Inside the Monitor Model: Processes of Default and Challenged Translation Production

Michael Carl and Barbara Dragsted
Copenhagen Business School
Dalgas Have 15, DK-2000 Frederiksberg

It has been the subject of debate in the translation process literature whether human translation is a sequential and iterative process of comprehension-transfer-production or whether and to what extent comprehension and production activities may occur in parallel. Tirkkonen-Condit (2005) suggests a "monitor model" according to which translators start in a literal default rendering mode, and a monitor interrupts the default procedure when a problem occurs. This paper proposes an extension of the monitor model in which comprehension and production are processed in parallel by the default procedure. The monitor supervises text production processes, and triggers disintegration of the translation activity into chunks of sequential reading and writing behavior.

To investigate this hypothesis, we compare text copying with translation activities under the assumption that text copying represents a typical literal default rendering procedure. Both, translation and text copying, require decoding, retrieval and encoding of textual segments, but translation requires in addition a transfer step into the target language. Comparing reading and writing behaviour obtained in the copying and translation experiments, we observe surprisingly many similarities, which also suggests similarities in the underlying processes. Copyists deviate from the default literal text reproduction into more effortful text understanding, and much of the translators' behaviour looks like simple text copying. During translation as well as during text copying we observe that translators and copyists resort to sequential reading and writing patterns which seem to be triggered through target text production problems.

1   Introduction

In his seminal book, Gile (1995) suggests a stratificational translation process model, in which a translator iteratively reads a piece of the ST and then produces its translation. First the translator creates a "Meaning Hypothesis" for a ST chunk (i.e. a Translation Unit) which is consistent with the "context and the linguistic and extra linguistic knowledge of the translator" (p. 107) before the translation is produced. Similarly, Craciunescu et al. (2004) claim that "the first stage in human translation is complete comprehension of the source language text". Only after this complete (i.e. *deep*) comprehension is achieved can the translation be produced. Also Angelone (2010) supports that translators process in cycles of comprehension-transfer-production and that "uncertainties" of translators can be attributed to any of the comprehension, transfer, or production phases.

Some scholars challenge this view, stating that translation processes can also be based on a *shallow* understanding and that ST understanding and TT production can occur in *parallel*. According to Ruiz et al. (2008) "the translator engages in partial reformulation while reading for the purpose of translating the source text". They assume that in parallel processing "code-to-code links between the SL and TL [are involved] at least the lexical and syntactic level of processing". Similarly, Mossop (2003) claims the existence of "direct linkages in the mind between SL and TL lexicogrammatical material, independent of 'meaning'", and that a translator "automatically produces TL lexical and syntactic material based on the incoming SL forms".

In a study comparing reading behaviour for different purposes, Jakobsen & Jensen (2008) investigate (among other things) the difference between test persons reading a text for comprehension and reading a similar text in preparation for translating. Their study showed that reading purpose has a "clear effect on eye movements and gaze behaviour" and they suggest "that a fair amount of pre-translation probably enters into the reading of a text as soon as it is taken to be the source text for translation" [p.116].

Although it is unclear what is exactly meant by "pre-translation", such findings are obviously in contrast with the eye-mind[1] hypothesis when assuming a stratificational model of translation. Reading with "a fair amount of pre-translation" implies certainly different mental activities than reading for understanding. Since in both cases the eyes remain on the ST it may be difficult (if not impossible) to disentangle which fixations in the logged gaze data are to be linked to text understanding and which ones are due to pre-translation, and hence either the eye-mind hypothesis has to be weakened or the stratificational model of translation has to be reconsidered.

We assume, with Tirkkonen-Condit, that "literal translation is a default rendering procedure, which goes on until it is interrupted by a monitor that alerts about a problem in the outcome. The monitor's function is to trigger off conscious decision-making to solve the problem" (Tirkkonen-Condit 2005: 407-408). In our interpretation, the literal default rendering procedure implies parallel, tightly interconnected text production and comprehension processes: while the mind is engaged in the production of a piece of text, the eyes search for relevant textual places to gather the required information needed to continue the text production flow. When this default procedure is interrupted by the monitor, can we observe gaze patterns on the ST or on the TT which indicate comprehension- or production-related translation problems. Similarly, Ruiz et al. (2008) suggest that "lexical and syntactic properties may follow different time courses", so that different levels of understanding are obtained at different stages in the concurrent reading/production processes.

In this paper, we will show examples of unchallenged, parallel processing in text copying and translation production. We will also point to passages of conscious, effortful text production which suggest a more disentangled relation between comprehension and production. Similar to Ruiz et al. (2008), who distinguish between shallow/parallel and deep/sequential[2] processes in translation,[3] we find that translators switch between the two modes, but more frequently engage in shallow/parallel processing. In addition, we find that deeper understanding in the decision-making processes is triggered through translation production problems, rather than difficulties in ST understanding. This finding coincides with Gile (1995) who reports that deeper understanding of the ST may emerge through problems in TT production, rather than when first reading a ST passage. He points out that the translation practice indicates processing from a production-based perspective:

> Oftentimes, the translator does not test Meaning Hypothesis until after verbalising it in the target language (...) Frequently, he or she only realizes there is a problem when trying to read the first target-language version (...) in other words, when already in the reformulation phase. [p.110]

---

[1] there is no appreciable lag between what is fixated and what is processed (Just & Carpenter, 1984)
[2] Shallow-deep refer to the level of the translator's (conceptual) representation during the translation process, while parallel-sequential refer to the observable coordination of comprehension (eye gaze movements) and production processes (keystroke presses) in time.
[3] Respectively horizontal/parallel vertical/serial in in their terminology.

A clear-cut allocation of "uncertainties" to one of the stratificational processes becomes difficult, since such processes do not normally exist independently in the translator's mind. Not only is it infeasible (or impossible) to distinguish between comprehension and pre-translation activities during reading for translation, but also the borders between ST understanding and TT production problems become blurred.

In order to investigate the basic literal default rendering procedure, we first look at text copying. We take it that copying (i.e. re-typing) a text is much more a shallow/parallel process than translation: 1) apart from a lexical encoding and decoding (John, 1999), text copying does not, in theory, require any deep ST (or TT) understanding 2) copying can proceed in parallel to a maximal degree, since no revision[4] and no lexical or structural transfer is required. Typing patterns and speed would thus essentially depend on the typing skills of the copyist. Comparing copying behaviour and translation behaviour would reveal the additional effort of translation.

Our investigation is based on empirical data obtained in 10 copying sessions and 15 translation sessions. The experiments were recorded using the Translog 2006 software, which logs keystroke and gaze movements during a reading, translation or text production task. Section 2 gives an overview over the text type and the translators' activity data. In section 3, we first describe a cognitive model of text copying, which predicts word comprehension and typing time. We then illustrate the model with two copying examples. The first example shows an instance of unchallenged copying, the second example shows how text understanding plays a role, also when copying a text. Section 4 discusses four passages from a translation experiment and compares it to the copying data. We look at instances of unchallenged parallel and sequential translation processes and at two instances of challenged translation which requires re-reading of the source and the target texts. The examples show that much of the translation process resembles text copying, and that deeper understanding processes are triggered through text production problems. While these examples show individual patterns, the last section 5 provides evidence that distribution of pauses in translation is similar for different translators translating the same text.

2   Experimental data for copying and translation

We investigate and compare user activity data of two experiments, a copying experiment and a translation experiment. In sections 2 and 4, we re-use and analyze a subset of the translation data that was collected by KTH. Jensen for his PhD thesis (Jensen, 2011). The copying data which is discussed in section 2 and 3 was taken from an experiment which Annette C. Sjørup (2011) conducted for her PhD thesis. Some statistics about these translation/copying processes is shown in Figure 1. In section 5 we re-use translation data which was collected in the context of the Eye-to-IT project (http://cogs.nbu.bg/eye-to-it/). All process data was collected with different goals and for different purposes over several years in the CRITT/CBS under controlled conditions; all process data is based on processing short English texts between 110 and 170 words long.

2.1   A copying experiment

In the first experiment, an English financial text about bank loans was copied by 10 English L2 speakers. The text consisted of 9 sentences, totalling 169 words

---

[4] Some revision may be going on, for instance correction of typos, but these activities are of a different kind than most of those in translation revisions.

(punctuation marks are counted as one word), and 945 characters, which makes an average of 5.59 characters/word. For data acquisition, Translog-2006 was used (Jakobsen, 1999), which divides the screen into two windows. The upper window plots the source text, or in this case the text that had to be copied. The translator typed a copy of this text into the lower window. While most of the reproduced texts had the same number of words, none of them were identical to the source text. Besides a number of typos, there were also productions of semantically (or phonetically) similar words. In several instances, "votes" was reproduced as "voted"; "issuers ability" as "issuers' ability" (which was actually incorrect in the source text, and corrected in the reproduction), in one case "any" was retyped as "the", in another case "choke" as "chose" and "card" was reproduced as "credit card". This indicates to us, that a decoding and "understanding" of the texts has clearly taken place during copying where even instances of explication can be observed.

Figure 1 plots some of the process data for the 10 copying sessions. It shows that the text was reproduced in less than 3 minutes by one translator (4), and in more than 7 minutes by another translator (7), which is a factor of approx. 2.5 between the fastest and the slowest copyist. The figure also shows that the number of deletions varies between 27 (for the fastest copyist) and 94 (for one of the slowest copyist). The median typing speed of all 10 copyists is 282.81ms per character (average 266.27).

	Translation Task				Copying Task			
	Time	keys	del	ms/key	Time	keys	del	ms/key
1	265762	916	49	290.13	255342	1100	80	232.13
2	396390	999	49	396.79	388155	1113	94	348.75
3	654681	1053	107	621.73	193344	982	28	196.89
4	259094	918	40	282.24	169039	980	27	172.49
5	260644	848	22	307.36	260530	1021	48	255.17
6	350854	973	67	360.59	278291	1044	59	266.56
7	411156	959	49	428.73	430353	1009	40	426.51
8	349750	965	47	362.44	228016	994	33	229.39
9	379272	967	58	392.22	288249	1035	54	278.50
10	209142	937	57	223.20	273975	1069	69	256.29
11	322110	979	41	329.02				
12	380643	971	33	392.01				
13	316730	1083	81	292.46				
14	352497	969	76	363.77				

Figure 1: Translation and copying time, processed number of keystrokes (keys), number of deletions thereof (del), as well as average inter-key duration for the 15 translators and 10 copyists.

2.2   A translation experiment

In the second experiment, 15 translators translated an English (L2) text into Danish (L1), using a similar Translog setting. This text was a news text about the Olympic Games in Beijing, which consisted of 5 sentences and 160 words (845 character length) with an average word length of 5.32 characters/word. Figure 1 (left) shows the translation process data of the 15 translation sessions. The median inter-keystroke delay in this translation task was 362.44 ms/character (average 360.19), which is about 1.4 times more than in the copying task. The time spent by the slowest translator (654 seconds) is more than 3 times higher than that of the fastest translator (209 seconds), but looking at the difference of inter-keystroke delay, this difference reduces to 2.79, since many more keystrokes (particularly deletions) are produced by

the slow translators than by the fast translators. There are slightly more corrections (deletions) during translation than during copying: approx. 5.14% and 5.68% of all keystrokes are deletions in the copying and translation tasks respectively.

The analysis suggests that differences between copying and translation are not as large as one might expect: ST reading and TT production activities occur in both the copying and the translation tasks, whereas transfer from source language to target language only takes place in the translation task. If we assume a stratificational translation process model, which proposes that only one activity – ST understanding, TT production, or transfer – can take place at any one time, we are left with less than 100ms per ST character for transfer from the source to target language. While this might be an appealing model for some, we give evidence below that reading (ST understanding) and typing (TT production) can occur in parallel, which leads us to a different set of conclusions.

2.3   Gaze behaviour during reading, copying and translation

As mentioned earlier, it has been found that different reading purposes imply different gazing patterns (Jakobsen & Jensen, 2008): e.g. the gaze behaviour in reading for understanding is different from reading for translation, which is different from reading for text copying. Dragsted (2010), for instance, observes on average 3.2 and 5.7 fixations per word by expert and student translators respectively in a translation task, while less than one fixation per word was measured during reading for comprehension for normal college students.[5] That is, there are approximately 3.5 times more fixations on the ST in a translation task than in a text comprehension task. Dragsted assumes that the additional ST fixations are likely to result from the planning of TT production and the effort of transforming ST expressions into meaningful TT[6]. Thus, more than 2/3 of the ST reading effort during translation seems to be linked to translation-related activity and the planning/validation of TT production.

The number of ST fixations during copying in our data was, on average, 1.8 fixations/word, and varied between 207 and 538 fixations for the 169 word text. Text copying thus requires approximately twice as many fixations on the ST than reading for comprehension. During translation drafting we measured between 309 and 595 fixations on the ST. For a text of 160 words, the average was 2.4 fixations per ST word. These figures are much lower than those reported by Dragsted (2010). One explanation might be that we only count fixations during translation drafting; there are additional ST fixations during initial orientation and in the revision phases.

3   Text Copying

In this section we will first introduce a cognitive model of text copying, i.e. reading and re-typing. We will then empirically evaluate the model in two copying examples. Whereas the first example in section 3.1 confirms the model for unchallenged, smooth copying, the second example, section 3.3 shows that comprehension difficulties may play a role in the copying process.

3.1   A cognitive model of typing

John (1999) suggests a three step model for text copying. First a perceptual operator perceives a written word. Then a cognitive operation retrieves the spelling of the

---

[5] http://www.learninginfo.org/eye-movements-reading.htm
[6] But we maintain that it might be impossible in detail to figure out which fixations belongs to which of the processes.

word from long-term memory, and finally a motor operator finds a key on the keyboard and hits it. John makes a distinction between copying of single (sequences of) characters and more complex symbols. For the more complex symbols, like words and syllables, a cognitive operator is required to retrieve the spelling of the word from memory long-term memory and to initiate the typing of each character. Hence, the perception of a word and its typing requires the encoding (perception) and decoding (memory retrieval) of the symbol. With the assumption that a skilled typist produces about 30 gross words per minute (gwpm), John arrives at the following model:

1. A perceptual operator reads a word of about six letters and encodes it in 340 ms
2. The cognitive operator has a cycle time of 50 ms to retrieve the spelling and to activate the typing of the characters from the spelling list
3. A motor operator needs 230 ms on an alphanumeric keyboard at a rate of about 30 gwpm

In addition, John assumes that each of the operators works serially in themselves (only one keystroke can be processed at any one time) but that they can work in parallel with each other, with the serialisation restrictions that:

- perception has to be complete before getting the spelling or initiation of a character can begin.
- once a character has been initiated with a cognitive operator, the motor operator cannot be stopped.
- the perceptual processor (eyes) stays three words (chunks) ahead of the cognitive processor.

John uses this model to analyse which of the three operations is the limiting factor in text copying. She finds that the overall typing speed depends primarily on the time needed for motor activity, rather than for perception or cognitive control. In line with other researchers, John assumes that the 50ms for cognitive cycle time are constant, and hence typing skills are often the limiting factors.

Comparing this model with the figures in the previous section, we note that some of the typists type much quicker than the time predicted for 30gwpm and are able to copy a text in 75% of the time predicted by John. The average measured typing time is, however, about 115% of the predicted 230ms, which suggests that some of our copyists need more time for the motor actions, and/or they engage in additional (reading) activities during which the copyist performs no typing. The next two sections will look at some examples in more depth.

3.2 Copying an easy text

The first example illustrates the typing process of the 3rd sentence of the copied text:

(1) The rise in unemployment has spattered a once-profitable business with red ink.

The sentence consists of 13 words (including sentence final full stop) with 80 characters (including inter-word blank spaces). One of the copyists copied the sentence in 21 seconds with 5 typos. Figure 2 shows the progression graph: the vertical Y-axis plots the original sentence which was to be copied; the horizontal X-axis represents a time line in ms in which the reading and typing activities take place. Single fixations on the source text are marked by a blue circle. Typing activities consist of text insertions (in black) and deletions (in red) shown on the time line.
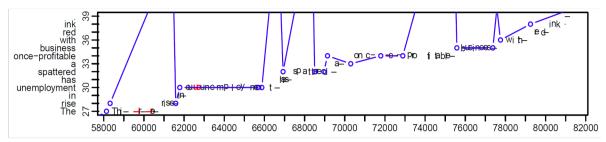
Figure 2: A progression graph for unchallenged text copying

The figure shows a time segment of 21 seconds between ms 58,000 to around 81,000 in which the sentence is being copied. At the beginning, the typist first gazed at the two words "The" and "rise" before starting to type. Two typos occurred in the first word when reproducing "The". These typos were immediately corrected. Perrin (2003) suggests a short-hand form to represent writing activities, where corrections are represented in square brackets. In this notation, the typing pattern would be represented as: "Th[i-][r-]e", which is read as follows: First the typist writes "Thi-".[7] Then "i-", the blank space (i.e. "-") and the "i" are deleted, then "r-" is typed and deleted, until finally the correct "e" is typed. There are thus 4 correcting keystrokes in the production of "The". The typist goes then on immediately with the typing of "rise", without looking back into the source text. There are two fixations just before ms 62,000, one on "rise" and one on "unemployment", the latter while already typing "in". From the progression graph it appears that the word "in" was actually not looked at; however, it is likely that this word was in the parafoveal scope of the fixation on "rise".

The copying process then goes on smoothly. There are two more typos and deletions, but the typist seems to copy the text without much hesitation, looking in general only one word ahead in the source text. In some cases the typist verifies the spelling of a word that is being typed (e.g. "spattered" around time stamp 69,000), and in other instances she scans the next word already while still typing the previous word (e.g. also "spattered" around 67000 ms).

According to John's model, a skilled typist would minimally need 80 * 230ms for keying the 80 characters, plus an initial 390ms for reading the first word, 340ms for perception of the first word, and 50ms for retrieval of the spelling list. Counting the 7 typos as 2 keystrokes each, the predicted typing time, according to John's model, amounts to approximately 22 seconds. Compared with the observed typing time of 23 seconds, the prediction is pretty good with an error rate of less than 5%. While the model thus seems to be quite exact for predicting the overall time needed when typing activities go smoothly, it does not seem to be so precise for predicting the gaze activities and the structure of the gaze/keystroke coordination: John's model predicts a three words look-ahead; however, in many cases only one word is looked ahead from the word that is currently being copied, and longer (or more difficult) words, e.g. "unemployment" and "once-profitable", may trigger re-fixations, whereas some short words are not fixated at all.

3.3   Copying difficult texts passages

The following example (2) illustrates the impact of reproducing unusual idioms on the coordination of gaze and typing activities. More difficult/unusual or surprising source text passages trigger extended reading behaviour and reduce the copying speed. Example (2) shows the 6th sentence taken from the same text as Example (1):

---

[7]   The blank space is represented as a dash "-" in the graph and in figures below.

*(2) The industry's claim that the **bill will choke off access to credit is a bit rich given its** own rush to reduce its unsecured lending.*
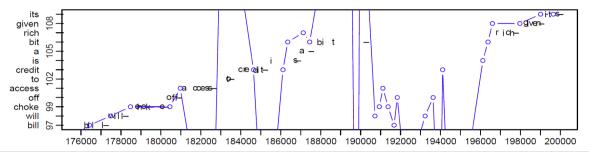

Figure 3: A writing progression graph for challenged text copying

Figure 3 shows the progression graph of the fragment in bold, which contains the expression "to be (a bit) rich".[8] The fragment has 13 words, which are copied in 24 seconds. It consists of 61 characters which, according to John's model, should be typed in approx. 14.5 secs. However, the typist needs 24 secs, with a pause of approx. 8 seconds between seconds between seconds 189 and 197, just before copying the word "rich". Note that "rich" is the head word of the idiom "to be a bit rich", and the idiom might not be known to the typist, English being her/his L2.

We observed twice as many fixations in the more difficult copying in example 2, as in the unchallenged copying activity of the 13 word sentence in Example 1 (see Figure 2). Copying of the fragment in example 1 was achieved with 21 fixations, which amounts to 1.6 fixations per word. In contrast to this, there were 39 fixations on the ST when typing the more difficult 13 words fragment in Figure 3, which amounts to an average of 3 fixations per word.

Since we can safely assume that there is no problem in decoding, encoding or typing the word "rich", we believe that the extended reading activities of the ST segment indicate a verification or clarification of the idiom meaning, which seems to be unnecessary in a copying task.

It is also interesting to note that the additional reading activity occurs immediately before "rich" is being typed, despite the fact that "rich" was already previously gazed at around ms 187,000. During its first reading "rich" was obviously not recognised as problematic – at least there was no extended reading of the context. The monitor seems to run in the background and triggers conscious re-reading when passages are unclear, even during text copying. This indicates that (1) text understanding also takes place during text copying and (2) the understanding of more challenging expressions takes place when keying in the words, rather than (or in addition to) when initially reading the source text. This is in line with the citation of Gile (above) who states that production processes seem to trigger additional testing of "Meaning Hypotheses" which does not occur when initially reading the text. In the next section we will look at reading behaviour during translation and observe a similar pattern, where reading activities are triggered by TT production problems.

---

[8] According to the free dictionary (http://idioms.thefreedictionary.com/rich): "to be (a bit) rich" is something that you say when someone criticizes you to show that you do not think they are being fair because they are as bad as you

4   Translation

In this section we look at translation activities. We distinguish between sequential and parallel activities. The term "sequential" is used when a translator either reads (the ST) or writes (the TT) at any one time. During "parallel" activities, the translator reads and writes at the same time.

4.1   Parallel reading and writing

Figure 4 shows an example were parallel translation activities occur. It represents a translation progression graph for the translation of an English source sentence into Danish in example (3):

> *(3) English source sentence:*
> *Police officer Chris Gregg said that Norris had been acting strangely around the hospital*
>
> *Danish translation:*
> *P[i]olitiins[ep]pekt[rør]ør Chris Gregg sagte at Norris havde opført sig sært på h[i]o[p[si]s]sp[o]italet*

As in the previous progression graphs, the vertical axis plots the ST words while the horizontal axis represents a time interval of 28 seconds (seconds 149-177) during which the translation is being produced.
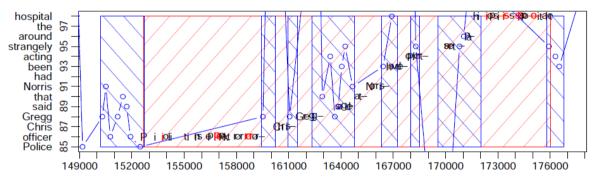

Figure 4: The translation progression graph shows parallel reading and text production.

In a copying task the produced text is identical to the ST which has been copied,[9] and therefore each ST word can be trivially aligned with its identical word in the produced copy. This is different in a translation task, where a single word in the produced translation can represent several words in the ST and vice versa, and discontinuous strings further complicate the translations process. For instance, Danish "Politiinspektør" is the translation of two English words. Thus,it is a 2-to-1 alignment, the translation of which appears in Figure 4 on the line of "officer" whereas the "Police" line is empty.

Typing activities are clustered into production units (PUs), marked in (red) upwards hatched boxes. A PU consists of successive typing activities, where no pause of more than 1000ms occurs between successive keystrokes. A PU is intended to represent coherent writing activities of a segment of text. The value of one second was chosen in line with a previous study (Carl and Kay, 2011) which report that a production pause of more than 1000ms in text production is likely to represent a shift of attention towards another segment. A PU thus subsumes all subsequent

---

[9] This should be in principle the case. However, we also found cases where copyists inserted or deleted additional punctuation marks, typos and even words.

keystrokes which are part of the same chunk. According to this criterion, the entire sentence in figure 4 was typed as one coherent production unit.

The (blue) circles are ST fixations and the downwards hatched boxes represent "fixation units". A fixation unit (FU) consists of a sequence of coherent ST[10] fixations, where no pause of more than 400ms occurs between the end and the beginning of two successive fixations. A FU thus represents a ST reading chunk. There are two large FUs in the progression graph in Figure 4. The first occurs during the time between seconds 149 and 152, at the beginning of a new sentence. The translator's eyes moved back and forth in the chunk "Police officer Chris Gregg said that Norris". After this the translation "Politiinspektør" was typed, including a number of typos which were immediately corrected (deleted characters are in red), but the following typing of the proper noun goes smoothly.

Obviously reading is not steady. The eyes jump over two, sometimes three words, back and forth, until a segment of text is understood well enough to start typing a translation. Note that the measured reading time in the sentence onset (circa 3 seconds) approximately corresponds to the perception and decoding time predicted in John's model: the 7 words would require 2380ms to perceive (at a rate of 340ms per word) plus the retrieval operation brings us to 2.5 seconds.

A second FU occurs between seconds 162 and 165. Here the fragment "Gregg said that Norris had been acting strangely" is read by jumping back and forth in the chunk. In contrast to the first FU, this reading activity occurs while typing the translation of "Gregg said that". Such parallel translation activities characterize experienced translators (Carl & Kay, 2011), capable of typing a translation while already reading ahead in the ST, whereas translation students resort frequently to an sequential mode, as will be discussed in section 4.2.

The Danish translation in example 3 consists of 12 words with 79 characters and 12 typos. According to the copying model above, a typist would need 24 seconds to key in the sentence. If we subtract the 3 initial seconds between time stamp 149,000 – 152,000 where the beginning of the sentence was scanned, and which is supposedly due to the translation activity, we measure 25 seconds production time vs. 24 seconds predicted by Johns model, an error of less than 5%. The translation was thus produced approximately at an expert copyists' speed, plus an initial orientation which amounts to 10% of the typing time. The second FU shows that reading, translating and writing can take place in parallel, and transfer does not necessarily require additional (measurable) time, as compared to the copying task. However, there is more gaze activity on the ST than in the case of unchallenged copying in example 1.

We may call this translation activity *transcoding*, i.e. substituting words or phrases in one language system by corresponding expressions in the target language system. It is an instance of the literal default rendering procedure, as discussed earlier.
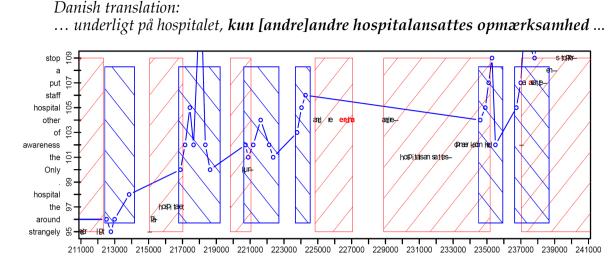
4.2   Sequential reading and writing

Figure 5 shows an example of mainly sequential translation activity. The translator is either reading a ST segment or writing a piece of TT, and the average translation time per character increases dramatically. The figure plots the translation progression of a sentence-final segment and the first words of the next sentence. The produced translation is shown in example 4:

(4) English source text:
    *… strangely around the hospital. **Only the awareness of other hospital staff** …*

---

[10] We only show here fixations on the ST. TT fixations are currently not processed with the software.

Danish translation:
… underligt på hospitalet, **kun [andre]andre hospitalansattes opmærksamhed** *…*

Figure 5: The translation progression graph shows sequential reading and text production.

Presumably, due to a syntactic reordering in the translation of "the awareness of other hospital staff" (see Figure 6), there is some reading activity before the translator starts translating the second sentence and between seconds 221 and 225 just after typing "kun"(translation of "only"). As noted above, a translation progression graph visualises the temporal development of translation equivalents. Units of translation equivalences (i.e. alignment units) were manually aligned in the final translation product, and the allocation of each keystrokes to one of those alignment unit are computed based on an algorithm described in (Carl, 2009). The alignment units may be more complex than merely a 1-to-1 relation. Figure 6 shows the more complex translation where an insight-out alignment is intermingled with a discontinuous English phrase. Four English words (a101, a103, a105, a106), the compound noun "hospital staff" together with the definite article "the" and the preposition "of" are translated as b99: "hospitalsansattes". The keystrokes in time which produce this translation appear in the translation progression graph of figure 6 on line 101. The words "awareness" and "other" change the order in the Danish translation and appear respectively on line 102 and 105, but are typed in the reversed order. Figure 6 shows the label "s", a semantic equivalence between the English function words and their Danish translation, while the compound noun component has no label. Kroman (2003) gives more background on this alignment schema, but we do not make use of the labels in this investigation. Obviously, the translator had to read 6 words ahead to produce this translation. However, s/he did not read further ahead than that, and thus took into account just enough, but not more context than was required to produce the translation.

The bold part in the translation example 4 consisted of 7 words with 42 characters, and was produced in approximately 19 seconds (timestamps 217000 - 236000). Including the 5 typos ("andre" (*other*) was first typed then deleted and then typed again) the copy model predicts 12.3 seconds for typing.
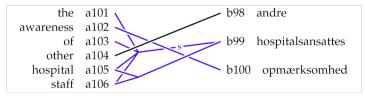

Figure 6: Translation fragment from figure 5

There are two pauses of 3 secs and 3.8 secs, in each of which the translator reads more or less the same seven words ahead of the ST word that is currently being translated. The first reading activity fits approximately with Johns' reading model, which predicts 2380 ms for perception of seven ST words, the additional reading time (approx. 0.5 second) may be allocated to transfer activities. We assume that the second pause, after typing "kun", is triggered by the "monitor" leading to a non-literal translation and a token reordering in the target text. The translation challenge is relative small syntactic re-phrasing of the English "of" construction into a Danish genitive. We are thus in a position to deconstruct the 19 seconds measured translation time of the 7-word segment into 1) ST reading time, approx 2.5 seconds 2) transfer and monitor time, approx. 4.3 seconds and 3) typing time, 12.2 seconds. Note that the 4.3 seconds correspond approximately to the transfer time of 42 characters, as predicted in section 2.2., and 12.2 seconds measured typing time are, again, pretty close to John's 12.3 seconds prediction.

In contrast to example 3, where we looked at parallel reading and writing, we observe a sequential reading/typing behaviour, in which only one of the activities is carried out at the same time. While there might be approximately the same amount of ST reading in a parallel and in a sequential translation mode, the overall translation time doubles in the latter one, since activities occur sequentially. Even though there seems to be some sort of "alerting" from the translator's cognitive monitor, it is perhaps not appropriate to assume a "deep" understanding of the text when reading 3 to 6 words ahead. In the next section we will look at more challenging instances of translation.

4.3   Instances of challenged translation

While the examples in Figures 4 and 5 show instances of unchallenged translation, where translation production proceeds smoothly in a parallel (example 3) or in an sequential mode (example 4), with only some words look-ahead, we will now look at more challenging translation problems which are more time consuming. This translation activity is characterised by delayed text production and associated with extended reading activities into the ST or TT context, beyond 5 or 6 words from the current translation position, or a production pause exceeding by far the expected decoding time predicted by John.

In this and the next section we look at two examples illustrating translation problems which trigger extensive re-reading of the TT (Example (5) and Figure 7) and re-reading of the ST (Example 6 and Figure 8). The fragments in example 5 consist of 30 English words which are translated into 38 Danish words. From this sentence, we investigate more closely a ST segment of 13 words, which is reproduced in bold in example (5):

> *(5) English source sentence:*
> *Although developing countries are understandably reluctant to compromise their chances of achieving better standards of living **for the poor, action on climate change need not threaten economic development.***
>
> *Danish translation:*
> *Selv om udviklingslandene forståeligt nok tøver med at risikere at ødelægge deres muligheder for at opnå en bedre levestandard **for de fattige, behøver tiltag over for klimaforandringer ikke at udgøre en trussel mod deres økonomiske udvikling**.*
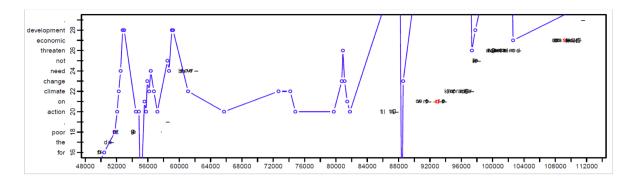
Figure 7: Translation progression graph of a translation problem

The translation progression graph in figure 7 plots the unfolding of the translation of the fragment in bold in example 5. The first three words "for the poor" did not pose any problems: the translator produced a monotonous, literal word-for-word translation into Danish, with only very little look-ahead into the ST. This unchallenged translation production was followed by a scanning of the entire main clause between time stamps 52,000 and 60,000. With the main clause at the second position, the finite verb needs to be clause initial in the Danish translation while the structure is SVO in the English source. The translator thus typed the translation of the finite verb ("need"), obviously without having a clear idea of how to render the remaining parts of that sentence. There is hence a second phase of more than 20 seconds between time stamps 62,000 and 86,000, in which the translator mostly re-reads the target text. The progression graph shows few fixation points, particularly between timestamps 62,000 and 80,000. The figure only shows the gaze movements on the ST window, but from studying the log file we know that the translator was concerned with re-reading the produced target text, obviously trying to find an appropriate translation solution for the current ST segment. Between 80,000 and approx. 82,000 the translator re-read the problematic ST chunk, maybe to verify or confirm a translation solution, before starting typing the translation from time 86,000.

Figure 8 shows the alignment of the translation which corresponds to the time segment in figure 7 and the bold part in example 6. While figure 7 illustrates *how* the
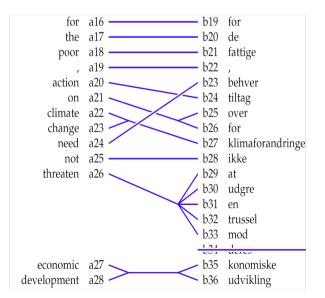

Figure 8: The alignment of the fragment in figure 7

translation was produced, figure 8 shows *what* was produced in the target text. It shows several cases of re-ordering: i) a syntactic re-ordering of the finite verb "need" ("behøver"[11]) moves into the first position in the Danish clause. Note that only after this was done, the long pause described above occurs. ii) The progression graph in figure 7 suggests that the main problem here was the rendering of "action on" which does not have a suited literal translation into Danish and requires a lexical choice. The alignment in Figure 8 also shows that iii) "threaten" is translated into an entire phrase, and and extra words ("deres") is inserted which do not have an English equivalent.

Despite the large number of modifications, here too, the translator only reads the minimum amount of text that is required to continue text production. Note that, as in the previous example, the translation of the finite verb is already typed before the translator knows how to render the remaining part of the clause. We assume, here too, that there is no problem in understanding the English source text, rather re-reading of the texts is triggered through production problems, for each of the syntactic and lexical re-organization problems. Parts of the translation are successively typed out as soon as enough context has been read, even if the continuation of the translation is still unclear. This anticipation may sometimes be misguiding and require a re-reading of the ST as the next example will show.

4.4   Erroneous translation anticipation

Another example of a challenged translation is shown in Figure 9. This progression graph plots an instance of a translation problem, where the translator re-reads the (already translated) context of the source sentence. Similar to the example in figure 2, the translator scans past context in the ST to retrieve information that is necessary to continue production.
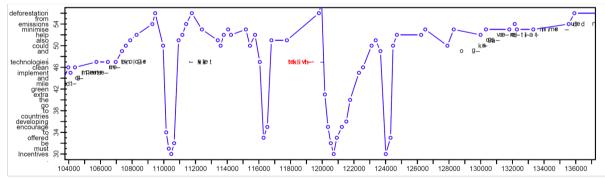

Figure 9: Translation progression graph with backward looking activities

The translation progression graph in figure 9 plots the production of the bold part in example 6, in which the English sentence consists of 28 words, and its translation into Danish consists of 31 words:

(6) *English source sentence:*
*Incentives must be offered to encourage developing countries to go the extra green mile **and implement clean technologies, and could also help minimise** emissions from deforestation.*

Danish translation:
Der skal tilbydes incitamenter til at opfordre udviklingslandene til at tage det ekstra

---

[11] diacritical characters (åøœ, etc.) as in "behøver", "udgøre", "økonomiske" etc. are lost in figure 8.

> *grønne skridt **og implementere rene teknologier[, hvilket] og kan også være med til at minimere** udledninger fra skovrydning.*

Similar to the previous example, here too, the beginning of the plotted part shows unchallenged translation. Already while typing the translation for "technologies" around time stamp 108,000, the translator reads ahead until the end of the sentence. She then looks at the beginning ("Incentives must be offered"), and reads the end of the sentence again. Then the translator writes the relative pronoun "hvilket" (around time stamp 112,000) which is deleted 6 seconds later (between timestamps 118,000 and 120,000). After repeatedly re-scanning the entire source sentence, the translator continues with fluent typing at time stamp 129,000 and produces an almost literal, word-for-word translation of the English original.
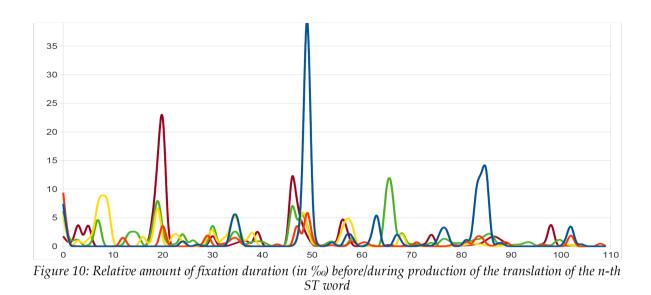
In this process, which lasted about 30 seconds, it is interesting to note how reading and text production interact. Presumably, the lack of a subject in the English clause "and could also help minimise emissions from deforestation" and its sentence final position has made the translator think that this might be a relative clause of "technologies", despite the fact that English "and" does not translate into a Danish relative pronoun. It is only after writing the relative pronoun ("hvilket") and after re-scanning the source sentence several times that the translator realises that the English clause is a main clause conjunction (with an omitted subject "Incentives"), and not a relative clause of "technologies".

The typing of the relative pronoun suggests a misguided expectation on the part of the translator which had to be sorted out and which required extra reading effort in the ST context. The example shows that there may be only a partial understanding of the text when a translator starts typing its translation, and that more advanced meaning representations are generated "on the fly" when needed, and only to the extent that they are necessary to proceed with text production. As shown in the previous examples, this extended need for understanding is triggered by problems associated with text production rather than a lack of ST comprehension. In reading research it is well known that the reader constructs an analysis of the syntactic structure of a sentence in a highly incremental manner, usually on a word-by-word basis (Staub, 2007). Our data show that this might be similar in translation production, going back to the source text when this procedure does not work out.

5   Distribution of ST fixation

As discussed in the previous examples, the fixations and fixation units on the ST are not equally distributed over the ST. At some points we observe long fixations and re-fixations on the texts, while other parts are translated smoothly without much problems. We have also seen that – in our examples – the extended gazing duration on the ST is triggered by production rather than comprehension problems.

While extended gaze activities on the ST or TT reflect, thus, text production problems, the question remains whether different translators translating the same text face similar problems at the same text positions. We therefore looked at the amount of ST gaze activity between the typing of every two successive alignment units *n* and *n+1*, and compared the relative amount of ST gaze duration of five translators before typing the translation of a word *n*. The result for the five translators is plotted in Figure 10.

Figure 10: Relative amount of fixation duration (in ‰) before/during production of the translation of the n-th ST word

That is, for each of the 5 translators *T=1...5* we calculate the gaze duration *GT* on the ST window and the gaze duration *GTn* just before and during the production of the translation unit *n*. We then compute for every translator *T* the portion (in ‰) of gaze time for each unit *n* by the ratio *GTn*1000/GT*[12]. For instance, the ST gazing time during the 8.5 sec pause in Figure 9 before and during the translation of word #49 "and" will be allocated to the *GT49*, and appear as a peak in Figure 10.

The graph in figure 10 overlays the proportion of the gaze time that each translator spends on the ST before typing the translation of a word. The horizontal X-axis enumerates all the words of a 110 word ST. The vertical Y-axis shows the relative amount of ST gaze time spent before typing the translation of the ST words. The Y-axis overlays the proportion of ST fixation duration (in ‰) before/during translation production for the five translators. For instance, around ST word #48, one translator spends almost 40‰ of her observed ST gazing time before she produces a translation of the word, while the other translators spent between 5‰ and 14‰ of their gaze time at the same location. The source sentence between ST word number 30 to 56 is represented in Figure 9, while the graph plots the translation progression between words 43 to 52. In Figure 9 we see that much gaze time is spent before "og", the translation of ST word 49 ("and"), is typed. Figure 10 shows that all translators spent substantial time at this position, indicating that they all had presumably similar problems identifying the reference of the verb in the clause starting with "and". Note that while the sentence starts at word position 30, it is only when the conjunction must be translated at word position 49 that all translators have extended reading activities in the ST.

Another example of more than average ST gazing time is around word #19 in Figure 10: all but one translator showed increased reading activity before typing "og", i.e. the translation of ST word #19 "need". Figure 7 shows a translation progression graph of the fragment between ST word #16 and #28. It shows that in some cases all translators spend increased gaze time on the same ST sequences, for other sequences none seems to invest much reading effort and there are still other sequences at which some translators spend much and others no reading time.

---

[12] The amount of gaze time for each unit *i* also normalised (divided) by the number of ST words that this unit contains.

6   Conclusion

We have compared two experimental settings, a copying task and a translation task. In the copying task, a copyist read an English text and typed the same text on a keyboard, while in the translation task the translators produced a Danish translation of an English source text without making use of additional aids like dictionaries or collocation tools. Keystrokes and gaze movements were recorded using the Translog 2006 software. While translators mostly copied and translated the text smoothly, looking only a few words ahead in the ST, we also observed at some instances that the text production triggered extended reading activities, e.g. when the meaning of an idiom was unclear. This behaviour can be observed during translation and when copying a text. Interestingly, the extended reading activity did not occur until the idiom was to be reproduced (typed), and not when it was first read.

We discuss instances of unchallenged text production, which are similar in text copying and in translation (Examples 1 and 3). We also provide examples that show difficulties when formulating (render and address) the translation rather than a ST comprehension problem. Example 2 showed that word meaning is processed also during text copying. A production pause occurred when typing an expression that was difficult to understand, rather than when reading it the first time. Similarly, also during translation a "monitor" triggers extended reading when faced with production problems. That is, translations of a phrase are already typed before the translator exactly knows how to render the remaining part of that phrase. Whereas in most cases the initial translation "guess" is appropriate and the translation could go on, example 9 shows that this is not always the case. We take this as an indicator that translation production (and text copying) may start with a partial (literal) understanding, and that (ST) meaning emerges and consolidates as the translation develops. Two types of behaviour can be distinguished:

- much of the translation drafting is unproblematic and approximately within the time limits predicted for text copying by Johns' TYPIST model[13]. Translators look only a few words ahead into the ST from the position which they are currently translating. Literal translations are produced in such a parallel reading/writing mode decoding and encoding are processed simultaneously. We suspect that the degree of parallel activity depends on experience and typing skills of the translator. A touch typist would more likely exhibit parallel processing behaviour, similar to the one in figure 3, while a translator with less developed typing skills would show more sequential translation patterns.

- at some points in the translation extensive reading behaviour could be observed, signalling more serious translation problems. Depending on the type of problem, it may be necessary for the translator to re-scan the ST or the TT. In both cases, the increased reading activity seems to be triggered by a TT production problem rather than by a ST comprehension problem. That is, in the examples discussed above, we observed that the ST was understood, and meaning hypotheses were generated only to the extent required to keep on producing target text. If, for whatever reason, TT production cannot go on smoothly, and the typing flow is interrupted, the missing information needs to

---

[13] This conclusion is based on our translation material? from English into Danish, two relatively close languages with similar word order.

be retrieved. This may lead to the re-reading of a ST passage with a view to verification or reinterpretation, and/or the revision of the produced TT.

We do not exclude that a translator may re-read a difficult ST fragment several times before starting to translate it, but we did not encounter examples of such behaviour in our data. All behavioural patterns that point to translation problems (i.e. re-reading of ST or TT passages) were triggered by *production problems.* There are more pauses during translation than during copying, indicating more sequential reading/writing processes in translation, while the absence of such pauses indicates more parallel reading and writing activities during copying. This confirms the hypothesis that there are more interventions of the monitor during translation than during copying. Looking at gaze activities we found that the number of ST fixations during parallel unchallenged translation activities is approximately identical to those of unchallenged copying while there were more ST fixations during sequential translation activity.

In a stratificational comprehension-transfer-production theory of translation this behaviour is difficult to explain. If the ST would first have to be completely understood before a translator starts translating it, why would the translation activity have an impact on the ST reading behaviour? Instead, we assume that "Meaning Hypotheses" are constructed to the extent and at the moment they are needed to continue the task at hand. Different meaning hypotheses are required for different kinds of activities, e.g. a technician reading a car repair manual needs a different kind of understanding than a translator translating the text into another language. The reading purpose, thus, determines what kind and depth of meaning representation is required. During translation and text copying, the ST meaning is often only elaborated and tested in the writing process – which leads to the surprising conclusion that comprehension does not precede, but follows text production.

7   References

Angelone, Erik. 2010. "Uncertainty, uncertainty management and metacognitive problem solving in the translation task". In Translation and Cognition, Shreve, Gregory M. and Erik Angelone (eds.), 17–40.

Carl, Michael and Martin Kay. 2011. Gazing and Typing Activities during Translation: A Comparative Study of Translation Units of Professional and Student Translators, Accepted for publication in META

Carl, M. and Jakobsen, A. L. 2010. Towards statistical modelling of translators' activity data. International Journal of Speech Technology, 12(4), 124-146.

Craciunescu, Olivia; Constanza Gerding-Salas, Susan Stringer-O'Keeffe. 2004. "Machine Translation and Computer-Assisted Translation: a New Way of Translating?" *Translation Journal* 8 (3). http://translationjournal.net/journal/29computers.htm

Gile, Daniel. 1995. La Traduction. La comprendre, l'apprendre.   Paris: Presses Universitaires de France

Dragsted, B. 2010. "Coordination of reading and writing processes in translation: An eye on uncharted territory". In G. M. Shreve, and E. Angelone (eds), Translation and Cognition, pp. 41–62, Amsterdam/Philadephia, Benjamins

Jakobsen, Arnt Lykke. 1999. „Logging target text productioin with Translog". In *Probing the Process in Translation: Methods and Results. (Copenhagen Studies in Language 24)*, ed. by Hansen, Gyde, 9-20. Kopenhagen: Samfundslitteratur.

Jakobsen, A. J. and Jensen, K. T. H. 2008. "Eye movement behaviour across four different types of reading task". In *Looking at Eyes. Eye Tracking Studies of Reading and Translation Processing, CSL* 36, ed. by S. Göpferich et al., 103-124.

Jensen, Hvelplund, K. T. 2011. *Allocation of cognitive resources in translation: an eye-tracking and key-logging study*. Ph.d.-afhandling. Copenhagen Business School. Copenhagen.

John, Bonnie E. 1996. Typist: a theory of performance in skilled typing. Hum.-Comput. Interact., 11(4):321–355.

Just, M.A., & Carpenter, P.A. 1984. "Using eye fixations to study reading comprehension". In *New methods in reading comprehension research*, ed. by D.E. Kieras & M.A. Just, 151–182. Hillsdale, NJ: Erlbaum.

Kromann, M. T. 2003. The Danish Dependency Treebank and the DTAG treebank tool. In *Proceedings of the Second Workshop on Treebanks and Linguistic Theories (TLT 2003)*, 217-220. Växjö.

Mossop, Brian. *An Alternative to 'Deverbalization'*. Technical report, York University, 2003. http://www.yorku.ca/brmossop/Deverbalization.htm

Pavlovic, Nataša, and Kristian T. H. Jensen. 2009. "Eye tracking translation directionality". In *Translation Research Projects* 2, eds. Anthony Pym and Alexander Perekrestenko, 93-109. Tarragona: Intercultural Studies Group. Available at: http://isg.urv.es/publicity/isg/publications/trp_2_2009/index.htm.

Perrin, Daniel. 2003. "Progression analysis (PA): investigating writing strategies at the workplace". *Pragmatics* 35, 907–921.

Ruiz, C.; N Paredes, P Macizo, M T Bajo. 2008. "Activation of lexical and syntactic target language properties in translation". *Acta Psychologica* 128 (3), 490-500.

Sjørup, Annette C. Cognitive Effort in Metaphor Translation : An Eye-Tracking Study In *Cognitive Explorations of Translation*, ed. Sharon O'Brien, 197-214. London : Continuum International Publishing Group Ltd.

Staub, A. and Rayner, K. 2007. "Eye movements and on-line comprehension processes". In *The Oxford Handbook of Psycholinguistics*, ed. by G. Gaskell, 327-342. Oxford: Oxford University Press.

Tirkkonen-Condit, Sonja. 2005. "The Monitor Model Revisited: Evidence from Process Research". *META* 50 (2), 405-414.