

## Exploring Translation and Interpreting Hybrids. The Case of Sight Translation

Barbara Dragsted et Inge Gorm Hansen

Volume 54, numéro 3, septembre 2009

URI : <https://id.erudit.org/iderudit/038317ar>

DOI : <https://doi.org/10.7202/038317ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)

1492-1421 (numérique)

[Découvrir la revue](#)

Citer cet article

Dragsted, B. & Hansen, I. G. (2009). Exploring Translation and Interpreting Hybrids. The Case of Sight Translation. *Meta*, 54(3), 588–604.  
<https://doi.org/10.7202/038317ar>

Résumé de l'article

Le présent article fait état d'une étude comparative de la traduction écrite et de la traduction à vue. Elle est fondée sur des données expérimentales qui associent un enregistrement de la frappe, une étude oculométrique ainsi qu'une évaluation de la qualité de traductions orales et écrites produites par des traducteurs et des interprètes professionnels. La comparaison des modalités orale et écrite met en évidence des différences majeures. L'évaluation des traductions montre en effet que le débit faible observé pour la traduction écrite ne garantit nullement une qualité accrue. Par ailleurs, l'analyse comparative fait état de différences fondamentales de comportement entre interprètes et traducteurs. De façon générale, la traduction orale semble pouvoir contribuer de façon significative à l'économie de temps et d'effort sans compromettre la qualité, ce qui justifierait une accentuation de son rôle, et même une pleine intégration, dans la formation des traducteurs.

## Exploring Translation and Interpreting Hybrids. The Case of Sight Translation

**BARBARA DRAGSTED**

*Copenhagen Business School, Copenhagen, Denmark*  
bd.isv@cbs.dk

**INGE GORM HANSEN**

*Copenhagen Business School, Copenhagen, Denmark*  
igh.isv@cbs.dk

### RÉSUMÉ

Le présent article fait état d'une étude comparative de la traduction écrite et de la traduction à vue. Elle est fondée sur des données expérimentales qui associent un enregistrement de la frappe, une étude oculométrique ainsi qu'une évaluation de la qualité de traductions orales et écrites produites par des traducteurs et des interprètes professionnels. La comparaison des modalités orale et écrite met en évidence des différences majeures. L'évaluation des traductions montre en effet que le débit faible observé pour la traduction écrite ne garantit nullement une qualité accrue. Par ailleurs, l'analyse comparative fait état de différences fondamentales de comportement entre interprètes et traducteurs. De façon générale, la traduction orale semble pouvoir contribuer de façon significative à l'économie de temps et d'effort sans compromettre la qualité, ce qui justifierait une accentuation de son rôle, et même une pleine intégration, dans la formation des traducteurs.

### ABSTRACT

This article reports on a comparative study of written translation and sight translation, drawing on experimental data combining keystroke logging, eye-tracking and quality ratings of spoken and written output produced by professional translators and interpreters. Major differences in output rate were observed when comparing oral and written modalities. Evaluation of the translation products showed that the lower output rate in the written condition was not justified by significantly higher quality in the written products. Observations from the combination of data sources point to fundamental behavioural differences between interpreters and translators. Overall, working in the oral modality seems to have a lot to offer in terms of saving time and effort without compromising the output quality, and there seems to be a case for increasing the role of oral translation in translator training, incorporating it as a deliberate practice activity.

### MOTS-CLÉS/KEYWORDS

processus cognitif de traduction, traduction à vue, interprétation, enregistrement de frappe, étude oculométrique  
cognitive translation processes, sight translation, interpreting, keystroke logging, eye tracking

## 1. Introduction

This paper presents the findings of a small-scale comparative study of written translation and sight translation. It explores the notion of *speaking* a translation instead of *writing* it, and draws on experimental translation process data as well as product data. Process data from eight professional translators and interpreters working under varying processing conditions were analysed with keystroke logging and eye-tracking and the quality of the translation output was rated by three experts. The study forms part of a larger project<sup>1</sup> on comprehension and production processes in translation and interpreting hybrids.

The interpreter participants spent approximately 4-5 minutes on the sight translation, whereas some of the translators spent up to an hour on the written translation of the same text. In view of this very considerable difference in production time, it seems relevant to investigate the added value (if any) of the translators' written translations compared with the interpreters' sight translations, and perform intra-subject analyses of the translators to observe the time/quality relationship between their written translation and their sight translations.

Analyses of the target texts will be made on the basis of quality ratings performed by three experts. If the written translations are not (significantly) better than the oral output, or if the oral output is rated higher than the written output, it will be relevant as a next step to discuss whether translators, in addition to saving time and physical effort, may improve their overall performance by adopting some of the techniques and strategies traditionally associated with the oral modality (Riccardi 2002: 26). Thus, we will explore and compare characteristics of the interpreters' oral translation, the translators' oral translation and the translators' written translation.

In addition to the primary aim of investigating and discussing different processes in written and sight translation, a secondary aim of this study is to test and evaluate new methodologies and technologies in translation studies, in particular the use of eye-tracking data combined with keystroke logging.

## 2. Oral and written modalities

It is commonly assumed that oral and written translation are two fundamentally different processes, and that the working methods and strategies applied by interpreters are different from those applied by translators. Sight translation is a hybrid between written translation and interpreting in that the source text (ST) is written and the target text (TT) is spoken (Agrifoglio 2004: 43; Setton and Motta 2007: 203).

Traditionally, sight translation has been treated as a type of simultaneous interpreting, although it differs from both consecutive and simultaneous interpreting in various ways (Agrifoglio 2004: 44). First of all, the source text segment continues to be visually accessible to the translator (Agrifoglio 2004: 44; Gile 1997: 204), which means that there is no memory effort of the kind involved in traditional simultaneous and consecutive interpreting (Gile 1997: 203). Moreover, sight translation is not paced by the source language speaker, meaning that the interpreter has more flexibility in terms of the speed of delivery (Gile 1997: 203). However, it seems that the interpreter will, under normal circumstances, be intent on producing a smooth delivery (Agrifoglio 2004: 45; Gile 1995: 166; Mead 2000: 90; Mead 2002: 74, 82). Finally, as mentioned by Gile, the risk of linguistic interference may be higher in sight translation

than in simultaneous and consecutive interpreting, because the source language words fade more rapidly from memory in interpreting (Gile 1997: 204).

Despite these differences, sight translation appears to be closer to interpreting than to translation because the output is oral and because the oral modality carries an expectation of speedy delivery. In our experiments, the interpreters were able to apply oral-to-oral interpreting strategies, such as reformulation and condensation, whereas translators (with one exception) when asked about differences between writing and speaking their translation, stated that they found the oral translation process unfamiliar and quite different from their usual working method.

Translators, according to Gile (1995: 101-106), process one translation unit after another: They first read and comprehend a source language (SL) unit, developing a meaning hypothesis which they test and reconstruct until they are satisfied. Subsequently they write a target language (TL) rendition of the translation unit, which they then check for fidelity and editorial acceptability, and change if they find it necessary. The degree to which the translation process is as sequential as described here may depend on the individual translator's level of expertise (de Groot 1997: 30-32; Dragsted 2005: 54-67). Yet, there is no doubt that translators process segments fundamentally differently from interpreters. Because of the time constraints present in interpreting, including sight translation, interpreters have to start producing TL output simultaneously with comprehending SL input (Gile 1995: 169-170) and either wait for longer SL units of meaning to be spoken before they can start producing TL output, or apply anticipation strategies.

Below, we will investigate how these differences between translator and interpreter behaviour are reflected in our data by exploring the participants' orientation on the screen (eye movements) and the segmentation of the written and oral target text output (distribution of pauses in speech and in writing) under varying conditions.

Another aspect which will be considered in our data analysis is the notion of proximity, i.e., how close the target text is to the source text. Although most translation and interpreting scholars seem to agree that fidelity is a central criterion in translation/interpreting quality assessment (e.g., Gile 1995: 50), it seems that translators tend to search tirelessly for "the right word" (Schäffner 2004: 7) to a much higher degree than interpreters, who in turn, due again to the time constraints, often have to compromise and apply various strategies such as compression or condensation (Pöchhacker 2004: 134), leading to a less literal translation.

A more form-based strategy in interpreting was found in a study by Shlesinger and Malkiel comparing written translation and simultaneous interpreting. Here it was found that "the cognate seems to be the default option in interpreting," i.e., in the special case of cognates interpreters are more likely to opt for the literal translation, which is easier and faster (Shlesinger and Malkiel 2005: 184).

Differences between oral and written language at a more general level are likely to influence both the level of proximity and the distribution of pauses. Although sight translation is indeed different from spontaneous speech as the translator/interpreter is highly restricted by the requirement to reproduce the content of the source text, we expect the oral (sight translation) output to contain features of the oral mode which are not found in the written translations, because of expectations of a smooth delivery.

Chafe and Danielewicz (1987) stress differences in the variety of vocabulary between spoken and written output by focusing on

[...] how speakers and writers choose words and phrases appropriate to what they want to say. [...] speakers must make such choices very quickly whereas writers have time to deliberate, and even to revise their choices when they are not satisfied. As a result, written language, no matter what its purpose or subject matter, tends to have a more varied vocabulary than spoken. (Chafe and Danielewicz 1987: 86)

In other words, writers and translators, can take, in principle, as long as they want to find an appropriate word, whereas speakers, and interpreters/sight translators, “may typically settle on the first words that occur to them” (Chafe and Danielewicz 1987: 88).

### 3. Research design and methods

All experiments were conducted in the summer of 2006. The results of a pilot study have been reported in Dragsted and Hansen (2007).

#### 3.1. Experimental set-up and procedure

The experiments were conducted in a sound-proof video recording studio. Initially, the participants were given a brief written outline of the project, including an instruction to the experimental set-up and to the tasks they were expected to complete, as well as a brief oral introduction to the eye-tracker and Translog (see below). In order to capture eye movement data, each participant was subjected to a calibration procedure (see O’Brien 2006a).

Gaze data from all participants were collected with a Tobii eye-tracker,<sup>2</sup> a system which registers a person’s eye movements on a computer screen throughout any given process. The recordings from the eye-tracker were analysed with Tobii’s *ClearView* software to track and study the participants’ visual orientation and gaze behaviour. (For a more thorough introduction to the Tobii eye-tracking system and its use in translation research, see O’Brien 2006a, and for an introduction to the use of eye-tracking during reading, see Radach *et al.* 2004; Rayner 1998.)

Eye-tracking is a relatively new technology in translation studies, and we encountered various problems related to our research design. Firstly, our pilot tests revealed that the system’s registration of gaze fixation was sometimes imprecise. This could be off-set by increasing the line spacing and/or the font size. Secondly, synchronization of gaze data with oral output had to be done “manually.” Finally, we had to manually determine the lexical content related to a given gaze fixation on the screen. In the Eye-to-IT project,<sup>3</sup> additional modules have been developed which enable: 1) automatic synchronization of gaze data and audio output (integration of an audio component in *Translog*); 2) automatic identification of words on the basis of gaze fixations (with the Gaze-to-Word Mapping [GWM] module). For a discussion of research design methodology using eye-tracking in translation and reading studies, see Göpferich *et al.* (2008).

The translators’ written translation process was logged in *Translog*, a program which performs keystroke logging and registers pauses in the translation process (Jakobsen and Schou 1999).

Furthermore, each participant's performance was video-recorded, and finally, after conducting the actual experiments, we interviewed the participants about the tasks they had completed.

The experimental set-up was far from identical to the normal working conditions of translators and interpreters, and it must be borne in mind that factors such as video monitoring, keystroke logging and, in the case of the interpreters, the absence of an audience impinge upon the data. This was reflected in comments made by some of the participants in the subsequent interviews.

What is presented here is a small-scale study, and our findings are preliminary. A larger number of participants, more language pairs and a range of different text types would make our results more generalizable, and enhanced technology (GWM and audio components, see above) would strengthen the reliability of our data.

### *3.2. Participants*

All participants (four interpreters and four translators) were native speakers of Danish with at least 10 years' experience as professional interpreters and translators. The interpreters work primarily as interpreters, but two of them are also practising translators. None of the translators take on interpreting jobs although they are all sworn translators and interpreters.

The participants were instructed to perform the tasks as they would have done if they had been commissioned by a customer.

### *3.3. Texts and transcription*

The source text was a political speech given in English in an EU context (1289 words). For the purpose of the present study, the interpreters performed only one task: a sight translation into Danish of the first 634 words.<sup>4</sup> The translators performed two different tasks and were divided into two groups of two; one group was asked to make a sight translation of the first 634 words and a written translation of the next 655 words, and vice versa for the other group.

The data from the eight sight translations (four interpreters and four translators) were transcribed using the 'bysoc' criteria developed for a national corpus of spoken Danish<sup>5</sup> (adapted version). The raw transcriptions were then transformed into texts which resembled written translations in that punctuation was added and off-the-cuff remarks were deleted. However, apart from these changes, the texts were not improved or amended in any way.

### *3.4. Quality assessment*

"Quality is an elusive concept, if ever there was one" (Shlesinger 1997: 123), and quality assessment in translation and interpreting immediately raises the question of quality for whom and from which perspective. Various models, sets of assessment criteria and norms for evaluation of the quality of written translation and interpreting respectively have been developed and discussed extensively (see e.g., Garzone 2002; House 1997; Kalina 2002; Pöchhacker 2002; Schjoldager 1996), but there is no overarching definition of quality applicable to both modalities of translation – see Anderman's

proposal to “set up a project to decide on criteria for the assessment of translation and interpretation” (Schäffner 2004: 46, see also Setton and Motta 2007: 201-202).

Hence, our choice of evaluation criteria and method was not easy in that our data comprise both written translations (by translators) and an artificial/hybrid product, i.e., transcript-based texts produced from oral translations delivered by both translators and interpreters.

In quality assessments, we have taken a product-oriented approach (Pöchhacker 2002: 96). Pöchhacker cites core criteria of interpreting quality such as accuracy, clarity and fidelity and refers to Gile’s notion of the target text as “a ‘faithful’ image” of the original discourse. In written translation the notion of fidelity has also been central in the discussion of quality (e.g., Kalina 2002: 120).

In our attempt to find common ground, we decided to ask three raters to evaluate the eight sight translations and the four written translations. One of the raters is a professor and translation teacher at Copenhagen Business School (CBS), one is an external lecturer at CBS as well as a freelance translator/interpreter, and one is a full-time translator. Information on whether the texts were produced by an interpreter or a translator and whether it was a sight translation or a written translation was not available to the raters. The three raters were asked to assess the quality of the translation/sight translation products as follows:

- a) Give an overall assessment of the adequacy of the target text (as a translation of the source text) (on a scale from 1-5, 1 = highest quality, 5 = lowest quality).
- b) Indicate the proximity of the target text to the source text (on a scale from 1-5, 1 = very close to source text, 5 = very free).
- c) Indicate errors in the text and state the severity of the error on a scale from 1-5 (5 = most serious). Please take into account that the text is a speech (political).<sup>6</sup>
- d) Indicate on a scale from 1-5 the approximate time it will take to edit/revise the text so that it will be an adequate translation of the source text (1 = approx. 0-10 min., 5 = approx. 40-50 min.).

### 3.5. *Pause analysis*

Pause analysis of the written translation output was performed on the basis of pause data from *Translog* supplemented by eye-tracking data. The analyses of the oral translations were based on the video recordings of each participant combined with eye-tracking data. Audio files (wav) of the output texts were extracted from the videos and analysed in the speech analysis and synthesis program PRAAT<sup>7</sup> using an add-on script developed for pause detection.<sup>8</sup>

Our definition of a pause was based on written translation output, where pauses of 1-2 seconds have been shown to indicate some translation task-related cognitive processing (Baddeley 1986; Dragsted 2005; Jakobsen 1998, O’Brien 2006b). For purposes of comparison between written and oral output, we have defined a pause in both written and sight translation as an interruption in the writing or speech process of 2 seconds or more. When analyzing oral output, a distinction is often made between filled and unfilled pauses (Ahrens 2005; Mead 2000), where filled pauses typically consist of hesitation markers, while unfilled pauses are defined as silence intervals. In the present analysis we will focus on unfilled pauses.

## 4. Results and data analysis

### 4.1. Output rate and quality assessment

Table 1 below shows significant differences in processing time between the two groups.

TABLE 1

Speed of production by interpreters and translators in words per minute (WPM)

<i>Interpreters, sight translation</i>		<i>Translators, sight translation</i>		<i>Translators, written translation</i>	
<i>Interpreter</i>	<i>WPM</i>	<i>Translator</i>	<i>WPM</i>	<i>Translator</i>	<i>WPM</i>
<i>I1</i>	147	<i>T1</i>	35	<i>T1</i>	11
<i>I2</i>	133	<i>T2</i>	79	<i>T2</i>	11
<i>I3</i>	160	<i>T3</i>	66	<i>T3</i>	12
<i>I4</i>	127	<i>T4</i>	114	<i>T4</i>	34
<i>Mean (SD)</i>	142 (14.77)	<i>Mean (SD)</i>	74 (32.70)	<i>Mean (SD)</i>	17 (11.34)

The translators produced the written translation at a rate of 11-34 words per minute. Under the sight translation condition the translators worked considerably faster (35-114 WPM) but on average only about half as fast as the interpreters who processed the text under the same task conditions at a pace of 127-160 WPM (see also Jakobsen, Jensen *et al.* 2007: 233). Note that the interpreters constitute a fairly homogeneous group, whereas one translator, T4, stands out by working faster than the other translators both under the written and the sight translation condition. Yet, all translators, including T4, worked at a considerably lower output rate than the interpreters when completing both tasks, in particular when producing the written translation.

In order to measure whether the additional time spent by translators producing written translations seems worthwhile, i.e. whether the quality of their written output is correspondingly higher than that of the interpreters' – and their own – sight translation output, we correlated the time data in Table 1 with the quality assessments given by the raters.

TABLE 2

Quality scores and estimated correction time<sup>9</sup>

<i>Interpreters, sight translation</i>					<i>Translators, sight translation</i>					<i>Translators, written translations</i>				
<i>Quality score</i> <sup>10</sup>														
	<i>R1</i>	<i>R2</i>	<i>R3</i>	<i>Mean</i>		<i>R1</i>	<i>R2</i>	<i>R3</i>	<i>Mean</i>		<i>R1</i>	<i>R2</i>	<i>R3</i>	<i>Mean</i>
<b>I1</b>	2	1	1	<b>1.3</b>	<b>T1</b>	4	4	3	<b>3.7</b>	<b>T1</b>	4	4	4	<b>4</b>
<b>I2</b>	3	3-4	2-3	<b>3</b>	<b>T2</b>	3	2-3	2-3	<b>2.7</b>	<b>T2</b>	3	2	2	<b>2.3</b>
<b>I3</b>	5	4	5	<b>4.7</b>	<b>T3</b>	2	2-3	2	<b>2.2</b>	<b>T3</b>	3	2	1	<b>2</b>
<b>I4</b>	5	4	5	<b>4.7</b>	<b>T4</b>	2	3	1	<b>2</b>	<b>T4</b>	2	2	1	<b>1.7</b>
<i>Time needed to correct the text</i> <sup>11</sup>														
	<i>R1</i>	<i>R2</i>	<i>R3</i>	<i>Mean</i>		<i>R1</i>	<i>R2</i>	<i>R3</i>	<i>Mean</i>		<i>R1</i>	<i>R2</i>	<i>R3</i>	<i>Mean</i>
<b>I1</b>	1	1	1	<b>1</b>	<b>T1</b>	4	3-4	3	<b>3.5</b>	<b>T1</b>	3	4	5	<b>4</b>
<b>I2</b>	2	3	2-3	<b>2.5</b>	<b>T2</b>	2	1-2	2-3	<b>2</b>	<b>T2</b>	2	2	2-3	<b>2.2</b>
<b>I3</b>	3	3	5	<b>3.7</b>	<b>T3</b>	1	1-2	2	<b>1.5</b>	<b>T3</b>	1	1-2	1	<b>1.2</b>
<b>I4</b>	4	3-4	5	<b>4.2</b>	<b>T4</b>	1	2	1	<b>1.3</b>	<b>T4</b>	1	1-2	1	<b>1.2</b>



As appears from the ratings in Table 2, the translators' written translations are not unambiguously better than the interpreters' sight translations or the translators' own sight translations. Thus, the additional time spent under the written translation condition does not seem to be justified by a considerably higher quality of the translation product.

I1, who is the fastest of all the participants, is also the one who gets the best quality score, whereas I3 and I4 are closer to the other end of the 1-5 quality scale. In the written condition, T1 spends almost an hour to complete the written translation task, and gets a very low score.

The picture with respect to the interpreters' quality ratings is uneven in that the group comprises both the highest and the lowest scores. Two of the interpreters produce high quality output at high speed, and I1 in particular stands out as an "elite interpreter" (Ericsson 2000: 211) whose performance is remarkable, but looking at the performance of the group as a whole it cannot unambiguously be maintained that working in the oral modality enhances the quality of the translation product. However, the striking differences in output rate shown in Table 1 suggest that an analysis of differences in strategic behaviour under the oral and written translation modalities may be worthwhile.

An intra-subject comparison of the translators' performance under the sight vs. the written condition shows that T1 performs marginally better in the sight translation than in the written translation even though she sight translates 3 times faster. The other three translators take 3-7 times longer processing the written translation, and receive marginally better scores under this condition. In two cases, T1 and T2, the time consumption to correct the TT and bring it up to an acceptable standard is higher for the written translation than for the sight translation.

Generally, it might be argued that the substantially lower output rate in the translators' written translation compared with their sight translation does not seem to be justified by higher quality output. This lends support to the idea of speaking one's translation and exploring the potential of oral translation to play a more prominent role in translator training curricula by incorporating sight translation as a deliberate practice activity (Ericsson 2000: 212).

#### *4.2. Eye movements in sight vs. written translation*

In this and the following sections we will analyse and discuss differences in strategic behaviour among translators and interpreters. In view of the results on the output rate/quality-relationship reported above, we will discuss whether it may be useful for translators to draw on some of the strategies typically associated with interpreters and the oral modality, and whether it may even be worthwhile to include interpreting strategies and tasks in our curricula for translator training (see Hansen and Shlesinger 2007 on interpreter training at CBS).

The eye-tracking data provided us with insights into the participants' orientation on the screen during the translation process. The figures below show so-called hotspot analyses highlighting focus areas of a selected segment, i.e., the gaze behaviour for a paragraph of the text in the interpreters' sight translations and the translators' written translations. The hotspot image shows a grey background with highlighted areas indicating where the participants have been looking. The gaze intensity is reflected

in the size and the colour of the spots, i.e., the longer the gaze, the darker and larger the spot (the “hotter” the spot).

FIGURE 1

#### Hotspot analysis for interpreters' sight translation (ST)

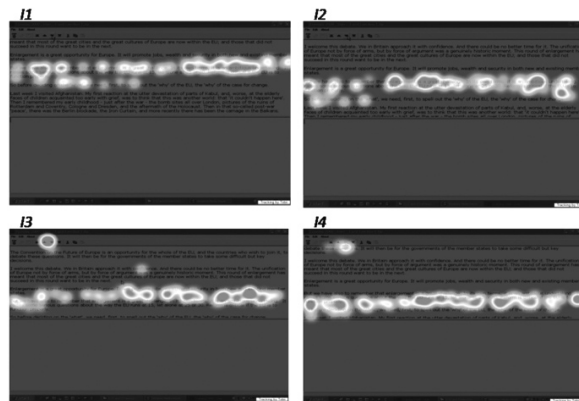
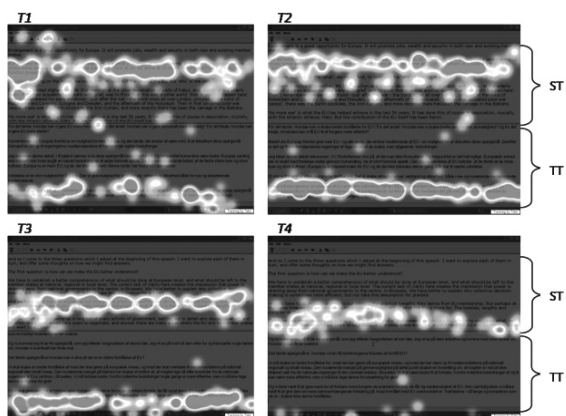


FIGURE 2

#### Hotspot analysis for translators' written translation (ST and TT)

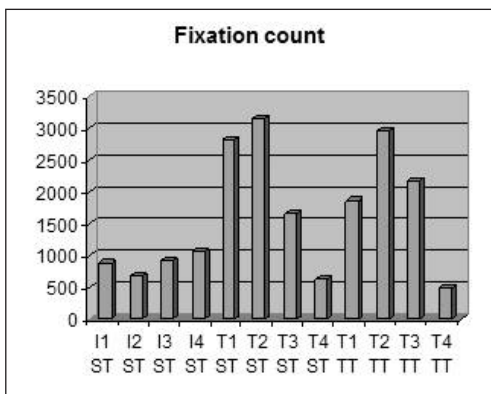


This visual representation of the participants' gaze behaviour indicates that the interpreters' (I1-I4) reading and comprehension proceeds in a fairly “controlled” and linear manner where the gaze stays on the area of the source text that is currently being translated. Hotspot analyses for the translators' sight translation (not displayed here) reveal a pattern similar to the one found in the interpreters' sight translation. The translators' (T1-T4) gaze fixations, by contrast, are split between the source and target text window (ST and TT), but the gaze pattern in the source text window (also when seen in isolation) is characterized by scattered fixations, indicating that the translators are not focussed solely on the segment being translated. Generally, interpreters seem to concentrate more linearly on comprehension and production than translators, who adopt a more global strategy with plenty of regressions and back-tracking (Rayner 1998).

A possible explanation for this behavioural difference may be that the interpreters were affected by time pressure and the concern for smooth delivery, whereas the translators, who felt no time constraint, searched larger areas of the screen in order to look for help in previously translated TT segments or further down in the ST.

Further support for the distinction between the interpreters' focussed and continuous reading vs. the translators' more discontinuous and disrupted reading of the ST was found when calculating the number of fixations in the source text window (see Figure 3). For the interpreters, we looked at source text fixations (I1 ST, I2 ST etc.) and for the translators we divided our analysis into fixations on the source text (T1 ST, T2 ST etc.) and on the target text (T1 TT, T2 TT etc.).

FIGURE 3  
Number of gaze fixations



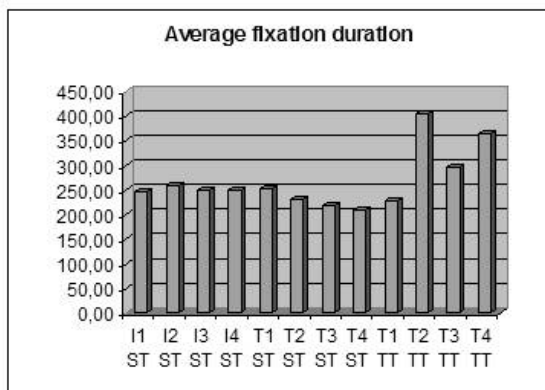
With one exception (T4), the translators had far more ST fixations than the interpreters. This indicates, again, that the interpreters proceed in a more linear manner. In broad terms, once they have completed the translation of a segment, they move on to the next without further backtracking. The translators, on the other hand, return to the same word or segment repeatedly, reconsidering and sometimes rephrasing it. This behavioural difference is in line with Gile (1995) (see section 2).

There seems to be a tendency for the TT fixation count to follow the ST fixation count. For instance, T2, who has the highest number of ST fixations, also has the highest number of TT fixations. At the other end of the scale, we find T4 with the same pattern, but remarkably few fixations in both the ST and the TT window. This explains T4's hotspot analysis (Figure 2), which resembles the interpreters' gaze behaviour (in a selected paragraph) with fixations concentrated almost exclusively in the ST window. The video recordings showed that T4 is a skilled touch typist, and it may be the case that the automation of the typing process frees resources for other efforts, and minimizes the need for ongoing TT editing.<sup>12</sup> When interviewed, T4 stated that she teaches oral communication and is comfortable with working in the oral modality. These competencies might explain why T4 stands out as an "elite" translator (Ericsson 2000: 211) both under the oral and the written condition.

As a final parameter for gaze behaviour, we analyzed the average fixation duration (see Figure 4).

FIGURE 4

## Average fixation duration



The interpreters constitute a very homogeneous group in this respect, with average fixation durations between 240 ms and 260 ms averaging 252 ms. By comparison, Rayner cites mean fixation duration values of 225 ms for silent reading and 275 for oral reading (Rayner 1998: 373). There is a slight tendency for the translators to have shorter average fixation durations on the source text (average 228 ms) than the interpreters; the translators are generally more focussed on the TT output with average fixation durations generally between 300-400 ms, which is close to the average fixation duration of 400 ms found for typing (Rayner 1998: 373).

A comparison of the ST fixation count and fixation duration values reveals that interpreters had few but slightly longer gaze fixations whereas the translators had a higher number of ST fixations of shorter average duration. The eye movement record as a whole (hotspot analysis, fixation count and fixation duration) generally points to the same conclusion, that is to say, the interpreters' translation process is more continuous than that of the translators. The translators' gaze behaviour is disrupted, reflecting that they have to relate to two texts on the screen. Also, since translators are not under immediate time pressure, they have time to scan other segments than the one being translated to retrieve information.

According to Clifton, Staub *et al.* (2007) predictable words "are looked at for less time than words that are not predictable" and readers more frequently skip over high predictable words than low predictable words (Clifton, Staub *et al.* 2007: 347). The interpreters' lower number of gaze fixations and more focussed fixation pattern in general may perhaps be explained by their application of anticipation strategies, which, when successful, enable them to generate expectations guiding the comprehension process (Pöschhacker 2004: 119, 133) and allowing them to quickly scan or even skip words in the text. According to Ericsson, skilled interpreters can be expected to behave like top pianists, who are able to sight-read music, leaving out less important notes but preserving the core melody, and to continue playing the music and successfully predict notes that were blacked out (Ericsson 2000: 213).

Clifton, Staub *et al.* go on to argue that "[...] how long readers look at a word is influenced by the ease or difficulty associated with accessing the meaning of the word" (Clifton, Staub *et al.* 2007: 248). The translators' high number of gaze fixations

and their long fixation duration on ST and TT collectively reflect higher cognitive effort associated with the comprehension and production processes than is apparently exhibited by the interpreters.

Putting more emphasis on the oral modality in translator training might force translators to become more focussed and opt for their first choice (like the interpreters) rather than searching meticulously for the perfect word. In some cases, the fast-and-ready solution may not be preferable, and will have to be corrected afterwards (Schäffner 2004: 1-2), but quite often, the translator's first intuition may also result in the best translation.

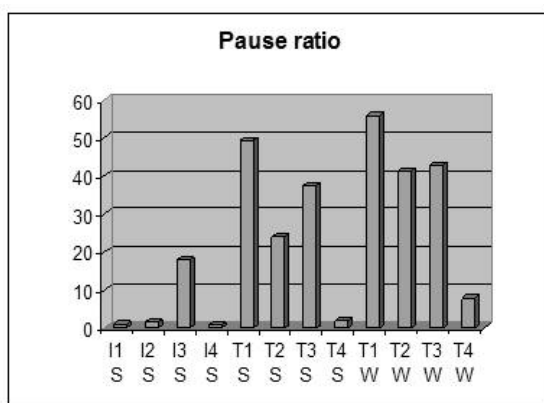
In an attempt to take the gaze analysis one step further, we have experimented with calculating the participants' "eye-voice span" in sight translation (see ear-voice span in simultaneous interpreting). Future perspectives for our research include new experiments to elicit data on the coordination between comprehension and production in oral and written translation.

#### 4.3. Pauses

As shown in Table 1, the interpreters completed their tasks substantially faster than the translators, which is also reflected in the number and duration of pauses registered by *Translog* and PRAAT. In addition to the unfilled pauses detected by PRAAT, the transcriptions of the oral (sight translation) output revealed that, unlike the interpreters, the translators had a large number of filled pauses, which will not be included in our analysis at this point.

FIGURE 5

Proportion of pausing (for 2 seconds or more) in relation to the total production time (per cent)<sup>13</sup>



The remarkable differences in Figure 5 reflect that the translators had considerably more and longer pauses than the interpreters. For example, I1 and I4 had only one pause of more than 2 seconds, whereas T1 and T2 had 213 and 188 pauses respectively. The pause ratios are indicative of fundamental behavioural differences between translators and interpreters. The translators producing written translation appear to

be hesitant, pausing sometimes before each word, whereas the interpreters seem more self-assured producing the text fluently with very few pauses. An intra-subject comparison of the translators indicates less hesitation in the oral mode for all translators, although all except T4 still have a considerably higher pause ratio than the interpreters also under the sight translation condition.

The differences between pause patterns in the written and oral modalities may be ascribed to many factors. The time constraint felt by both interpreters and translators when working in the oral modality is the most obvious, along with the fact that in order to produce coherent oral discourse without having the target text on the screen, the translator/interpreter has to produce and monitor larger units of meaning rather than smaller segments.

But the pause patterns and disfluencies observed in the translators' oral output may also be explained by the translators' working habits and tradition of focussing almost exclusively on the written modality of translation, which is reflected in the way translation is taught. There is no clear-cut relationship between the pause/hesitation pattern and the quality of the output; yet, T3 and T4, who have the lowest number of pauses in the translator group, are also the ones who receive the best quality scores (Table 2). A tentative conclusion could be that very hesitant and disfluent production compromises the quality of the output, which again lends support to the recommendation to emphasize the oral modality in translator training.

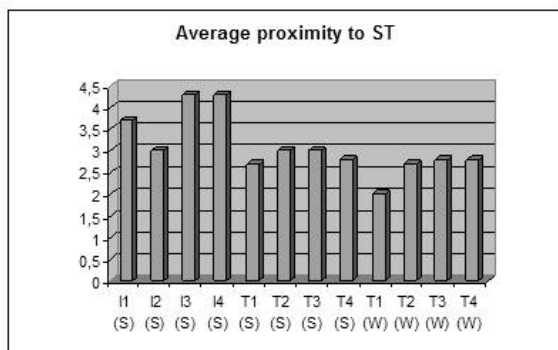
#### 4.4. Proximity

By contrast to the process-oriented aspects discussed so far, the analysis of proximity is based on the written and oral translation products.

When asked to give their judgments of the proximity of the target text to the source text, the raters gave the scores shown in Figure 6, where 1 indicates that the translation is very close to the ST, and 5 indicates a very free translation. The notion of proximity was not explained or defined to the raters, and we did not ask our raters to consider whether a very free translation, for instance, indicated high or low quality. Yet, the inter-rater agreement was high, with SD between 0 and 0.5 (in two cases 1), indicating a shared perception of the notion of proximity among the raters.

FIGURE 6

Proximity to source text<sup>14</sup>



The interpreters' target texts (with averages between 3 and 4.3) are more free than the translators' target texts in general, and the translators' sight translations are slightly more free than their written translations. Hence, there is a slight tendency for the translation product to become freer as the orientation towards the oral modality becomes more prominent. This is in accordance with what we expected based on Schäffner (2004) and Pöchhacker (2004). Although Shlesinger and Malkiel (2005: 179-182) found that, compared with translators, interpreters apply a more form-based strategy when processing cognates, their findings were limited to the lexical level and did not necessarily generate any conclusions with regard to strategies on a broader textual level.

Of all the translation products, T1's written translation is considered to be closest to the ST, whereas I1's (sight) translation is considered to be very free. It is interesting to note that T1's written translation received a very low quality score, and I1's sight translation received the highest quality score (Table 2). However, I3 and I4, who received the lowest quality scores, also produced the most free translations. Hence, there seems to be no correlation between quality and the level of proximity in our data, and the tendency for oral translations to be more free speaks neither for nor against augmenting the oral modality in translator training.

## 5. Conclusions and perspectives

In the present article, we set out to investigate sight translation, a hybrid between translation and interpreting, with a view to exploring the notion of speaking your translation. We found large differences in output rate when comparing oral and written modalities: interpreters produced 142 words per minute on average, whereas translators produced 74 words per minute under the sight translation condition and 17 words per minute under the written translation condition. We did not find that the lower output rate under the written translation condition was justified by significantly higher quality output. We then went on to explore behavioural characteristics of sight versus written translation in order to investigate the potential for incorporating oral translation in translator training curricula as a supplement to traditional written translation practice.

Gaze data (hotspot analyses, source text fixation count and average fixation duration) generally pointed to the same conclusion: the interpreters appear to process segments more consecutively and to work in a more focussed manner than the translators, who have more regressions and backtracking, searching other areas of the screen to look for clues for the current segment.

The pause analyses yielded similar results of fundamental behavioural differences between translators and interpreters in the oral and written modalities. Under the written translation condition, the translators appeared to be hesitant, pausing sometimes before each word. The interpreters seemed more self-assured and produced their target text fluently with very few pauses. In the oral translation condition, the translators were less hesitant although most of them still had considerably higher pause ratios than the interpreters.

There appeared to be a slight tendency for the translation product to become freer as the orientation towards the oral modality became more prominent. There seemed to be no correlation between quality and the degree of ST/TT proximity in our data.



Our experimental design, combining keylogging, eye-tracking and quality ratings of spoken and written output, serves to throw light on translation and sight translation processes from different perspectives. Observations from the combination of data sources point to fundamental behavioural differences between interpreters and translators: interpreters seem to be more focussed and they translate faster and more continuously than the translators, who appear hesitant and process the text discontinuously, in particular in the written translation condition.

Our analyses support our overall assumption that working in the oral modality, *speaking your translation*, using speech recognition technologies, which are rapidly gaining ground, seems to have a lot to offer in terms of saving time and effort without compromising the output quality significantly. Indeed, by adopting strategies associated with the oral modality, translators might acquire alternative working habits, processing larger ST segments and producing TT segments more fluently and continuously. Where appropriate, they may opt for fast-and-ready solutions rather than tirelessly searching for “the perfect word.” In some circumstances, depending on text type and domain familiarity, the end product might even gain from being produced orally. In other cases, a written version of the translator’s spoken translation could be produced by means of speech recognition software and serve as a draft for subsequent revision. However, the time saved by speaking the translation leaves plenty of room for revision.

Professional translators must be prepared to invest time and effort in changing their working routines and training the speech recognition software. Yet, our results indicate that the effort is worthwhile and point to the relevance of using speech recognition in translation, not only for professional translators trying to meet the ever increasing demands for more value for less money, but also for universities responsible for training translators. By incorporating oral translation in translator training as a deliberate practice activity, the next generation of translators may blaze the trail for speech recognition technology as a valuable new translation tool in the translator’s toolbox, which may give translators a competitive edge in an ever more demanding market, where the players are not only other human translators, but also fully automated translation solutions.

## NOTES

1. “Comprehension and production processes in translation and interpreting hybrids” conducted under the auspices of the Center for Research and Innovation in Translation and Translation Technology (CRITT) at Copenhagen Business School (<<http://www.critt.dk>>). The general aim of the research project is to explore the field between translation and interpreting and the potential for convergence between the written and oral modalities of translation.
2. Eye tracking system from Tobii (<<http://www.tobii.se>>, visited on June 08, 2009).
3. Official Eye-to-IT website (<<http://cogs.nbu.bg/eye-to-it/>>, visited on June 11, 2009).
4. For the purpose of the project described in note 1, the interpreters also rendered a simultaneous interpretation of the rest of the speech where subtitles were provided for part of the speech, but only data from their sight translations will be investigated here.
5. Description of Bysoc and transcription criteria (<<http://bysoc.dyndns.org/>>, visited on June 11, 2009).
6. An error analysis will not be included at this point.
7. Praat website with information and downloads (<<http://www.fon.hum.uva.nl/praat/>>, visited on June 11, 2009).
8. Praat scripts developed by Miette Lennes (<<http://www.helsinki.fi/~lennes/praat-scripts/>>, visited on June 11, 2009).



9. As appears from the figures in Table 2, the inter-rater agreement is high. This is confirmed when calculating the standard deviation (SD) on each parameter, which generally ranges from 0 to 0.6 (in few cases, it is 1 or 1.2).
10. Number 1 indicates highest quality, whereas 5 indicates lowest quality.
11. Number 1 indicates the lowest time requirement (approx. 0-10 min.), whereas 5 indicates the highest (approximately 40-50 minutes).
12. In future studies, the typing speed variable may have to be factored into the calculations.
13. S = sight translation, W = written translation.
14. SD ranges between 0 and 0.6 and in a single case (T2, sight translation), it is 1.

## REFERENCES

- AGRIFOGLIO, Marjorie (2004): Sight translation and interpreting: A comparative analysis of constraints and failures. *Interpreting*. 6(1):43-67.
- AHRENS, Barbara (2005): Prosodic phenomena in simultaneous interpreting: A conceptual approach and its practical application. *Interpreting*. 7(1):51-76.
- BADDELEY, Alan (1986): *Working memory*. Oxford: Clarendon Press.
- CHAFE, Wallace and DANIELEWICZ, Jane (1987): Properties of spoken and written language. In: Rosalind HOROWITZ and S. Jay SAMUELS, eds. *Comprehending Oral and Written Language*. San Diego: Academic Press, 83-113.
- CLIFTON, Charles, STAUB, Adrian, and RAYNER, Keith (2007): Eye Movements in Reading Words and Sentences. In: Roger P.G. van GOMPEL, Martin H. FISCHER, Wayne S. MURRAY, and Robin L. HILL, eds. *Eye Movements: A window on mind and brain*. Amsterdam: Elsevier, 341-371.
- DE GROOT, Annette M. B. (1997): The cognitive study of translation and interpretation: Three approaches. In: Joseph H. DANKS, Gregory M. SHREVE, Stephen B. FOUNTAIN, and Michael K. McBEATH, eds. *Cognitive Processes in Translation and Interpreting*. London: Sage, 25-56.
- DRAGSTED, Barbara (2005): Segmentation in translation: Differences across level of expertise and difficulty. *Target*. 17(1):49-70.
- DRAGSTED, Barbara and HANSEN, Inge G. (2007): Speaking your translation. Exploiting synergies between translation and interpreting. In: Franz PÖCHHACKER, Arnt L. JAKOBSEN, and Inger M. MEES, eds. *Interpreting Studies and Beyond. Copenhagen Studies in Language*. Vol. 35. Copenhagen: Samfundslitteratur, 251-274.
- ERICSSON, K. Anders (2000): Expertise in interpreting. An expert-performance perspective. *Interpreting*. 5(2):187-220.
- GAMBIER, Yves (2004): Translation studies: A succession of paradoxes. In: Christina SCHÄFFNER, ed. *Translation Research and Interpreting Research: Traditions, Gaps and Synergies*. Clevedon: Multilingual Matters, 62-70.
- GARZONE, Giuliana (2002): Quality Norms in Interpretation. In: Giuliana GARZONE and Maurizio VIEZZI, eds. *Interpreting in the 21st Century. Challenges and Opportunities*. Amsterdam: John Benjamins, 107-119.
- GILE, Daniel (1995): *Basic Concepts and Models for Interpreter and Translator Training*. Amsterdam: John Benjamins.
- GILE, Daniel (1997): Conference interpreting as a cognitive management problem, In: Joseph H. DANKS, Gregory M. SHREVE, Stephen B. FOUNTAIN, and Michael K. McBEATH, eds. *Cognitive Processes in Translation and Interpreting*. London: Sage, 196-214.
- GÖPFERICH, Susanne, JAKOBSEN, Arnt L. and MEES, Inger M., eds (2008): *Looking at Eyes: Eye-tracking studies of reading and translation processing. Copenhagen Studies in Language*. Vol. 36.
- HANSEN, Inge G. and SHLESINGER, Miriam (2007): The silver lining: Technology and self-study in the interpreting classroom. *Interpreting*. 9(1):95-118.
- HOUSE, Juliane (1997): *Translation Quality Assessment. A Model Revisited*. Tübingen: Gunter Narr.

- JAKOBSEN, Arnt L. (1998): Logging Time Delay in Translation, LSP Texts and the Translation Process. *Copenhagen Working Papers in LSP*. Vol. 1, 73–101.
- JAKOBSEN, Arnt L. and SCHOU, Lasse (1999): Translog documentation. In: Gyde HANSEN, ed. *Probing the Process in Translation: Methods and Results. Copenhagen Studies in Language*. Vol. 24. Copenhagen: Samfundslitteratur, 151–186.
- JAKOBSEN, Arnt L., JENSEN, Kristian T. H., and MEES, Inger M. (2007): Comparing modalities: Idioms as a case in point. In: Franz PÖCHHACKER, Arnt L. JAKOBSEN, and Inger M. MEES, eds. *Interpreting Studies and Beyond. Copenhagen Studies in Language*. Vol. 35. Copenhagen: Samfundslitteratur, 217–249.
- KALINA, Sylvia (2002): Quality in Interpreting and Its Prerequisites – A Framework for a Comprehensive View, In: Guiliiana GARZONE and Maurizio VIEZZI, eds. *Interpreting in the 21<sup>st</sup> Century. Challenges and Opportunities*. Amsterdam: John Benjamins, 121–130.
- MEAD, Peter (2000): Control of pauses by trainee interpreters in their A and B languages. *The Interpreters' Newsletter*. 10:89–102.
- MEAD, Peter (2002): Exploring hesitation in consecutive interpreting. In: Guiliiana GARZONE and Maurizio VIEZZI, eds. *Interpreting in the 21st Century: Challenges and Opportunities*. Amsterdam: John Benjamins, 73–82.
- O'BRIEN, Sharon (2006a): Eye-tracking and translation memory matches. *Perspectives: Studies in Translatology*. 14(3):185–205.
- O'BRIEN, Sharon (2006b): Pauses as indicators of cognitive effort in post-editing machine translation. *Across Languages and Cultures*. 7(1):1–21.
- PÖCHHACKER, Franz (2002): Researching interpreting quality: Models and methods. In: Guiliiana GARZONE and Maurizio VIEZZI, eds. *Interpreting in the 21st Century: Challenges and Opportunities*. Amsterdam: John Benjamins, 95–106.
- PÖCHHACKER, Franz (2004): *Introducing Interpreting Studies*. London: Routledge.
- PYM, Anthony (2007): On Shlesinger's proposed equalizing universal for interpreting. In: Franz PÖCHHACKER, Arnt L. JAKOBSEN, and Inger M. MEES, eds. *Interpreting Studies and Beyond. Copenhagen Studies in Language*. Vol. 35. Copenhagen: Samfundslitteratur, 175–190.
- RADACH, Ralph, KENNEDY, Alan, and RAYNER, Keith (2004): *Eye Movements and Information Processing During Reading*. New York: Psychology Press.
- RAYNER, Keith (1998): Eye Movement in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*. 124(3):372–422.
- RICCARDI, Alessandra (2002): Interpreting research: Descriptive aspects and methodological proposals. In: Guiliiana GARZONE and Maurizio VIEZZI, eds. *Interpreting in the 21st Century: Challenges and Opportunities*. Amsterdam: John Benjamins, 73–82.
- SCHÄFFNER, Christina (2004): Researching translation and interpreting. In: Christina SCHÄFFNER, ed. *Translation Research and Interpreting Research: Traditions, Gaps and Synergies*. Clevedon: Multilingual Matters, 1–9.
- SCHJOLDAGER, Anne (1996): Assessment of simultaneous interpreting. In: Cay DOLLERUP and Vibeke APPEL, eds. *Teaching Translation and Interpreting*. Vol. 3. Amsterdam: John Benjamins, 187–195.
- SETTON, Robin and MOTTA, Manuela (2007): Syntacrobatics. Quality and reformulation in simultaneous-with-text. *Interpreting*. 9(2):199–230.
- SHLESINGER, Miriam (1997): Quality in simultaneous interpreting. In: Yves GAMBIER, Daniel GILE, and Christopher TAYLOR, eds. *Conference Interpreting: Current Trends in Research. Proceedings of the International Conference on "Interpreting: What Do We Know and How?" (Turku, August 25–27, 1994)*. Amsterdam: John Benjamins, 123–131.
- SHLESINGER, Miriam (2004): Doorstep inter-subdisciplinarity and beyond. In: Christina SCHÄFFNER, ed. *Translation Research and Interpreting Research: Traditions, Gaps and Synergies*. Clevedon: Multilingual Matters, 116–123.
- SHLESINGER, Miriam and MALKIEL, Brenda (2005): Comparing modalities: Cognates as a case in point. *Across Languages and Cultures*. 6(2):173–193.