Eye tracking translation directionality

NATAŠA PAVLOVIĆ

University of Zagreb, Croatia

KRISTIAN T. H. JENSEN

Copenhagen Business School, Denmark

The paper reports on a study investigating directionality in translation processes by means of eye tracking. The following hypotheses are tested: (1) in both directions of translation, processing the TT requires more cognitive effort than processing the ST; (2) L2 translation tasks on the whole require more cognitive effort than L1 tasks; (3) cognitive effort invested in the processing of the ST is higher in L1 translation than in L2 translation; (4) cognitive effort invested in the processing of the TT is higher in L2 translation than in L1 translation; and (5) in both directions, students invest more cognitive effort in translation tasks than do professionals. The hypotheses are tested through a series of experiments involving student and professional subjects who translate two comparable texts, one into their L1 (Danish) and the other into their L2 (English). The following data from the translation tasks are analyzed: gaze time, average fixation duration, total task length and pupil dilation, all of which are assumed to be indicative of cognitive effort. Only the first hypothesis is found to be wholly confirmed by our data; the remaining hypotheses are only partially confirmed, that is, confirmed by some indicators and not by others, or confirmed for only one group of subjects.

Key words: directionality, translation processes, eye tracking, cognitive effort, gaze time, average fixation duration, pupil dilation, pupillometry.

Introduction

Three areas of research converge in this study: research on translation processes, eye-movement research, and research on translation directionality. Research on translation processes has been conducted for more than 20 years, focusing on various issues and using a variety of research methodologies (two key volumes dealing with methodological issues are Alves 2003, and Tirkkonen-Condit and Jääskeläinen 2000; for a good overview, see Jääskeläinen 2002). Most recently, scholars have started to use eye tracking as a methodology for research on translation processes, including O'Brien 2006 and Jakobsen et al. 2007, applying insights from eye-movement research to study translation. At the same time, Translation Studies has

From *Translation Research Projects 2*, eds. Anthony Pym and Alexander Perekrestenko,
Tarragona: Intercultural Studies Group, 2009. pp. 93-109. ISBN: 978-84-613-1620-5.
http://isg.urv.es/publicity/isg/publications/trp_2_2009/index.htm

broadened its scope to become less prescriptive, less Eurocentric in its approach. Some practices that Western translation theorists had traditionally considered to be simply "wrong" have recently become hot topics of research. One of the issues researchers have thus started focusing on has been the issue of directionality—whether translation is done into the translator's first language (L1 translation) or from that first language into the second (L2 translation). This issue is becoming increasingly important in the globalizing world, as professional translators are increasingly called upon to do L2 translation, particularly but not exclusively in those settings that use a "language of limited diffusion". Directionality has thus been the topic of two forums and their subsequent proceedings (Grosman et al. 2000; Kelly et al. 2003). Attempts have also been made to isolate the differences between the two directions of translation with L2 translation training purposes in mind (e.g. Pavlović 2007). This study continues along the same lines, using eye tracking to investigate the differences between L1 and L2 translation processes of students and professionals. The aim of the study is therefore to see what insights eye tracking has to offer to our knowledge of translation processes with particular regard to translation directionality.

Assumptions and hypotheses

We are assuming that the observable, measurable data that can be gained from eye tracking are indicators of unobservable cognitive processes happening in the subjects' mind during the translation tasks. In this assumption we rely on previous research on eye movements, a good overview of which is Rayner 1998. We are furthermore assuming that the data related to the subjects' focus on the source text (ST) section of the screen are indicators of ST processing (reading, comprehension), while those data related to the subjects' focus on the target text (TT) section of the screen are related to TT processing (production, revision).

We thus used four kinds of data obtainable from eye tracking in order to gain insights into the cognitive processes of our subjects. The following data were used:

a) "gaze time", that is, the total time a subject spent focusing on a particular (ST or TT) section of the screen;
b) "average fixation duration", which is based on the gaze time value and the total number of fixations;
c) "total task length", that is, the total time it took the subjects to complete the given translation task;
d) "pupil dilation", dilation of the subjects' pupils during the task.

All of the above are assumed to be indicators of the subjects' cognitive effort in the given translation task.

With these assumptions in mind, we formulated the following hypotheses:

1. In both directions of translation, processing the TT requires more cognitive effort than processing the ST;
2. L2 translation tasks on the whole require more cognitive effort than L1 tasks;
3. Cognitive effort invested in the processing of the ST is higher in L1 translation (where the ST is an L2 text) than in L2 translation (where the ST is an L1 text);
4. Cognitive effort invested in the processing of the TT is higher in L2 translation (where the TT is an L2 text) than in L1 translation (where the TT is an L1 text);
5. In both directions of translation, students have to invest more cognitive effort in translation tasks than do professionals.

Research design and methodology

In order to test the above hypotheses, we created the following research design. The central part of the research was a series of experiments in which the same subjects were asked to translate two texts, one into their L1 (Danish) and one into their L2 (English). Both source texts were accompanied by a realistic task description (brief). The subjects' gaze behavior was recorded by an eye tracker, and their translation processes recorded by Translog (see below for details of both methodologies). The order of the tasks was reversed for different subjects in order to counter the possibility of "retest" or "acclimatization" effect influencing the data. The subjects were additionally given short warm-up tasks prior to the two main tasks to help them get used to the experimental setting, the computer, the eye tracker, and so on. The two tasks took place on the same day, after a short break.

Source texts

One of the main challenges of this research was to find two source texts that could be considered comparable. Finding comparable texts is a tall order even when they are written in the same language. For the purposes of this study, the two source texts obviously had to be in two different languages, Danish and English, which made comparability even more difficult to test.

Having the subjects direct their gaze at places other than the ST or TT would have made data analysis too complicated. For this reason, the texts we used in the experiments could not be so difficult as to require the use of external resources. The texts we selected were thus non-domain specific (non-technical), and they both belonged to the same genre: they were two reviews from reputable newspapers, of books dealing with a political topic.

The review of Olav Hergel's *Flygtningen*, written by Lars Bonnevie, appeared in *Weekendavisen* on March 17, 2006. The review of *A Russian Diary* by Anna Politkovskaya, written by Thomas de Waal, was published in *The Sunday Times* on April 1, 2007. The articles also appeared in the online versions of the newspapers. We shortened both reviews to around 250 words, and made some minor changes to make them more comparable.

In addition to length and genre, texts can be compared in terms of readability. There are many methods and formulae for measuring readability (the relative ease with which a text can be read) among them, the Kincaid formula, the Flesch reading ease formula, the Fog index, and so on. Problems arise, however, when these tools (which are freely available on the Internet) are used to compare texts written in different languages, as is the case in studies involving directionality of translation. To what extent are the grades obtained by the various formulas comparable across languages? Björnsson (1983), the author of the Lix formula (see Bedre Word 2007), compared readability of newspapers in 11 languages, and the results indicate that for texts of the same genre from comparable newspapers the scores varied widely from language to language. Luckily for the authors of this study, English and Danish were found to get very similar scores, so that the Lix formula could be applied to both our source texts. The formula measures word length and sentence length to arrive at a difficulty assessment ranging from (below) 25 to (over) 54. According to this formula, our two texts belong to the same readability category; namely *medium* level of difficulty. The score for the Danish text was 40 and for the English text 41.

We additionally tested our source texts by means of SMOG, a formula developed by McLaughlin (1969, 2007; see also Trottier 2007), which uses syllable count and sentence length to measure difficulty. According to this formula, our texts were again rated the same degree of readability (12). According to the SMOG scale, full comprehension of the two texts presupposes that the reader has at least 12 years of schooling.

Test subjects

A total of 16 subjects participated in the study. Of these 16 subjects, eight were final year students of translation and eight were professional translators. All subjects had Danish as their L1 and translated primarily into English as their L2.

The subjects' L1 and L2 competences were tested by means of Dialang (www.dialang.org), a language-assessment application based on the Council of Europe's (2001) *Common European Framework of Reference*. Data on the subjects' experience in translation was also elicited.

The presence of *Brownian motion* (see below) in our experiment contaminated the data to such an extent that 50 percent of our data had to be discarded. This left us with four final-year students of translation and four professional translators, which arguably still is a sufficient pool of data for statistical analysis.

Eye-tracking equipment

The tracking of our test subjects' eyes was carried out with the Tobii 1750 eye tracker (www.tobii.se), which is a remote tracker that allows unrestrained head movement. For our type of translation-oriented experiment, unrestrained head movement was deemed essential because we wanted to imitate a translation situation that resembles a translator's normal work environment as much as possible. By using this type of eye tracker instead of a tracker that relies on supporting the test subject's head and thereby obstructing head movement, we achieve a relatively higher level of ecological validity. The main disadvantage of using a remote eye tracker is, however, that the level of eye tracking quality in terms of spatial accuracy is lower (up to 1 degree of inaccuracy) than that of a head supported tracker, e.g. the EyeLink tracker (www.sr-research.com), which has an inaccuracy of between 0.15 and 0.5 degrees. However, despite the reduced spatial resolution, ecological validity was considered more important than accuracy, and a remote eye tracker, such as the 1750, is thus the most suitable type of tracker on the market for our type of naturalistic study.

Eye tracking data analysis and settings

Research shows that the mean fixation duration during silent reading is around 225 milliseconds (Rayner 1998:373). At the same time, Rayner notes that there is considerable variability between readers, which means that fixations can last anywhere from under 100 milliseconds to over 500 milliseconds during silent reading (1998: 376). Therefore, to include a maximum of gaze data directly related to the translation task, the lower fixation threshold that we used to discriminate fixation from non-fixation was set to a temporal resolution of 100 milliseconds and a spatial resolution of 40 pixels. This means that what we consider to be fixations representing reading must consist of a sequence of at least five gaze samples[1] that are located within a radius of 40 pixels from each other.

Having located our fixation threshold, ClearView, which is Tobii's data analysis software, can now analyze the raw tracking data recorded by the eye tracker. ClearView allows the experimenter to extract basic numerical values from the eye-tracking session, among those the total number of fixations during a translation and the total amount of time spent gazing at predefined

---

[1] Cf. Tobii 1750's 50 Hz sampling rate, which equals a gaze sample recorded every 20 milliseconds (i.e. 50 gaze samples each second).

spatial areas of the screen (areas of interest or AOI). For our study of directionality in translation, it was relevant to define two AOIs: an ST AOI and a TT AOI. The expanse of these two AOIs was based on the principle that the potential gaze area which could be directly related to either the ST or the TT should be included into the respective AOIs, so naturally, the ST AOI would include the ST area of the screen and the TT AOI would include the TT AOI of the screen leaving the remaining parts of the screen unassigned.

With ClearView we are able to calculate three of the four values that we use as indicators of cognitive effort, namely: (a) total gaze time, (b) average fixation duration, and finally (c) total task length. We now have three sets of data for our two AOIs.

ClearView does not contain a tool that analyses (d) pupillometric values, i.e. pupil dilation. These values had to be extracted from ClearView's exported data-log files by manually identifying where in the log file the relevant task starts and ends, as suggested by O'Brien (2006: 191).

Indicators of cognitive effort

The four values we use as indicators of cognitive effort are described below. Three of these indicators (a, b, d) are directly related to the test subjects' gaze and pupil behavior while one (c) is related to the overall time it took to complete the task.

(a) Total gaze time

Total gaze time is the combined duration of fixations alone. This means that saccades and the amount of time spent looking away from the screen do not serve as basis for calculating this measurement. Relative distribution of gaze time at ST and TT may be considered an indicator of the distribution of attention and thus an indicator of cognitive effort.

(b) Average fixation duration

The average fixation duration indicator, which is based on total gaze time and the absolute number of fixations, is an indicator of cognitive effort in that an increase in average fixation duration is considered synonymous with increased cognitive effort.

(c) Total task length

The total amount of time it takes to complete a translation task is considered synonymous with increased cognitive effort in that we equate processing time with cognitive effort.

(d) Pupil dilation

Finally, relative change in pupil dilation is considered an indicator of change in cognitive effort. Based on research by Iqbal, Adamcyzk, Zheng and Bailey, O'Brien assumes that the higher percentage change in pupil dilation, the more cognitive effort is expended in the processing of a TM match (2006: 191). For the purposes of this study, we have adopted and modified this assumption, so that we assume that higher percentage change in pupil dilation is synonymous with more cognitive effort being invested into a given translation task.

Problems with data analysis

Fifty percent of the data that we collected with the eye tracker had to be discarded. This 50 percent contained a high level of *Brownian motion*, which is eye-tracking gaze data that are rich in noise and artifacts. This noise may be detected in ClearView's dynamic playback of the eye-tracking session and is characterized by many abnormally short fixations (<200 milliseconds) and erratic vertical saccadic-like motions linking the fixations. This behavior is misrepresentative of *true* gaze data, which consists of primarily horizontal gaze paths (in linear reading) and average fixation durations of at least 200 milliseconds. The source of Brownian motion in our experiments is unknown; however, data from this study and a comparable study by Jakobsen et al. (2007) using some of the same test subjects suggest that Brownian motion most likely is not subject-dependent but rather equipment-dependent. In the Jakobsen et al. study, one test subject exhibited distinct Brownian motion while no Brownian motion could be detected in her recording from this study.

Other data protocols

The recording of the test subjects' keyboard activity was done by the process monitoring software application Translog (www.translog.dk). Translog logs all keyboard and mouse activity which can then be analyzed offline alone or in parallel with other protocols such as TAPs, eye tracking protocols, EEG protocols. With the purpose of the present paper in mind, however, Translog data will not be subjected to analysis.

Findings

Hypothesis 1

Our first hypothesis, that the processing of the TT requires more cognitive effort than the processing of the ST in both directions of translation, was confirmed by all three relevant indicators. In L1 and L2 tasks alike, the subjects spent considerably more time (81.2 percent more in L1 translation and 118 percent more in L2 translation) gazing at the target AOI than they did at the source AOI (Table 1). Their average fixation duration values were higher by 53.1 and 55.1 percent respectively (Table 2). The pupil dilation values were also higher for the target AOI in both tasks (2.4 and 2.6 percent higher respectively; see Table 3).

	L1	**L2**
ST	212798	173790
TT	385497	378840

Table 1. Gaze time (mean values)

	L1	**L2**
ST	258	247
TT	395	383

Table 2. Average fixation duration (mean values)

	L1	**L2**
ST	3.37	3.42
TT	3.45	3.51

Table 3. Pupil dilation (mean values)

While Tables 1-3 compare the mean values, Table 4 shows individual data for all eight subjects and the results of a statistical analysis (paired t-tests). As we can see from Table 4, all p-values are well below 0.05, which means that our first hypothesis was confirmed in a statistically significant way.

L1 task	**gaze time**		**av. fixation duration**		**pupil dilation**	
	ST	**TT**	**ST**	**TT**	**ST**	**TT**
subject 1	547686	620968	331	438	3.1	3.15
subject 2	165545	389075	209	390	4.12	4.22
subject 3	159318	273951	219	305	3.4	3.47
subject 4	176778	309164	265	567	3.51	3.59
subject 5	228392	442833	229	304	2.74	2.81
subject 6	153394	217014	230	235	3.38	3.42
subject 7	134235	422356	259	548	3.44	3.53
subject 8	137033	408614	245	513	3.31	3.41
p-value (< 0.05?)	0.000859077		0.004387401		0.000027503	

L2 task	**gaze time**		**av. fixation duration**		**pupil dilation**	
	ST	**TT**	**ST**	**TT**	**ST**	**TT**
subject 1	354239	504581	303	404	3.16	3.18
subject 2	167989	385183	193	298	4.17	4.3
subject 3	162843	246993	213	295	3.6	3.65
subject 4	175172	455901	253	542	3.61	3.64
subject 5	120542	266096	252	302	2.73	2.89
subject 6	163913	297685	242	247	3.34	3.44
subject 7	140221	346209	251	566	3.5	3.6
subject 8	105400	528070	250	534	3.28	3.42
p-value (< 0.05?)	0.000962781		0.009184741		0.00171975	

Table 4. t-tests for Hypothesis 1

Hypothesis 2

We further hypothesized that L2 translation tasks on the whole require more cognitive effort than L1 tasks. This hypothesis was only partially confirmed, by two of the four indicators of cognitive effort (task length and pupil dilation), of which only the pupil dilation data showed statistical significance in favor of L2 tasks (see Table 9). For both students and professionals, L2 tasks on average lasted longer than L1 tasks (0.9 percent more for students and 2.8 percent more for professionals; see Table 7). Pupil dilation values were also higher in the L2 tasks, for both groups of subjects (2.5 percent higher for students and 0.6 percent higher for professionals; see Table 8).

As far as the average fixation duration is concerned (Table 6), the hypothesis was confirmed for professional subjects, whose protocols showed 5.6 percent higher values in their L2 tasks. The protocols of students, however, showed 7.9 percent higher values in the *opposite* direction. As a result, the total values were in fact slightly (in a statistically insignificant way) in favor of L1 translation. Also surprisingly, gaze time values were higher in L1 translation for both groups of subjects (7.7 percent higher for students and 8.9 percent higher for professionals; see Table 5).

	Students	**Professionals**	**All**
L1	660621	535968	598295
L2	613225	492034	552630

Table 5. Gaze time (mean values)

	Students	**Professionals**	**All**
L1	343	320	333
L2	318	338	327

Table 6. Average fixation duration (mean values)

	Students	**Professionals**	**All**
L1	959517	819618	889568
L2	968505	842686	905595

Table 7. Task length (mean values)

	Students	**Professionals**	**All**
L1	3.57	3.25	3.41
L2	3.66	3.27	3.47

Table 8. Pupil dilation (mean values)

Hypothesis 3

Thirdly we hypothesized that in L1 translation the processing of the ST is more demanding in terms of cognitive effort than it is in L2 translation. The reasoning behind this hypothesis is that in the former task, the ST is a text in the subjects' second language, which should be more difficult to process than the L1 ST from the latter task.

ST + TT	gaze time		av. Fixation duration		task length		pupil dilation	
	L1	L2	L1	L2	L1	L2	L1	L2
Stud. 1	1168654	858820	384	354	1496403	1198588	3.12	3.17
Stud. 2	554620	553172	300	246	807794	855309	4.17	4.23
Stud. 3	433269	409836	262	254	915293	978915	3.43	3.62
Stud. 4	485942	631073	416	398	618578	841207	3.55	3.62
Prof. 1	671225	386638	266	277	1067437	907888	2.77	2.81
Prof. 2	370408	461598	232	245	869125	1058486	3.4	3.39
Prof. 3	556591	486430	403	408	671995	594705	3.48	3.55
Prof. 4	545647	633470	379	392	669916	809664	3.36	3.35
p-value (< 0.05?)	0.472802096		0.354158407		0.809423976		0.035329603	

Table 9. t-tests for Hypothesis 2

Surprisingly enough, this hypothesis was not uniformly confirmed either. Only one of the relevant indicators, gaze time, yielded expected values (22 percent and 23.2 percent higher gaze time values in L1 tasks for student and professional subjects respectively; see Table 10). However, when the t-tests were done on the data, the difference in favor of L1 translation was not found statistically significant (see Table 13).

	Students	Professionals	All
L1	262332	163264	212798
L2	215061	132519	173790

Table 10. Gaze time (mean values)

Average fixation duration values were expectedly higher in L1 translation when it came to the student group (11 percent), but not in the case of professionals. For the latter group, the values were in fact 4.2 percent higher in *L2* translation (see Table 11). The total score was slightly in favor of L1 translation, but not in a statistically significant way (see Table 13).

	Students	Professionals	All
L1	273	238	258
L2	246	248	247

Table 11. Average fixation duration (mean values)

As far as the pupil dilation values are concerned, the professional group showed the expected results, albeit barely so (0.3 percent difference in favor of L1 translation), while the student data in fact suggest a 2.8 percent greater cognitive effort in L2 translation (see Table 12). The overall difference in favor of L2 translation is statistically insignificant (see Table 13).

	Students	Professionals	All
L1	3.53	3.22	3.37
L2	3.63	3.21	3.42

Table 12. Pupil dilation (mean values)

ONLY ST	**gaze time**		**av. fixation duration**		**pupil dilation**	
	L1	**L2**	**L1**	**L2**	**L1**	**L2**
Stud. 1	547686	354239	331	303	3.10	3.16
Stud. 2	165545	167989	209	193	4.12	4.17
Stud. 3	159318	162843	219	213	3.40	3.60
Stud. 4	485942	175172	265	253	3.51	3.61
Prof. 1	228392	120542	229	252	2.74	2.73
Prof. 2	153394	163913	230	242	3.38	3.34
Prof. 3	134235	140221	259	251	3.44	3.50
Prof. 4	137033	105400	245	250	3.31	3.28
p-value (< 0.05?)	0.106607715		0.558054882		0.120428573	

Table 13. t-tests for Hypothesis 3

Hypothesis 4

Our fourth hypothesis stated that L2 TT production requires more cognitive effort than L1 TT production. Only one of the three relevant indicators confirmed this claim, namely the pupillometric indicator (Table 16), which showed a 1.7 percent higher value for average pupil dilation in the L2 translation task compared to the L1 translation task. When a t-test was done on the pupil dilation data, the difference in favor of L2 translation was found to be statistically significant (Table 17).

However, the remaining two indicators showed the opposite: both gaze time and average fixation duration values were on average *lower* in L2 translation, in spite of the mean values for the professional group being slightly higher (Tables 14 and 15). Neither difference in favor of L1 translation was found to be statistically significant (Table 17).

As we can see, these findings do not provide consistent evidence that TT processing requires more effort in L2 translation than in L1 translation.

	Students	Professionals	all
L1	398290	372704	385497
L2	398165	359515	378840

Table 14. Gaze time (mean values)

	Students	Professionals	all
L1	413	378	395
L2	377	390	383

Table 15. Average fixation duration (mean values)

	Students	Professionals	all
L1	3.61	3.29	3.45
L2	3.69	3.34	3.51

Table 16. Pupil dilation (mean values)

ONLY TT	**gaze time**		**av. fixation duration**		**pupil dilation**	
	L1	**L2**	**L1**	**L2**	**L1**	**L2**
Stud. 1	620968	504581	438	404	3.15	3.18
Stud. 2	389075	385183	390	298	4.22	4.3
Stud. 3	273951	246993	305	295	3.47	3.65
Stud. 4	309164	455901	567	542	3.59	3.64
Prof. 1	442833	266096	304	302	2.81	2.89
Prof. 2	217014	297685	235	247	3.42	3.44
Prof. 3	422356	346209	548	566	3.53	3.6
Prof. 4	408614	528070	513	534	3.41	3.42
p-value (< 0.05?)	0.875037204		0.3224893		0.011099053	

Table 17. t-tests for Hypothesis 4

Hypothesis 5

Our fifth hypothesis stated that students of translation need to invest more cognitive effort in a translation task of either direction compared to professional translators. The rationale for this hypothesis is that students have not developed strategies and skills that will effectively help in reducing the amount of time and effort needed to complete the translation task. Our study confirmed the hypothesis for three of the four indicators, as illustrated in Tables 18-20. Students gazed at the AOIs 23.3 percent and 24.6 percent more (in L1 and L2 translation respectively) than did professionals. The task length indicator also confirms our hypothesis: students spent 17.1 percent more time translating the L1 text and 14.9 percent more time translating the L2 text compared to professional translators. Similarly, the pupillometric data suggest that students invest more cognitive effort in the translation tasks, as their pupils were 9.8 percent more dilated in L1 translation and 11.9 percent in L2 translation compared to the professionals' data. The difference in the values for the last indicator was not statistically significant (Table 22).

	L1	**L2**
Students	660621	613225
Professionals	535968	492034

Table 18. Gaze time (mean values)

	L1	**L2**
Students	959517	968505
Professionals	819618	842686

Table 19. Task length (mean values)

	L1	**L2**
Students	3.57	3.66
Professionals	3.25	3.27

Table 20. Pupil dilation (mean values)

In contrast to these three indicators, the final indicator, average fixation duration, only provides partial confirmation of our hypothesis (see Table 21). In L1 translation, the average fixation duration is 7.2 percent longer in students than in professionals, but in L2 translation it is reversed, and the professionals' average fixation duration is 6.3 percent longer than the students'.

	L1	**L2**
Students	343	318
Professionals	320	338

Table 21. Average fixation duration (mean values)

Total values S vs. P	**gaze time**		**av. fixation duration**		**task length**		**pupil dilation**	
	stud.	**prof.**	**stud.**	**prof.**	**stud.**	**prof.**	**stud.**	**prof.**
L1	2642485	2143871	1362	1280	3838068	3278473	14.13	12.87
L2	2452901	1968136	1252	1322	3874019	3370743	14.65	13.10
p-value (< 0.05?)	0.008964971		0.949844573		0.033701446		0.065469196	

Table 22. T-tests for Hypothesis 5

Conclusions

To summarize our findings, only one of our five hypotheses has been wholly confirmed by the data we have collected in this study. As we can see from Table 23, our first hypothesis, that cognitive effort invested in the processing of the TT is greater than that invested in the procession of the ST in both directions of translation, has been confirmed by all the relevant indicators: gaze time, average fixation duration and pupil dilation.

The second hypothesis, that L2 translation tasks require more cognitive effort than L1 translation tasks, has not been confirmed by all four indicators. L2 tasks did last longer and showed an increase in pupil dilation for both student and professional subjects in comparison with L1 tasks, but the remaining two indicators, gaze time and average fixation duration, failed to confirm this hypothesis. At this stage of our research it is difficult to explain the discrepancies between the various indicators of cognitive effort when it comes to L1 and L2 tasks on the whole. Student data, in particular, are ambiguous in this respect. It seems that for students, who are equally inexperienced in both L1 and L2 translation, both directions of translation might be just as demanding in terms of cognitive effort. This coincides with introspective data reported on in Pavlović (2007: 169), where more students actually found L2 translation (subjectively) easier than L1 translation. It is certainly intriguing to find that L2 translation may not necessarily be "more difficult" than translation into L1, as is widely assumed.

It is also widely assumed that ST processing requires more cognitive effort in L1 translation (where the ST is an L2 text) than in L2 translation and, conversely, that TT processing requires more cognitive effort in L2 translation (where the TT is an L2 text) than in L1 translation. When we tested these two related hypotheses, we again found that our data did not provide conclusive evidence to prove the claims. It seems that ST processing in L2 translation can be just as demanding as in L1 translation. Again, this is a finding that coincides with that reported in Pavlović (2007: 160), where the subjects' concurrent verbalizations in collaborative translation protocols suggested that the construction of ST meaning is as important in L2 translation as it is in L1 translation.

Our final hypothesis, that students require more cognitive effort for the same translation tasks than do professionals, was mostly (but not completely) confirmed. One of our indicators, average fixation duration, in fact displayed higher cognitive effort-related values for professionals.

Indicators: / Hypotheses:	Gaze time	Av. Fixation duration	Task length	Pupil dilation
1. L1& L2: TT > ST	+	+	N/A	+
2. L2 task > L1 task	-	-	+	+
3. ST L1 > ST L2	+	+	N/A	-
4. TT L2 > TT L1	-	-	N/A	+
5. L1&L2: Stu > Pro	+	-	+	+

Table 23. The hypotheses / indicators matrix

Our findings would seem interesting in that they challenge traditional assumptions about L1 and L2 translation. However, it would be premature to draw any definitive conclusions from them, for a number of reasons.

First of all, our pool of data was relatively small (eight valid test subjects in all). With such a small sample, any free variable can cause havoc in the data. The statistical tests in particular might have suffered from this limitation. If we add to that the fact that we used highly sensitive equipment that is still insufficiently tested in translation research, it becomes obvious that much more data are needed before we can make even tentative generalizations. Another cause for concern may be the (in)comparability of the source texts. Other texts (and of course, other language pairs) should be used in future studies to corroborate our findings.

In spite of all the limitations of our conclusions, we believe that the findings from our study are intriguing enough to invite further research on the topic of directionality in translation processes, as well as further research on other translation-related topics that will make use of eye tracking.

References

Alves, Fábio (ed.) 2003. *Triangulating translation. Perspectives in Process Oriented Research*. Amsterdam and Philadelphia: John Benjamins.

Bedre Word. Visited January 2009. *Templates and program enhancements that add new functions to Microsoft Word.* http://bedreword.dk

Björnsson, C. H. 1983. "Readability of newspapers in 11 languages". *Reading Research Quarterly* 18(4): 480-497.

Council of Europe. 2001. *CEFR, Common European Framework of Reference for Languages*. Cambridge: Cambridge University Press. Also available at: http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf, www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf

Grosman, Meta, Mira Kadric, Irena Kovačič and Mary Snell-Hornby (eds.) 2000. *Translation into non-mother tongues. In professional practice and training.* Tübingen: Stauffenburg Verlag.
Jääskeläinen, Riitta. 2002. "Think-aloud protocol studies into translation: An annotated bibliography". *Target* 14(1): 107-136.
Jakobsen, Arnt L., Kristian T.H. Jensen, 2007. *Coordination of comprehension and text production in written and oral translation tasks*. Poster presented at AMLaP conference 27 August 2007.
Kelly, Dorothy, Anne Martin, Marie-Louise Nobs, Dolores Sánchez and Catherine Way (eds.) 2003. *La direccionalidad en traducción e interpretación. Perspectivas teóricas, profesionales y didácticas*. Granada: Editorial Atrio.
McLaughlin, G. Harry. 1969. "SMOG grading—a new readability formula". *Journal of Reading*. http://www.harrymclaughlin.com/SMOG_Readability_Formula_G._Harry_McLaughlin_(1969).pdf . Visited January 2009.
McLaughlin, G. Harry. Visited January 2009. Homepage: http://www.harrymclaughlin.com/SMOG.htm
O'Brien, Sharon. 2006. "Eye-tracking and translation memory matches". *Perspectives: Studies in Translatology* 14(3): 185-203.
Pavlović, Nataša. 2007. *Directionality in Collaborative Translation Processes.* Ph.D. thesis. Tarragona: Universitat Rovira i Virgili. Available on: http://www.tesisenxarxa.net/TDX-1210107-172129/
Rayner, Keith. 1998. "Eye Movements in Reading and Information Processing: 20 Years of Research". *Psychological Bulletin* 124(3): 372-422.
SR Research. Visited January 2009. Manufacturer of eye tracking equipment. http://www.sr-research.com/
Tirkkonen-Condit, Sonja and Riitta Jääskeläinen (eds.) 2000. *Tapping and Mapping the Process of Translation: Outlooks on Empirical Research.* Amsterdam and Philadelphia: John Benjamins.
Tobii. Visited January 2009. Manufacturer of eye tracking equipment. http://www.tobii.se/
Translog. Visited January 2009. Translog website: http://www.translog.dk/
Trottier, Alain. 2007. WordsCount. SMOG Calculator. http://www.wordscount.info/hw/smog.jsp . Visited January 2009.