

Studying Human Translation Behavior with User-Activity Data

Michael Carl¹, Arnt Lykke Jakobsen², and Kristian T.H. Jensen²

¹ Institut für Angewandte Informationsforschung, Saarbrücken, Germany
carl@iai.uni-sb.de

² Copenhagen Business School, Languages & Computational Linguistics,
Frederiksberg, Denmark
alj.isv@cbs.dk, kthj.isv@cbs.dk

Abstract. The paper introduces a new research strategy for the investigation of human translation behavior. While conventional cognitive research methods make use of think aloud protocols (TAP), we introduce and investigate User-Activity Data (UAD). UAD consists of the translator's recorded keystroke and eye-movement behavior, which makes it possible to replay a translation session and to register the subjects' comments on their own behavior during a retrospective interview. UAD has the advantage of being objective and reproducible, and, in contrast to TAP, does not interfere with the translation process. The paper gives the background of this technique and an example on a English-to-Danish translation. Our goal is to elaborate and investigate cognitively grounded basic translation concepts which are materialized and traceable in the UAD and which, in a later stage, will provide the basis for appropriate and targeted help for the translator at a given moment.

1 Introduction

With the technological changes over the past couple of decades, the conception of what constitutes translation has undergone considerable change. The movie and DVD industry, for instance, has created a huge new market for dubbing and subtitling skills. Post-editing of machine-translated text is another new field requiring a new set of skills, which combine traditional source-text-to-target-text translation skills, intralingual rephrasing and original text production skills, and insight into the way an MT system operates. Traditional human translation research has not been much concerned with the technological constraints and requirements in these new forms of translation, but has focussed much on seeing translation as 'functional' text production [19] and interlingual communication [8]. By the contemporary norm, the translator is no longer a neutral mediator, a mere passive reflector of meaning, but is seen as a person responsible for ensuring a loyal representation of source text meaning, of course, but also for guaranteeing readability and comprehensibility of the communication, not just in terms of making sense in a new language, but in terms of being easily accessible by target readers whose knowledge background may be radically different from that of the original text producer. Whenever expert-to-expert communication has to be communicated to non-experts, there is a need for meaning to be reformulated, often both interlingually

and intralingually. The transformations needed from the stage at which a text is circulated among experts until it reaches the end-user are typical of the kind of language skill translators are expected to have. How can the methods of human translation research contribute to modelling this expertise?

In this paper we first outline some traditional research methods in translation studies. In section 3 we introduce a new method for investigating human translation behavior which is based on User-Activity Data (UAD) and retrospective interviews. The goal in these interviews and in the analysis of the data is to link basic translation concepts, as, for instance, processing problems arising through translational divergences, with patterns that can be observed in the UAD. Section 4 gives an example of this.

2 Techniques in Translation Research

Over the past 25 years, human translation research has focussed increasingly on investigating translation processes, either from a cognitive perspective or from the perspective of managing the process in interaction with new technologies and colleagues in increasingly globalised organisations. Here, we will only deal with the study of cognitive translation processes, where a change can be observed from the earlier study of artificially elicited user data (as in think-aloud experiments) to the more recent study of user activity data (eye movements and keystrokes).

2.1 Think-aloud protocols

Using think-aloud as their preferred method of eliciting verbal data and viewing translation as fundamentally a decision-making process (for which the flow-chart was a suggestive analogy), the pioneers of process-oriented research Gerloff [7], Krings [10], and L'orscher [13] succeeded in establishing a complex inventory of meaning operations or strategies performed by translators. In the revised edition of Protocol Analysis, Ericsson & Simon [1] discussed and countered criticisms of their approach, which involves elicitation of data that is concurrent with and quite probably related to the cognitive processes they claim to study, but is nevertheless not a necessary activity. Despite the claims made by Ericsson and Simon, the method potentially skews the primary cognitive activity under scrutiny. Krings [11] found that think-aloud delayed translation by about 25% in his experiments, but did not suggest that the nature of the processing was affected. In the experiments reported by Jakobsen [3] this delaying effect was also documented. More importantly, his experiments also indicated that the think-aloud constraint had a degenerative effect on segmentation. Therefore, at least in translation experiments, the think-aloud condition appears to have a negative effect on processing, and there seems to be a processing price to be paid for verbalisation in terms of additional cognitive load.

2.2 Translation Rhythms

Taking advantage of the fact that in the 1990s most texts and most translations were typed on computer keyboards, software was developed to log the process by which

keystrokes were made in time (ScriptLog, Translog). By this method a complete log could be created of all the keystrokes made in producing a text, including typos, pauses, deletions, changes, mouse clicks, cursor movements, etc. A certain temporal patterning of text production was generally observable and assumed to reflect the cognitive rhythm with which processing takes place. Schilperoord [21] observed hierarchical temporal patterning of pauses between segments in oral dictation of routine letters, declining in duration from paragraph to sentence, from sentence to clause, from clause to phrase, and from phrase to word.

In translation experiments, the patterning is different for mainly two reasons. On the one hand, translators do not need to think about paragraph (or sentence) planning. Most often they can just be taken over from the source text. Secondly, the main obstacle to fluent translation is frequently to do with a local, e.g. semantic, problem occurring unpredictably in terms of structure, but often holding up production longer than the transition time from a sentence or paragraph to the next.

3 User Activity Data

The advantage of working with genuine user activity data (UAD), i.e. eye-movement and keystroke data, is considerable in that we have very direct access to the motor activity which results from the cognitive activity we wish to study. However, one disadvantage about keystrokes is that they are made at the tail-end of the translation or (post)editing process. First there is reading and construction of source text meaning. Then there is mapping of this meaning onto a representation in the target language, and then there is typing of that new representation. What is reflected in the typing activity is the discharge of a segment of information stored in working memory. Optimal human translation would involve that a constant supply of processed ST meaning and TT mapping was fed into working memory at a rate that would allow the translator to type continuously at maximum speed. However, since this situation rarely obtains for intervals longer than about half a minute [4], text production keystrokes tend to be clearly segmented into units reflecting the chunks of meaning that were processed either immediately before the keystrokes were made or starting before but overlapping to some extent with the period of typing.

3.1 Eye-movement Data

Eye movements, by contrast, are involved from the onset of the first reading activity. By definition, touch typists are capable of typing without simultaneously having to look at the keyboard, but even they frequently use their eyes to monitor their typing activity, either by occasionally looking at the keyboard (e.g. for rarely used keystrokes) or by monitoring text production on the computer screen from time to time.

A translator's eye movements give a detailed picture of the complex processing involved in constructing meaning from a string of verbal symbols and representing that meaning in the symbols of a new language. Fundamentally reading progresses from left to right (with left-to-right writing systems) along one line at a time and from the end of

a line to the beginning of the next line down, but reading is by no means a smooth succession of fixations strung together by forward-moving saccades. Kennedy, Pollatsek, Radach, Rayner [15, 18], and many others have shown that the calculation of saccade amplitude is a highly complex process depending not merely on parafoveal perception of word length, but also on parameters like probability of occurrence and familiarity with specific words and concepts. Whenever meaning construction fails temporarily, a regressive saccade moves the eyes back to a previous part of the text for reinspection. Fixations differ greatly both with respect to their duration in time and with respect to the number of times one and the same language item may be fixated. In the Eye-to-IT FET project (FP6 IST 517590), instances of multiple fixations within a word and/or returning refixation(s) of a word were assumed to indicate temporary failure of successful meaning construction or, in the case of translation, failure of successful mapping of constructed meaning onto a target language representation, calling for a prompt to be activated.

3.2 Reading Modalities

It should also be noted that reading while translating is different from reading continuously, e.g. for comprehension. Reading purpose and reading task are factors that strongly influence eye movement behavior. Reading a text for comprehension involves fewer fixations than reading a text out loud, for instance, while reading a text while typing a translation involves perhaps twice as many fixations merely on the source text. Additionally, the eyes also have to attend to the translator's emerging target text. Reading while typing a translation therefore involves constant transitions from the ST to the TT and back. This causes reading to be highly discontinuous and frequently results in several fixations before the original reading point is located.

3.3 Eye-movement and Processing Concepts

The relationship between what the eyes are doing at any given moment in time and what the mind is processing is not as straightforward as was originally assumed by Just and Carpenter [9]. Sometimes the mind is ahead of the eyes and is already processing information represented by a word the eyes have not yet fixated. Sometimes the eyes move ahead so fast that the mind lags behind and has to catch up. Such temporal misalignment may cause an earlier or a later word to be fixated longer even if the processing concerned a neighbouring item. Likewise there are at least three different ways in which the eyes may respond to processing difficulty: they may fixate an item longer, they may move on (and fixate a subsequent word while they wait for the mind to catch up), or they may execute a regressive saccade and refixate words already read. Liversedge & Findlay [12] have proposed to deal with such complexities by means of hybrid eye movement parameters which aggregate fixation patterns across several words. Our proposed analysis will follow the lines of such "regression path analysis" and seek to take the analysis a step further by mapping patterns of eye movement behavior onto processing concepts.

3.4 Basic Processing Concepts

Basic Processing Concepts (BPCs) are defined as major building blocks of actions at the level of mental representations [20]. Embedded in a hierarchical basic concept system, they bind together the functional, executional features of an action and the sensory characteristics that are perceived during action execution. The underlying theory states that actions are represented in functional terms as a combination of action execution and the intended and/or observed effect (see e.g. [2]). Therefore, BPCs can be regarded as cognitive tools for the execution of actions, such as those observed in highly skilled people and professional activities of experts, e.g. complex movement tasks in sports, specialist tasks in various crafts, but also in everyday actions that often also require a level of expertise we are hardly aware of (like driving a car, riding a bike or even tying shoelaces). Within these tasks, BPCs serve the purpose of reducing the degrees of freedom involved in action execution and thereby the cognitive effort necessary for controlling the action.

In long-term memory, a given task is built upon a hierarchical structure of BPCs that reflects the task experience and the level of performance of the individual and thereby the perceptual and cognitive content that can be linked to partial behavior and subtasks. In this respect, BPCs can be regarded as basic units of knowledge about the world ("Weltwissen"), accessible for mental control in volitional acts, but also as sensorimotor representations of movement effects. The number and granularity of BPCs in a given task depend on the task itself, on the level of expertise of the individual and on the way the task has been learned and trained. To find a suitable set of BPCs for a given task, it is necessary to observe the behavior thoroughly and to break it down into parts that relate perceptual input to cognitive content, and that can be labelled verbally and/or pictorially. It is therefore hardly possible to define BPCs without extensive feedback from and cooperation of subjects who perform the task at a sufficiently expertly level.

4 UAD in Translog

In this section we provide an example to illustrate our research strategy. A small text on politics of 125 words was to be translated from English into Danish, using the Translog program (www.translog.dk) in the version that logs eye movements as well as keystrokes. In the experiment, Translog separated the screen into two windows: the (English) source text was shown in the upper window. Subjects were asked to type a translation into the lower window as shown in figures 1 and 2.

An important feature of Translog is that registered UAD stored in a log file can be shown in a replay session, after the registration phase. While user activities are dynamically visualised, subjects can comment of their own gaze and keystroke behavior. Thus, keyboard activities and also the user's successive eye fixations can be replayed in real-time fashion. A screen shot of such an instance is shown in figure 1 and 2. In this way we can observe and study temporal patterns of eye-movement behavior and correlate these to properties of the source text, as well as to rhythms in text production. The replay tool is described in [6]. It allows the user to register retrospective interviews and associate BPC with the UAD. In the next subsection we will outline an experimental

setting to investigate the relation between textual properties and eye-movement behavior of translators. In future experiments we intend to take into account text production rhythms and BPC.

4.1 An Example

In a series of translation experiments three subjects were asked to translate several texts from English into Danish. One of those texts is shown below. Figures 1 and 2 represent accumulations of fixation points during the time span in which one subject starts reading a source language sentence and begins producing (i.e. typing in) a translation. We call such time intervals “translation pauses” since no keystrokes are observed. However, as pointed out previously, the mind is very active during those “pauses”, since the translator tries to understand (a fragment of) the source text and develops a translation strategy.

Figures 1 and 2 plot eye-tracking data of sentence initial translation pauses for the first clauses in the third and fourth sentence of the text. These segments are marked in **bold** in the text below:

In a gesture sure to rattle the Chinese Government, Steven Spielberg pulled out of the Beijing Olympics to protest against China’s backing for Sudan’s policy in Darfur. His withdrawal comes in the wake of fighting flaring up again in Darfur and is set to embarrass China, which has sought to halt the negative fallout from having close ties to the Sudanese government. **China, which has extensive investments in the Sudanese oil industry**, maintains close links with the Government, which includes one minister charged with crimes against humanity by the International Criminal Court in The Hague. **Although emphasizing that Khartoum bears the bulk of the responsibility for these ongoing atrocities**, Spielberg maintains that the international community, and particularly China, should do more to end the suffering.

The two segments consist of 10 and 14 words with 67 and 100 characters respectively, which amounts to 8.3% and 12.3% of the total characters and 8% and 11.2% of the total words in the text. Table 1 summarises the properties of the text and the two segments.

Table 1. Absolute and relative length of the text and the segments 1 and 2

Parameter	#words	#chars	% words	% chars
text	125	812	100.0%	100.0%
segment 1	10	67	8.0%	8.3%
segment 2	14	100	12.3%	11.2%

Both clauses have different degrees of difficulty. The difficulty when translating segment 2 into Danish is due to the subordinate clause which needs a subject and a

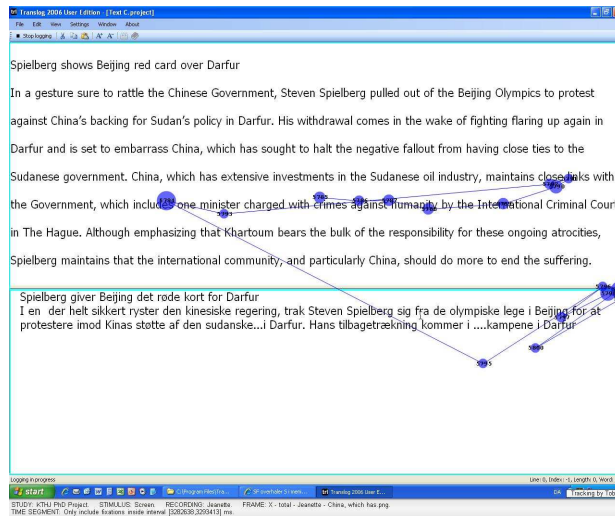


Fig. 1. The figure shows the number and durations of gaze fixation points accumulated during 10 seconds of translation pause when starting to read the third English sentence in the upper window. At this time the subject has already translated the beginning of the text into Danish in the lower window.

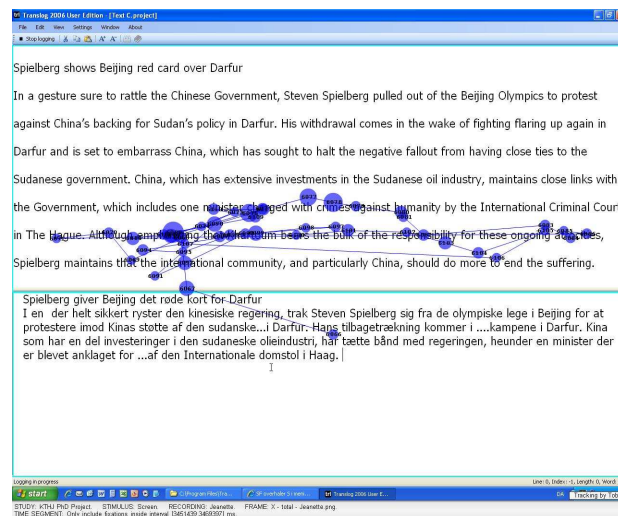


Fig. 2. The figure shows the number and durations of gaze fixations accumulated during the 18 seconds when starting to read the first segment in the fourth English sentence and typing in a translation. Note that, due to inaccuracies of the technical and natural devices, the recorded fixation points are not always above the words actually looked at. Figures are screen-shots taken with the Tobii software (www.tobii.se).

finite verb, as e.g. "Although *he emphasizes* that ...". More planning and restructuring of the translation is necessary than for segment 1.

Two of the translators (Transl₁ and Transl₃) were students, and Transl₂ was a professional translator. The students needed 493 sec. and 481 sec. to translate the entire text while the professional translator needed 61% of that time (303 sec.). Translator Transl₁ was a touch typist looking more than 65% of the time (324 secs.) on the screen, while the other two translators spent only 31% and 34% of their gazes on the screen, presumably the other two thirds on the keyboard. Accordingly, there are many more recorded fixations for Transl₁ (1327) than for the other two translators; and the lowest figure for the professional translator Transl₂ (418 fixations). However, the average fixation duration is relatively uniform among all translators. These figures are shown in table 2, also broken down for the two segment translations.

Table 2. UAD from three translators for the entire text and for clause 1 and clause 2, without postediting time.

	Transl ₁	Transl ₂	Transl ₃	entity
Figures of UAD for entire text:				
Translation Time	493	303	481	seconds
Gaze Time	324	95	169	seconds
Not Watching the Screen	169	208	311	seconds
Number of Fixations	1327	418	816	
Average Fixation Duration	0.244	0.227	0.208	seconds
Figures of UAD for Segments 1 and 2:				
Translation Time seg. 1	43	30	50	seconds
Translation Time seg. 2	77	42	56	seconds
Gaze Time seg. 1	30	11	17	seconds
Gaze Time seg. 2	54	14	23	seconds
Number Fixations seg. 1	121	45	69	
Number Fixations seg. 2	213	200	179	
Average Fixation Duration seg. 1	0.247	0.236	0.256	seconds
Average Fixation Duration seg. 2	0.213	0.200	0.179	seconds

For one of the translators (Transl₃), the UAD of the "translation pauses" for segments 1 and 2 are plotted in figures 1 and 2. We registered 10 seconds of translation pause for segment 1 and 18 seconds for segment 2, with the eyes moving several times back to the subordinate clause. Figure 2 shows that most of the eye fixations are located in the subordinate clause, indicating a difficulty in developing a translation strategy for this sentence. Once a decision was taken on how to translate the sentence and what form of subject should be introduced, the translation was written with only occasional reference to the source sentence.

The relatively long time span for reading and understanding followed by fluent and fairly quick production of the translation indicates that the entire sentence, consisting of 30 words, was constructed as one meaning unit. While translators often proceed chunk-wise, starting to translate a sentence without even having read it completely, this

sentence needed to be thoroughly scanned and understood before the translation could be started.

Table 3. Relative change in UAD for segments 1 and 2 compared to UAD for the entire text. The figures include translation pauses plus production time of the clause translation.

Parameter	Segment 1	Segment 2
Increase in Translation Time	16.4%	11.4%
Increase in Gaze Time	19.7%	26.4%
Increase in Number of Fixations	11.2%	44.3%
Increase in Average Fixation Duration	7.6%	-12.4%

To figure out an appropriate translation strategy for the second English segment, the three translators spent respectively 19.1%, 17.5% and 15.8% of all the fixations on the 14-word segment, while the first 10 word segment absorbed 9.1%, 10.8% and 8.5% of the fixations. The relative change of parameters, averaged over the three translators for the two text segments, and compared to the entire text is summarised in table 3. The relative number of fixation increase (+44.3%) in the structurally difficult segment 2 is presumably due to the syntactic reorganization that had to be processed. Despite the decrease in fixation length (-12.4%), the overall gaze time spent on the difficult passage was much longer (+26.4%), relatively, than that spent on the rest of the text and on segment 1 (+19.7%).

5 Conclusions

In this paper we have introduced a new cognitive research method, and described a tool for studying human translation behavior. Patterns of User Activity Data (UAD) such as eye-movement and keystroke behavior are associated with properties of the text.

In a first step we aim at investigating the impact of translation divergences on the UAD, i.e. whether and how average gaze behavior changes in different contexts, increases or decreases and how translation behavior changes in different conditions. We seek to detect patterns of translation behavior in the UAD and associate textual properties with it.

In a second step we intend to link basic translation concepts i.e. major building blocks of mental representation, with UAD and thus detect factors which contribute to the problems which translators face during their work.

We hope that the method can be sufficiently formalised so that it will eventually lead to machine-mediated processing, where programs assist translators in their tasks through the knowledge and emulation of human cognitive processes. The aim is to develop translation devices and translation help which intelligently interact with the translator or with a posteditor. This, we believe, can be achieved by knowing and formalising the basic processing concepts which translators implicitly possess and process when doing their work.

References

1. Ericsson, K.-A. & Simon, H.: Protocol Analysis: Verbal Reports as Data. Cambridge, Mass.: MIT Press. (1984) (2nd revised edition 1993).
2. Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W.: The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24, 849-878.
3. Jakobsen, A. L. : Effects of think aloud on translation speed, revision, and segmentation. In *Triangulating Translation. Perspectives in Process Oriented Research* (ed.) Fabio Alves. Amsterdam: Benjamins (2003), 69-95.
4. Jakobsen, A. L. : Investigating expert translators processing knowledge. In Dam, Helle V., Jan Engberg, Heidrun Gerzymisch-Arbogast, eds. (2005) *Knowledge Systems and Translation (Text, Translation, and Computational Processing 7)*, Berlin, New York: Mouton de Gruyter, (2005) 173-189.
5. Jakobsen, A. L. and K. T. H. Jensen: Coordination of comprehension and text production in written and oral translation tasks. *AMLAP 2007*.
6. Carl, M. & Jakobsen, A. L. and Spakov, O. : Towards an Annotation Tool for Eye Tracking Data, forthcoming (2008)
7. Gerloff, P.: Second Language Learners Reports on the Interpretive Process: Talk-aloud Protocols of Translation . In House, Juliane and Shoshana Blum-Kulka (eds.) *Interlingual and Intercultural Communication. Discourse and Cognition in Translation and Second Language Acquisition Studies*. Tübingen: Gunter Narr, (1986), 243-262.
8. Hatim, B. & Mason, I. : *The Translator as Communicator*. London & New York: Routledge. (1997)
9. Just, M.A. & Carpenter, P.A.(1980): A theory of reading from eye movements to comprehension. *Psychol. Rev.* 87, (1980) 329–354
10. Krings, H. : *Was in den Köpfen von Übersetzern vorgeht*. Tübingen: Gunter Narr (1986).
11. Krings, H., translated and edited by Koby, G.S.: *Repairing texts: empirical investigations of machine translation post-editing processes*. Kent State UP, (2001), Ohio, USA
12. Liversedge S.P & Findlay J.M. Saccadic eye movements and cognition. *TICS* 4 (1), (2000), 6-14
13. Lörcher, W.: *Translation Performance, Translation Process and Translation Strategies. A Psycholinguistic Investigation*. Tübingen: Gunter Narr (1991).
14. Pickering, M.J. & Traxler, M.J. (1998) Plausibility and recovery from Garden paths: an eye tracking study. *J. Exp. Psychol.* 24, 940-961
15. Radach, R., Kennedy, A. & Rayner, K. *Eye movements and information processing during reading*. Hove, (2004) East Sussex: Psychology Press.
16. Rayner K. (1975) The perceptual span and peripheral cues in reading. *Cognit. Psychol.* 7, 65-81
17. Rayner K. & McConkie G.W. (1976) What guides a reader's eye movements? *Vis. Res.* 16, 829-837
18. Rayner, K., Pollatsek, A. 1989. *The Psychology of Reading*. Englewood Cliffs: Prentice Hall. (1989)
19. Reiss, K. & Vermeer, H. J. : *Grundlegung einer allgemeinen Translationstheorie*. Tübingen: Niemeyer, (1984).
20. Schack, T. : The cognitive architecture of complex movement. *International Journal of Sport and Exercise Psychology; Special Issue Part II: The construction of action - new Perspectives in Movement Science*, 2 (4), (2004) 403-438.
21. Schilperoord, J.: *It's about Time. Temporal Aspects of Cognitive Processes in Text Production*. Amsterdam: Rodopi, (1996).