in the experimental task. If the difference is significant enough it will be possible to identify to what extent the sight translation skill develops at the two stages of training concerned.

4.　　Eye tracking – a new trek in process-oriented translation research

Process-oriented translation research has been conducted for over 20 years now, but thanks to the application of the eye tracking methodology the researchers have obtained a new window to the translator's mind. With the use of an eye tracker one can record the so-called *fixations* and *saccades* of the subject. "When we read, look at a scene or search for an object, we continually make eye movements called *saccades*. Between the saccades, our eyes remain relatively still during *fixations* for 200-300 ms" (Rayner 1998: 373, original emphasis). Fixations (but also saccades, e.g. in the form of regressions) are said to be representative of the amount of cognitive effort invested in processing given visual information (cf. Rayner 1998; Pavlović and Jensen. 2009; Doherty and O'Brien [no date]). For a few decades now, eye tracking has been extensively applied to research on reading patterns (for an overview see Rayner 1998), and we have seen the application of this methodology to translation research in recent years.

　　　Alves et al. (2010), for example, used eye tracking to research the performance of professional translators versus domain specialists, Pavlović and Jensen (2009) used it to research translation directionality, and Sharmin et al. (2008) to test the effects of time pressure

and text complexity on the performance of translators. Doherty and O'Brien ([no date]), on the other hand, applied eye tracking to the evaluation of machine translation output. More importantly for the subject-matter of this paper, Shreve et al. (2010) and Jakobsen and Jensen (2009) used eye tracking in their studies of sight translation that are briefly outlined below.

Shreve et al. (2010) set out to determine the cognitive differences between performing sight translation and written translation. They assumed that the comprehension effort could be disrupted by the manipulation of the source text, however the degree of such disruptions will be different in sight translation and written translation. In the study the source texts were manipulated in such a manner that they included both a paragraph with complex syntax and a paragraph written in simple syntax and "[e]ach text was sight translated and written translated by different participants and each participant performed both tasks (sight and written), but performed only one such task on any given text" (ibid.: 68). The independent variable was thus syntactic complexity, whereas dependent variables were eye movement patterns and errors in verbal output (for sight translation) and keystroke log patterns, pauses, and errors in written output (for the written translation). The study corroborated the hypotheses that in sight translation the complex syntax required more processing effort than non-complex syntax and that syntactic manipulation seems to have a greater impact on performance in sight translation than in written translation, as in the latter case the task exerts less pressure on short-term memory and the cognitive resources can be reallocated more easily over time. The results also indicated that sight translation is very sensitive to visual interference because of the constant presence of the source text.

Jakobsen and Jensen (2009) undertook to investigate whether visual attention (as manifested by eye movement patterns) would be distributed differently in four types of reading task, i.e. plain reading for comprehension (task 1), reading with an aim to later translate the text (task 2), reading the text and translating it orally at the same time (sight translation) (task 3) and reading while typing a written translation (task 4), and whether there would be differences in the cognitive effort expended by the two groups of subjects involved, i.e. professional translators and translation students. Given the purpose of our article, we will only focus on the data pertaining to sight translation. Jakobsen and Jensen (ibid.) found that, on the whole, sight translation required more cognitive effort than both tasks 1 and 2, as evidenced by longer task time and more fixations. What is more, when the average task time for sight translation was compared across the two groups of subjects, it turned out that the average task time for professionals was 154 seconds while for translation students it was 204 seconds (ibid.). This suggests that with practice the task of sight translating becomes less cognitively demanding. This could be treated as a point of departure for our study, whose objectives (research questions) are outlined below.

5.   Research questions

Given all of the above considerations concerning both the skill of sight translation as well as the new opportunities that are opened up by the eye tracking methodology, we set out to answer the following research questions:

1. How does the skill develop in the course of training?

2. Are reading patterns more efficient, i.e. do they generate less cognitive load for

   more advanced trainees?

3. Which text features (syntactic structures, lexical items) present the greatest

   difficulty for trainees?

6.   Research design and methodology

In order to answer the above-mentioned research questions, a study was conducted at Adam Mickiewicz University in June 2009. In the study two groups of participants, varying in the length of interpreter training that they had undergone, were asked to sight translate a source text in Polish into English, where both syntactic and lexical complexity had been manipulated. What follows is a detailed description of the source text, participants, equipment and settings, eye tracking metrics, and procedure.

6.1   Source text

We designed a study in which we used a text in Polish about swine flu. As swine flu was a topical issue at the time of conducting the experiment, we assumed that the students would be familiar with the subject matter. The text consisted of two pages and was 345 words long. We arranged the layout of the text as recommended by Gerganov (2007), i.e. font type and size was Tahoma 20, there were 90 characters per line, and the text was double spaced.

       The text was a newspaper article, but it was manipulated in such a way that it included both simple SVO sentences and complex non-SVO sentences, which were designated as our

areas of interest (AOIs) (which were of similar length, i.e. the simple SVO sentences AOIs totalled 68 words and the complex non-SVO sentences AOIs totalled 70 words). The reason why we chose such AOIs is that word order in Polish it is not as fixed as in English, and thus non-SVO sentences are quite common. We thus assumed that translating a non-SVO sentence into English will require greater processing effort when compared to translating SVO sentences. Similarly, we assumed that less cognitive effort will have to be expended to translate simple sentences in comparison with complex sentences (i.e. ones with embedded relative clauses), as in the case of the latter one usually has to read the whole sentence in order to grasp its meaning and store it in the short-term memory before translating it. In fact, this variable may be seen as combining two variables, i.e. sentence type (simple and complex) and word order (SVO and non-SVO). However, as we were not particularly interested in grammatical features, but rather in sentence difficulty as a factor influencing cognitive load, the design did not include all four potential conditions (i.e. simple SVO, simple non-SVO, complex SVO, complex non-SVO). Such a design can be used in further studies of sight translation with eye-tracking methods[1].

In addition, we designated as AOIs seven low frequency ('LF') lexical items: *układ odpornościowy* ('immune system'), *wydzielina* ('secretion'), *drogi oddechowe* ('respiratory tract'), *rozbicie* ('feeling under the weather'), *powikłania* ('complications'), *zapalenie oskrzeli* ('bronchitis), *zapalenie mózgu* ('brain inflammation'). The frequency of the words was verified in the IPI PAN corpus of Polish[2], where the average frequency of the lexical items was 1.7 per million words. All of the designated AOIs constituted our independent variables.

6.2   Participants

A total of 18 participants took part in the study (3 males and 15 females whose average age was 23.5). They were all full-time graduate students enrolled in a two-year conference interpreting programme at Adam Mickiewicz University (with Polish as their A language and English as their B language) and fell into two groups: one consisting of ten students in the first year of the course (end of second semester), who had completed approximately 600 contact hours of interpreter training (year 1, Y1), and the other consisting of eight students in their second year of the course (end of fourth semester), having completed some 1200 hours of such training (year 2, Y2). For the purposes of our study, the independent between-subjects variable was thus the duration of interpreting training. It should also be noted that prior to the

---

[1] We thank an anonymous reviewer for that comment.

[2] Available at http://korpus.pl (last accessed 1.12.2011).

experiment both first and second year students had been exposed to sight translation in the course of their training. Sight translation was used in various consecutive interpreting classes as an additional technique, but no separate course in sight translation was offered.

6.3   Equipment and settings

To record gaze behaviour of the participants we used the Tobii T60 eye tracker, which is a remote tracker allowing for unrestrained head movements. The screen resolution was set at 1280x1024 pixels and the fixation radius was 35 ppi (pixel per inch) (which is Tobii's default setting). The viewing distance was 50-60 cm from the screen. The collected data was analysed using Tobii Studio 2.0.8 analysis software.

6.4   Eye tracking metrics

The overriding assumption for this study was that obtaining observable, measurable data by means of eye tracking will help us uncover the underlying cognitive processes of our subjects. We used three kinds of measures in order to gain insight into the participants' cognitive processes:

1. 'fixation count', that is the number of fixations within an AOI;

2. 'fixation length', which is the length of the fixations within an AOI;

3. 'observation length', which is total time a person has looked within an AOI,

    starting with a fixation within an AOI and ending with a fixation outside the AOI;

The above-mentioned measures are said to be indicative of the cognitive effort invested by the subjects in processing the stimulus information, i.e. it is assumed the greater the values obtained for the measures, the more cognitive effort was invested by the participants in processing target information (cf. Rayner 1998; Pavlović and Jensen. 2009; Doherty and O'Brien [no date]).

    We analysed eye tracking data related to fixation count for warm-up, total task, the sentences and LF items, to fixation length for the sentences and LF items, and to observation length for total task, and as such these constituted our dependent variables.

To sum up, in the experimental design the variables were as follows:

1. Independent between-subject variable: duration of interpreting training

2. Independent variables: target sentence type (simple SVO sentences, complex non-SVO sentences), low frequency lexical items

3. Dependent variables: fixation count, fixation length, observation length

6.5   Procedure

Each participant was tested individually. They were seated in front of the screen at the required distance and their eyes had been calibrated to the eye tracker. They were presented with the same set of instructions in Polish (both orally and then again in writing on the screen). They were asked to sight translate the text from Polish into English. Before the participants started doing the sight translation proper they were given a warm-up task to help them become accustomed to the experimental settings, the eye tracker, etc. During the warm-up task the subjects could scan the first page of the stimulus source text for ten seconds. The rationale behind the warm-up task was that in a professional conference setting that interpreters frequently obtain speech texts shortly before they are delivered and interpreted. Thus, interpreters should be able to use this opportunity to familiarise themselves with the text to the greatest possible extent, i.e. by scanning the text for gist rather than by thorough and linear reading. After the warm-up the participants proceeded to sight translate the stimulus text. They had unlimited time to perform the translation and each session lasted approximately thirty minutes.

Prior to the analysis, the recordings were checked for Brownian motion, which can be detected by replaying an eye tracking session on the analysis software and it is characterised by erratic vertical (rather than horizontal) saccade-like movements linking fixations. Such data are considered to be rich in noise and thus deemed unsuitable for analysis (cf. Pavlović and Jensen 2009). As no such motion had been detected in our pool of data, all of the data could be subject to further analysis.

6.6   Hypotheses

The study was conducted to test the following four hypotheses.

H1: Y2 will demonstrate more fixations in the warm-up task than Y1.

We believe that more experienced trainees would use the time shortly before the sight translation task for gist reading to obtain as much general information about the text as possible. Beginners, on the other hand, would engage in linear reading from the very beginning of the text. This assumption is based on the authors' observation, during conference interpreting classes they conduct, that novices prepare the target language rendering of the first sentences instead of reading for gist in the preparation time before the ST task. The fact that Y2 would scan a greater amount of the text should result in a higher number of fixations.

H2: The total task time will be shorter for Y2 than Y1.

More advanced students will process the text and produce the translation more efficiently and quickly since they have had more exposure to interpreting tasks than their younger colleagues.

H3: There will be a general sentence effect across groups, i.e. simple sentences will generate fewer and shorter fixations than complex sentences.

It is assumed that the higher processing load required for the sight translation of complex sentences as compared to simple sentences will result in more numerous and longer fixations (in line with Rayner 1998).

H4: We will discover a group effect with fewer and shorter fixations on low-frequency items for Y2 compared to Y1.

As in the case of previous independent variables, advanced trainees should demonstrate less cognitive load (i.e. a lower number of fixations and shorter fixation durations) in processing more difficult lexical items in the text.

7.   Data analysis and results

We ran a series of independent samples t-tests for hypotheses including a comparison of both experimental groups and 2x2 mixed ANOVAs for sentence and group effects in the third hypothesis.

In the warm-up task, advanced students fixated less frequently (40.80 fixations on average) than novices (48.38), which is contrary to the assumption in H1 (Table 1). An independent samples t-test revealed no statistical significance of these results (t(16)=1.46, p>.05) so there is no difference between Y1 and Y2 in the number of fixations during the warm-up task.

Table 1. Fixation count in the warm-up task – descriptive statistics (SD – standard deviation, SEM – standard error of the mean).

Group	Mean	SD	SEM
Y1	48.38	15.380	5.438
Y2	40.80	5.308	1.679

We have used the heatmap tool in Tobii Studio to visualise the above data. Heatmaps are visualisations in which individual values are presented as colours on a scale from green (signifying less intensity) to red (signifying greater intensity). In fact, even a cursory glance at the heatmaps generated for all eye tracking data for Y1 and Y2 (Figure 1) shows that less advanced students used more scanning than their older colleagues. The gazes are more widely distributed in the heatmap for Y1 and more students managed to scan the bottom part of the text. It is also important to notice that the colour scale is different. Red symbolises 6 fixation counts for Y1 and 8 counts for Y2, which serves as further evidence of a wider gaze distribution for Y1.

Figure 1. Heatmaps for the warm-up task performed by Y1 (left) and Y2 (right).
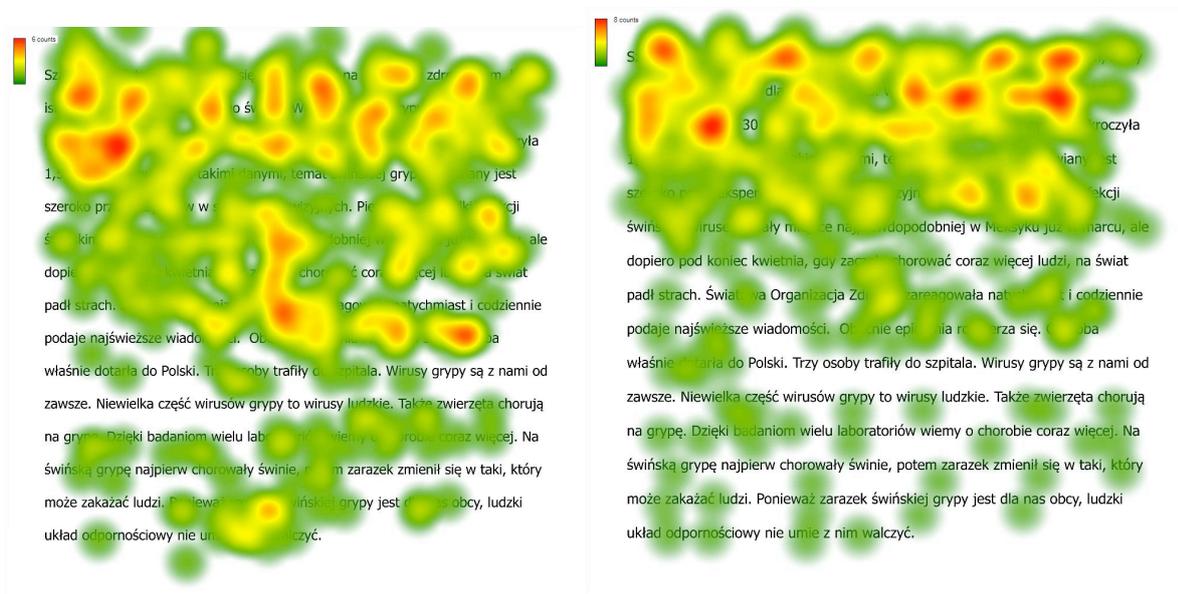
We used Tobii Studio's observation length data to measure total task time for both groups and found no group effect in an independent samples t-test (t(16)=1.81, p>.05) although Y2 performed sight translation slightly faster (351 seconds) than Y1 (387 seconds). Similarly, no group effect was found for the total number of fixations in the whole task (t(16)=.73, p>.05). Table 2 below presents detailed descriptive statistics.

Table 2. Observation time and fixation count for the whole ST task.

	Group	Mean	SD	SEM
Observation time	Y1	387	50.37	17.81
	Y2	351	33.71	10.66
Fixation count	Y1	995	219	77.49
	Y2	926	150	47.46

These data lead to the rejection of Hypothesis 2.

We further examined the length and number of fixations by both groups performed on target sentences. Table 3 presents the number of fixations by both groups on simple SVO and complex non-SVO sentences and the combined length of those fixations. We calculated these by identifying Areas of Interest on those parts of the screen that displayed the target sentences

Table 3. Mean count and length of fixations on target sentences.

	Group	Simple sentences	Complex sentences
Fixation count	Y1	193	179
	Y2	178	173
Fixation length	Y1	70	62
	Y2	61	54

Simple sentences generated more and longer fixations than complex sentences for both groups. The mean number of fixations for simple sentences was 193 by Y1 and 178 by Y2 whereas the number of fixations for complex sentences was 179 by Y1 and 173 by Y2. As for the fixation length, simple sentences generated fixations that lasted for 70 seconds by Y1 and 61 seconds by Y2, while fixations on complex sentences lasted for 62 seconds in the case of Y1 and for 54 seconds in the case of Y2. Fixations generated by Y1 were longer and more numerous for both types of sentences when compared to Y2. We conducted a 2 x 2 mixed ANOVA with group as a between-subjects variable and target sentence type as a within-subject variable. In analysing data for fixation count we found neither a sentence effect with $F_{(1,16)}=2.06$, $p>.05$ nor a group effect with $F_{(1,16)}=0.38$, $p>.05$. The ANOVA for the fixation length data revealed a sentence effect with $F_{(1,16)}=11.89$, $p<.05$ and no group effect with $F_{(1,16)}=3.57$, $p>.05$.

The analysis shows that simple sentences required more cognitive load than complex sentences and we found a sentence effect contrary to the assumption in Hypothesis 3 and our intuition. We thus decided to redefine our variable by examining the readability of target sentences, which might prove more important than their syntactic complexity.

Since no automatic readability tools are available for Polish, we calculated the index manually by counting the average sentence length and the average number of syllables per word and using the Flesch formula[3]. According to the reading ease (RE) index, the higher the RE (ranging from 0 to 100), the easier the text is to read. The index for simple sentences was 10 and for complex sentences was 22, suggesting strongly that complex sentences, despite their non-SVO word order, were more readable than seemingly easier simple sentences. We changed Hypothesis 3 accordingly:

H3: There will be a general sentence effect across groups, i.e. more readable (complex non-SVO) sentences will generate fewer and shorter fixations than less readable (simple) sentences.

---

[3] Available at http://www.readabilityformulas.com/flesch-reading-ease-readability-formula.php (last accessed 22.08.2012).

H3 in the above wording was partially corroborated in the analysis. The fixations for more readable sentences were longer, but not more numerous.

The last hypothesis pertained to the processing of low-frequency lexical items. We predicted a lower cognitive load manifested through less numerous and shorter fixations on AOIs with target words by Y2 as compared to Y1. An independent samples t-test revealed no statistically significant difference between Y2 (50) and Y1 (59) in the number of fixations on the target words (t(16)=0.92, p>.05). Similarly, there was no group effect in the length of fixations (t(16)=1.43, p>.05) between Y2 (17 seconds for all target words) and Y1 (23 seconds). Hypothesis 4 is thus rejected since no difference was discovered in the processing of low-frequency items by less and more experienced trainees.

8.   Discussion

H1 was rejected since the fixation count data for the warm-up task showed no difference between the groups. In fact, descriptive statistics revealed a non-significant tendency contrary to the first hypothesis. Younger students (Y1) seemed to apply a more efficient scanning technique to familiarise themselves with the text than their older colleagues (Y2). This result is difficult to explain by claiming that the scanning technique in sight translation is a skill that develops slowly and cannot be empirically proven by comparing groups of participants with a year-long difference in exposure to conference interpreting tasks. Rather, we might suggest some inter-subject differences across groups. Despite the fact that Y1 had less exposure to interpreting tasks (including ST) than Y2, the overall competence level in that group was higher than in Y2, which was only proven after the final examination taken by both groups (Y1 took their conference interpreting final exam a year after the collection of data for the present study and Y2 a few days after the experiment). The average final examination grade for Y1 was 3.5 and 3.2 for Y2 (both scores out of 5). Although the difference is not very distinct, it may offer at least some explanation for more efficient processing in the scanning phase by Y1 when compared to Y2.

The total task time for Y2 was not significantly shorter for Y1 and H2 was not corroborated. This may be explained by suggesting that the sight translation skill does not develop robustly and a clear difference cannot be documented by comparing groups with one year of training apart. The results obtained can be usefully juxtaposed with those obtained by Jakobsen and Jensen (2009) for ST performed by translation students and professionals whose total task time was significantly shorter than that of the students (for details see section 4). This could be used to corroborate our interpretation of the results that a one-year difference in the

training is too short a period to notice any significant differences in ST performance. Alternatively, it might be so that eye-tracking and/or the measures used in the study are not the best method to examine this kind of expected progression in the acquisition of skills. In that case, more traditional product-based studies of interpreting quality might turn out more suitable.

After redefining H3 and replacing syntax for readability index as a measure predicting processing load, H3 was partially corroborated, that is sentence effect was found for fixation length but not for fixation count. Fixation duration (both first pass duration and total reading time) is assumed to be indicative of syntactic processing (Demberg and Keller 2008). It seems that the SVO/non-SVO word order and simple/complex structure are not the only factors influencing the difficulty of sentence processing in sight translation. As revealed in the study, sentence readability (as reflected in sentence length and word length) could be a better predictor of processing load. Our findings, however, seem to be at variance with those obtained, for example, by Shreve et al. (2010) who found that in sight translation complex syntax required more processing effort than non-complex syntax (see section 4). It may be the case, though, that the difference in our findings results from differences in the study design, as in the case of our sentences there could be other factors influencing the processing effort (such as the lexis, which was rather constant in Shreve et al.'s study). Nonetheless, we believe that the differences in the two sets of results call for further studies involving both sentence complexity and readability as variables.

The fourth hypothesis pertained to the processing of low-frequency items. The lack of group effect in three types of data (fixation length and fixation count) shows that the difference between Y1 and Y2 students in the processing of difficult lexical items is not significant. The reason may be that those particular target words identified in the experimental text may have been too difficult for both groups and led to a bottom effect in the collected data. This suggests that further studies are necessary with a frequency of word usage as a controlled variable with at least three levels: high frequency, regular frequency, low frequency to compare how the two groups of students perform with such targets and if statistically significant differences can be revealed. Another explanation for the lack of group effect may be, as in the case of the second hypothesis, that the sight translation skill does not develop sufficiently enough between the first and the second year of the conference interpreter training programme to be visible in such a study. Hence, a study involving professional interpreters as compared to novices and interpreting school graduates would be an elegant continuation of this line of research.

9.   Conclusions and further research

The present study of eye tracking in sight translation constitutes a new trek in Translation Studies, one that by all means should be pursued by tapping into advanced technologies and research methods. This pioneering pilot study involving two groups of trainees at two stages of training produced some interesting results and contributions to research methodology, but was not successful in corroborating most of our hypotheses.

We generally found no group effect and it is thus impossible to tackle the question about ST skill development in the course of training in an exhaustive way. It may seem that a year difference in training is not a sufficiently long period to reveal noticeable differences in sight translation, efficiency of reading patterns, features that present the greatest challenge to the trainees and the moment they are mastered. Thus, studies involving more distinct groups of participants (novices, graduates and experienced professionals) should shed more light on the issue of ST skill development. Furthermore, longitudinal studies (involving the same group of participants tested at various intervals, such as novices and after a few years of professional conference interpreting experience) should bring valuable data since the participants' experience would be treated as a within-subject variable, the groups would be perfectly matched and no problems with inter-group differences, as in the present case, could skew the data.

This study revealed that researchers should be cautious when selecting their participant groups. The fact that one group (Y1 in our case) had much less exposure to training than another (Y2) does not mean that the former would perform less efficiently than the latter. Final examination results should be somehow considered in the data analysis process. However, this obviously hampers such research projects because researchers have to wait for their beginners to become graduates to collect examination data.

Another research trek to be pursued in this young research field is triangulation of eye tracking data with product research. In fact, the analysis of recordings of our participants' translations is a further stage of the above presented research study. By combining statistically analysed eye tracking data with quality assessment of the product (for instance by means of propositional accuracy score or error analysis (cf. Bartłomiejczyk 2010, Kurz and Färber 2003) and potentially with retrospective interviews, more could be revealed as to the processing of text in ST and the development of ST competence.

Finally, an interesting outcome of this experiment is identification of readability as potentially more important a factor in processing than a simple vs. complex sentence structure distinction. More research is needed in this area, especially in language-specific contexts. Additionally, this finding is directly applicable to training. When selecting texts for students to practice their ST skills teachers should use the readability index as a criterion for determining text difficulty level.

All in all, we believe that eye tracking is yet another interdisciplinary track in Translation Studies that allows researchers to go on new and exciting treks into the black box of translators and interpreters.

References

Agrifoglio, M. 2004. "Sight translation and interpreting: a comparative analysis of constraints and failures." *Interpreting* 6 (1): 43-67.

Alves, F., A. S. Pagano and da Silva, I. L. 2010, 18 Sep. Investing processing effort in the unfolding of the translation process: Methodological issues in empirical-experimental research using eye-tracking data. Paper presented at the 5th International Maastricht-Łódź Duo Colloquium on Translation and Meaning, Łódź, Poland.

Bartłomiejczyk, M. 2010. "Effects of short intensive practice on interpreter trainees' performance." In *Why Translation Studies Matters,* D. Gile, G. Hansen and N. K. Pokorn (eds.), 183-194. Amsterdam/Philadelphia: John Benjamins.

Chmiel, A. 2006. "A skill-based approach to conference interpreting." *Translation Ireland: New Vistas in Translator and Interpreter Training* 17: 47-64.

Craik, F. I. M., and Lockhart, R.S. 1972. "Levels of processing: A framework for memory research." *Journal of Verbal Learning and Verbal Behavior* 11: 671-684.

Demberg, V. and Keller, F. 2008. "Data from eye-tracking corpora as evidence for theories of syntactic processing complexity." *Cognition* 109 (2): 193-210.

Doherty, S. and O'Brien, S. [no date]. Can MT Output be Evaluated Through Eye Tracking? Available at: http://www.mt-archive.info/MTS-2009-Doherty.pdf (last accessed 26 January 2011).

Gerganov, A. 2007. Appendix A: Eye tracking studies with Tobii 1750 – recommended settings and tests. In Appendix A to D1.2 of the Consortium EYE-to-IT. Available at: http://cogs.nbu.bg/eye-to-it/?deliverables (last accessed 26 January 2011).

Jakobsen, A. L. and Jensen, K. 2009. "Eye Movement Behaviour Across Four Different Types of Reading Task." In *Looking at eyes - Eye Tracking Studies of Reading and Translation Processing*, S. Göpferich, A. L. Jakobsen, and Inger Mees (eds), 103-124. Copenhagen: Samfundslitteratur.

Kurz, I. and B. Färber, B. 2003. "Anticipation in German-English simultaneous interpreting." *Forum* 1 (2): 123-150.

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

Moser-Mercer, B., Lambert, S. Daro, V. and Williams, S. 1997. "Skill components in simultaneous interpreting." In: *Conference interpreting: Current trends in research,* Y. Gambier, D. Gile and Ch. Taylor (eds), 133-148. Amsterdam/Philadelphia: John Benjamins.

Pavlović, N. and K. Jensen. 2009. Eye tracking translation directionality. Available at: http://isg.urv.es/publicity/isg/publications/trp_2_2009/chapters/jensenpavlovic.pdf (last accessed 26 January 2011).

Pöchhacker, F. 2004. *Introducing Interpreting Studies*. New York: Routledge.

Rayner, K. 1998. "Eye movements in reading and information processing: 20 years of research." *Psychological Bulletin* 124: 372-422.

Sharmin, S., Špakov, O., Räihä, K.-J. and Jakobsen, A. L. 2008. "Effects of time pressure and text complexity on translators' fixations." In *ETRA '08 Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, 123-126.

Shreve, G. M., Lacruz, I. and Angelone, E. 2010. "Cognitive effort, syntactic disruption, and visual interference in a sight translation task." In *Translation and Cognition*, G. M. Shreve and E. Angelone (eds), 63–84. Amsterdam/Philadelphia: John Benjamins.

Viezzi, M. 1989. "Sight translation, simultaneous interpretation and information retention." In *Aspects of Applied and Experimental Research on Conference Interpretation*, L. Gran and C. Taylor (eds), 54-60. Udine: Campanotto.