

Eye tracking in translation process research: methodological challenges and solutions

Sharon O'Brien

Abstract

Eye tracking has been eagerly adopted as a technique in translation process research in recent years and, with it, comes a long list of methodological challenges, some of which are specific to translation process research, some of which are more general. This paper represents an attempt by a recent convert to eye tracking to document the various challenges and to offer suggestions for how some of them might be addressed by others who are interested in embracing this very interesting mode of investigation. It is difficult to comment only on those challenges related to eye tracking without mentioning issues that pertain to the more general questions of research design in the domain of translation process research as the two are inevitably linked. By highlighting challenges, one is inevitably exposing research design weaknesses. However, this should not be viewed negatively but should rather be seen as a means of improving the quality of combined research outputs with the aim of maturing the domain.

This paper is broadly divided into two parts: the first discusses the methodological challenges, which in turn are divided into the categories of research environment, research participants, ethics, data explosion and validity. Each section is then repeated in the second part of the paper, which offers suggestions on how the challenges might be addressed.

1. Methodological challenges

1.1 Research environment

“Research environment” refers to where the research takes place, the conditions in which it takes place and the equipment used. The challenges

presented by using an eye tracker in translation process research are discussed here under the headings of equipment, accommodation and familiarity.

Equipment

At the time of writing, the Tobii brand of eye tracker (1750, T60 or T120 models) was the most popular model in Translation Studies. This eye tracker is a desktop eye tracker with an in-built camera and infrared diodes. Its advantage is that it resembles a normal computer monitor and is therefore quite “uninvasive” for research participants, who are physically unaware that their eyes are being tracked (the participants would of course be made aware in advance that their eye movements were being tracked, see “Ethics” below). This type of eye tracker increases the ecological validity for the type of research projects conducted in Translation Studies and it is generally deemed to be more advantageous than the head-mounted eye tracker¹ because it does not burden the participant with physical equipment on the body. However, there is a trade-off between accuracy and validity here since head-mounted eye trackers are generally considered to be more accurate than desktop ones.²

Some of the challenges associated with using an eye tracker include (i) cost, (ii) rapid technological redundancy and (iii) learning curve. To date some translation research centres have managed to acquire funding to purchase an eye tracker. However, this technology is relatively expensive and funding is difficult to attain. The analysis software, which is used to analyse the data generated during eye-tracking sessions, may be frequently updated or, indeed, completely redesigned and there is sometimes a lack of backwards compatibility, meaning that eye-tracking sessions recorded in new versions cannot be viewed or analysed in older versions and vice

¹ Examples of current head-mounted eye trackers are the Arrington Research Mounted Eye Tracking System (<http://www.arringtonresearch.com/headmountframe.html>) or the ASL EYETRAC@6 head mounted optics (<http://www.a-s-l.com/site/Products/EYETRAC6Series/HeadMounted/tabid/57/Default.aspx>) [Last Accessed 12/05/09].

² For example, the Tobii 120 model can collect data at a rate of 120 Hz whereas the ASL head-mount model can collect data at 120, 240 and 360 Hz, according to the technical specifications on each company's website.

versa. A steep learning curve can also be associated with the use of the analysis software.

Accommodation

Once an eye tracker has been acquired, it needs to be housed somewhere. Eye trackers do not require vast amounts of space: a normal desk will suffice. The challenge is to find a space which meets specific requirements for conducting eye-tracking research. In particular, consideration must be given to light, sound and familiarity of the surroundings. If, as is the case for many translation process researchers, one is interested in measuring cognitive effort, then recording pupil dilations will be of particular interest. However, pupil dilation is known to be influenced by many factors, such as changes in light intensity, sound, caffeine, drugs, substance abuse, eye colour and even heavy eye make-up. It is highly recommended, therefore, that the lighting in the room where the eye tracking research takes place is controlled. At a minimum, blackout blinds and a consistent source of lighting are recommended. It has also been shown that the pupil reacts to sound so it is logical to control for this. Since soundproofing a room is beyond the means of most translation research groups, one should attempt to house the eye tracker in a quiet location and/or to conduct sessions when noise interruptions are least likely. Other factors, such as caffeine levels, substance abuse, eye make-up etc. are more difficult to control but participants can be made aware that these factors can influence measurements and can be asked to avoid caffeine prior to sessions or not to self-select if they have a record of substance abuse, for example.

Familiarity

Translation process studies, including those using eye-tracking equipment, frequently use professional translators as “research participants” in a bid to uncover professional work practices, strategies and differences between professionals and novices, among other things. This raises a challenge for ecological validity: we wish to observe what professional translators “normally” do, but we remove them from their “normal” work environments in order to do so. The fact that the eye-tracking monitor is most likely different in shape and size from their usual monitor, or that the operating system, software, version numbers, language packs, screen layout

or even keyboard type differ from their usual work environment may have an impact on their performance. The research community should not abandon research because of these challenges, but, where a “normal work environment” is an important factor, the eye-tracking environment ought to be set up in such a way that the participant is familiar and comfortable with it.

1.2 Research participants

Translation process research needs translators in order to observe, record and theorise the translation process. As mentioned, some researchers are interested in professional translators while others are interested in students of translation or in the differences between professionals and students. This focus on professionals presents many challenges, in general, as well as some specific challenges for eye tracking. We will discuss the general challenges first followed by the challenges relating to eye tracking.

Terminological considerations

Different terms are used in translation process research to designate people who earn their living by translating and people who are learning to do so. The first category is typically termed “professional”, a wide variety of names has been attributed to the second category, e.g. semi-professional, novice, student. Although the term “professional” is used consistently to refer to somebody who is a practising translator and not a student, the scope of the term varies. A professional could be a person who has more than 10 years of experience working full-time as a freelance translator or it could mean a graduate who has worked for less than one year in a translation agency. Therefore, when we make claims about how “professionals” operate, we are making claims about people with very different levels of experience. It should be acknowledged that even if the behaviour of two professional translators each with 10 years of experience was compared, it would not be surprising to see evidence of quite different behaviour. We are, after all, observing *humans*. Therefore, the research domain would benefit from agreeing criteria for the term “professional” so that we can make valid cross-study comparisons.

As mentioned, the term used for people who are studying to become translators also varies. The validity of the term “semi-professional” to

denote a final-year translation student is debatable. The same is true of the term “novice”. Until the research community can agree appropriate terms and criteria for assigning those terms, we ought to simply describe the profiles of our participants in detail. So, for example, if participants are post-graduate MA students in Translation Studies who are in the second and final semester of their training programme, then these are the exact details we should provide when describing our research participants.

Numbers

Since translation process research relies on human participants, one of the challenges is including an adequate number of participants who fit the specific profile we are hoping to investigate in order to enable researchers to make generalisable claims. If we wish to include professional translators in a study, then funding is required to pay such participants. Some professional translators will give freely and generously of their time. However, this raises questions about ecological validity: if somebody is donating their time freely in the name of research, one cannot be certain that they will behave “normally”, i.e. as one does in everyday translation work. Admittedly, it is also difficult to prove “normal behaviour” even when a person is being paid for their participation. However, it is reasonable to expect that they would behave normally if they are being paid. A certain aspect of the “normality” is of course compromised because the translator knows that the researcher is not a “real” client and that they have little to lose if they do not perform well. This is a methodological concern which has no easy solution.

Given limited funding, it is difficult to recruit adequate numbers of participants to make generalisable claims. Some studies have included five subjects, others ten, others 20. For example, in the research reported in a recent volume of the *Copenhagen Studies in Language* (Göpferich *et al.* 2008), the average number of participants in eye-tracking studies of translation was 12. Making valid generalisations on the basis of such numbers is questionable. Nonetheless, these studies have significant value. First, they are pioneering in their use of eye tracking in translation studies. Researchers are testing equipment, research designs, and methodologies. Secondly, such studies are valuable for generating hypotheses using small communities which can then be tested on larger communities.

Recruiting student translators is somewhat easier than professionals; students tend to be keen on learning about the research, methods and technology. Some reward for participating can be offered in the form of course credits (though this is not possible in all educational institutions) or in the form of books or book tokens, for example. However, the ethical considerations of offering rewards for participating in research studies also need to be considered (see "Ethics" for more discussion on this). Although student translators are easier to recruit than professionals, as a research community we should nonetheless be wary of applying the "professional" label to final-year translation students. We also need to acknowledge that, while we aspire to graduating translators who are ready to work in the professional translation domain, they are still at the bottom of the professional ladder when they graduate and have much to learn in the workplace.

Competence

Let us assume for one moment that research funding is abundant and we can afford to recruit, say, 50 professional translators and 50 student translators for a comparative study of translation strategies, cognitive load or translation competence. Let us also assume that we have agreed rigid definitions for our categories of "professional" (e.g. must have worked full-time as a translator in a specific specialised field for a minimum of 10 years) and "student" (must be in their second semester, at post-graduate level, with an average score in economic translation of 65 % and no fails in any subject). Can we assume that the 50 professional translators have equal competence or that the 50 students have equal competence? Our experience as translator trainers allows us to reasonably expect diverse levels of competence in different aspects of translation (e.g. SL comprehension, TL production, specialised domain knowledge, etc.) within each of these groups. The first challenge for translation process researchers is to *acknowledge* this diversity. We cannot claim that our participants are all equal. At the same time, having diversity within one group does not invalidate our studies. By having greater numbers, we can both expose and explore the diverse behaviour and competences and look for commonalities among the members of specific groups. The commonalities then allow us to make general claims about the nature of the translation process.

When it comes to using eye trackers, this diversity is all the more apparent. Even if we can build a group of research participants using tight control criteria, we are still faced with individual differences that have an impact on the study. Participants may have individual eye conditions, which reduce the quality of the data collected by the eye tracker. For example, it is claimed by the manufacturers of eye-tracking equipment that spectacles and contact lenses do not hinder eye tracking. In reality, the quality of eye data can be better for one person when wearing spectacles and for another when wearing contact lenses. Dark-coloured eyes may also have an impact on the ability of the eye tracker to differentiate the pupil from the iris. On some occasions, the quality of eye-gaze data might be so low that participants' data cannot be used in a study.

When is a suitable participant not suitable?

Even though a researcher may implement strict screening for the recruitment of individuals into a research study, it is sometimes the case that other issues arise which result in having to remove that person from the study. Such issues may include skills such as touch typing, language competence, ability to follow instructions, and propensity towards a “white coat effect”. Each of these will be discussed briefly.

An important quality when participating in eye-tracking studies is the ability to look at the screen. For text production studies, for example, the requirement is to look at the screen and not at the keyboard, so touch-typing is an essential skill. However, few people learn to touch type in the true sense of the term these days and, even if participants self-select on the basis that they can touch type, the researcher may well find that they spend a significant proportion of time looking at the keyboard. If this amount of time is significantly high, this participant's data cannot be used in the study.³

³ What qualifies as “significantly high” is not yet agreed among translation-process researchers. The threshold set for data quality will depend on many factors, such as the general objective of the study, how many participants one has, other measurement criteria being used, etc. As a very general and somewhat simplistic guideline, if less than 70 % of the time is spent looking at the screen, the researcher might want to consider removing that participant from the study. However, this could be seen as a lowest limit threshold.

Language competence is one of the basic criteria for selecting participants. Being a “native speaker” of the target language, for example, might be important in a research design. Participants may self-select for a study on the basis that they are native speakers of a particular language, but it can then transpire that they are bilingual, with one language more dominant than the other which may have an impact on their translation processes and product. It is therefore important to check with participants that they really meet the criteria specified for language competence.

Studies by Jääskeläinen (1987) and Tirkkonen-Condit (1989) demonstrated that the translation assignment, i.e. instructions to the subjects on what the translation was *for*, strongly influences the process. Therefore, it is important to consider the phrasing of such instructions. A frequent requirement in translation process research is that the participants behave as they would normally if a client had commissioned the translation. This requires the participants to suspend belief and pretend that the research assignment is “real”. A second challenge has to do with the ability (or the inclination) to obey instructions about the experiment itself. In particular when using an eye tracker, participants may be instructed to sit at a certain distance from the screen, not to move their heads too much, to press specific keys at certain points in the process, not to use hardback dictionaries, or not to interrupt their flow or talk to the researcher until they are finished. While these instructions are all relatively simple, some participants are better at following them than others.

Given that translation process experiments can be somewhat artificial in nature, it is important to put participants at their ease and to make sure they do not feel intimidated, judged, or in any way threatened. Ways of addressing these concerns include anonymising the data and informing participants that the study centres on *how* translation is done and not necessarily on *how good* the end product is (although product quality can be a feature of some studies). By conducting a pre-eye tracking briefing or interview and a warm-up session, participants can also be made to feel at ease. However, it is still the case that some participants are highly sensitive to being studied and a “white coat effect” ensues where they find themselves behaving differently to how they would normally behave. Sometimes participants will voluntarily inform the researcher of this after the session. However, the effect is sometimes undiscovered until some of

that participant's data has been analysed and it becomes clear that something went wrong during the session. Ironically, in this author's experience, this happens less with students and more with professionals.

1.3 Ethics

It is more and more the case that researchers wishing to recruit human participants for research purposes, even in Translation Studies, are required to obtain ethical approval from their organisation's ethics committee. While this might initially seem unnecessary, it is a very important step in research preparation. The ethics approval procedure requires the researcher to state in a comprehensible way the purpose of the research and precisely what is required from participants. A researcher could be tempted not to inform participants that their eyes are being tracked, in order to stave off the "white coat effect" mentioned above, but this would be considered unethical by many.

The ethics approval documentation also sets out participants' rights, i.e. (usually) that they will not be named, that their data will be protected and stored securely and that they can withdraw from the study at any stage without penalty. The latter right is of special importance for the two groups of participants that are commonly used in translation process research. In the case of student translators, especially when the researcher is also a teacher of the participant, it is important for the participant to know that removing themselves from the study will have no impact on other student activities.

Professional translators may also wish to remove themselves from a study for a number of reasons. If the professional's relationship to the researcher is one-to-one, this is not such a problem (except for the researcher who has lost a precious participant). However, this raises an interesting dilemma if the professional translator is a sub-contractor to an industrial sponsor of the research and the industrial sponsor is paying for the professional's time. Under normal working conditions, there might be a penalty if translators removed themselves from a commercial project prior to its completion. On the other hand, when that "project" is research being paid for commercially and the researcher is obliged to say that the participant can withdraw at any stage without penalty, then there is clearly

a contradiction in terms. There is no easy solution, but it is advisable to acknowledge and discuss these eventualities upfront in such circumstances.

A general ethical question should be addressed here also: translation process research effectively seeks to understand how translators work, what practices are “successful”, what counts as “efficient”. These represent the concerns of those who pay for translation services too. If our research exposes individuals as being incompetent, unprofessional, inefficient, etc., then we are arguably putting individual careers at risk. This is especially the case when the research is part-funded by industrial partners. It is therefore of the utmost importance to discuss these eventualities with all stakeholders prior to the commencement of a study and to ensure that there is no negative impact on individual livelihoods through our research activities.

1.4 Data explosion

We mentioned earlier that an important design challenge in translation process research is the small number of subjects and the resulting inability to generalise research findings. Funding for participants and ability to recruit participants of similar competence were mentioned as two impediments to working with larger numbers. The nature and extent of the data generated, especially when using eye tracking and keyboard logging, represents another challenge.

Translation process researchers have increasingly subscribed to recommendations on research methodologies expounded by, e.g., Frey *et al.* (1991: 124):

Measurement validity and reliability can be increased by combining quantitative and qualitative measuring procedures in the same research study, a practice referred to as triangulation.

Triangulation has been recommended and, indeed, adopted by many in the field (e.g. Krings 2001, Alves 2003, O'Brien 2005, etc.). Triangulation involves a weaving of results obtained using different tools and methods – some quantitative, some qualitative – such as interviews, questionnaires, think-aloud protocols, quality assessment, keyboard logging and eye tracking. Each one of these techniques produces a large data set which has to be recorded, transcribed, coded and analysed. Eye tracking alone

produces a veritable sea of data: there are gaze plots showing the number and sequence of fixations, hotspots showing the areas on the screen that were most frequently fixated, video files showing the eye gaze, reading and text production data which are sometimes overlaid with concurrent or retrospective protocols, and each eye-tracking recording also produces a very large data file with millisecond-based data on the position of the eyes according to the X, Y coordinates of the screen, the left and right pupil sizes, the validity of the data at any point in time, the fixation number, etc. If a keyboard-logging tool such as Translog is used in conjunction with the eye tracker, then researchers also have at their disposal a Translog replay file and a log file containing information about all the keyboarding and pause activity. All of this is produced per person, per recording. On the positive side, researchers have a rich set of data at their disposal from which conclusions can be drawn. However, the effect on the lone researcher, even with a small number of participants, can be overwhelming. How do we scale this up to include larger numbers of participants? While eye-tracking analysis programs such as Tobii's ClearView or Studio automate some aspects of the data analysis, it does not automate everything (e.g. pupil dilation calculations are still done "manually", i.e. the researcher has to figure out a way of slicing data and creating macros in a program such as MS Excel). Suggestions on how to manage the data explosion are given in "Addressing the Challenges".

1.5 Validity

Many of the above-mentioned challenges impact on the validity of research design. For translation process studies, one other issue deserves mention: what participants are translating and how. The capacity of one researcher to thoroughly analyse the amount of data that translation process methods can produce is limited. As a result, we tend to choose quite short texts for our research studies. The reasons for 200–300 word texts are numerous and valid. For example, participants could get tired or bored quickly, leading to a drop in motivation and an effect on the data; or, for eye tracking, scrolling down on a screen is problematic and therefore texts are selected so that the source and target can fit on the screen at the same time. If the researcher is interested in word fixations, the font size has to be big (16 or 18 font size,

for example) and this makes it impossible to select long texts. The challenge for translation process studies lies in the fact that, with the exception of the localisation industry, translators normally work with larger chunks of text and, therefore, one could argue that we are not investigating “real scenarios”. The validity of our research is, therefore, compromised.

Secondly, how participants are translating texts raises concerns about validity. Translators might spend most of their day working in MS Word, or MS Excel or in a particular Translation Memory environment. They use the World-Wide Web for research purposes, but also resort to their customised glossary and their hardback dictionary. They then volunteer for a research study and are asked to translate, without using the Web, a dictionary, or their own glossary, a text type they normally do not work with in a tool they have never encountered before. It is reasonable to expect that this will impact on their behaviour and we, yet again, are faced with the issue of researching non-standard behaviour while assuming it is standard. Translation process research does not find itself alone in this dilemma, but it is one that has been largely unacknowledged to date and which needs some consideration and controls.

A third challenge concerns the nature of the texts under study. Researchers who also translate or teach translation are not unfamiliar with the fact that the so-called “general” text can sometimes pose even more difficulties for the translator than the specialised text. At the same time, few translators earn their living by translating general texts or texts published in newspapers and, sometimes, the unexpected challenge of the apparently easy general text presented in a research study can come as a surprise to the highly specialised translator. Using unfamiliar text types or domains to investigate processes is valid if we are testing how participants cope with new or unfamiliar challenges. However, if we want to investigate what professionals normally do in their usual working environment, then we need to select texts that will help us answer that question.

2. Addressing the challenges

In this section we make some proposals for how at least some of the methodological challenges outlined above can be met.

2.1 Research environment

The setting up of a dedicated eye-tracking lab with control for light and sound is to be recommended. If a “normal” work environment is important in the research design, the lab should be made to look and feel like a standard office and not a lab.

The question of increased accuracy (using head-mounted eye trackers) over ecological validity (using desktop eye trackers) has received little consideration in translation process research to date. The tacit agreement, it seems, is that the desktop eye tracker, if combined with other measurement techniques such as verbal protocols, interviews, keyboard logging, etc., is accurate enough for the current aims.

Knowing one’s participants prior to commencement of the study is to be recommended. If the study is concerned with what translators do in their professional environment, then having information about what type of PC, operating system, software, versions and keyboards are commonly used is worthwhile, and efforts should be made to replicate the usual working scenario. Obviously, this can only be done to a certain extent, as it is not reasonable to present each participant with an exact replica of their own working environment. But knowing in advance, for example, that a participant has only ever used an “azerty” keyboard and is not used to a “qwerty” keyboard can save the researcher from losing a participant due to poor data quality caused by having to search for keys on the keyboard.

2.2 Research participants

It is important to decide on specific definitions and criteria for the participant community prior to commencement of the study. If the study is a hypothesis-generating one, then low numbers are acceptable. However, if the researcher wishes to generalise to the greater translation population community, then a larger number of participants should be included. Nevertheless, the ability to analyse the data within the specified project time frame also needs to be considered here.

It is reasonable to assume that there will be a 30 % drop-out rate (approximately) due to a lack of suitability (physical, competence, white coat effect, etc.). If specific skills are important, e.g. touch typing, test in advance. The warm-up task cannot be used for this purpose, as there is

insufficient time to analyse the data if it truly is a “warm-up” task administered immediately prior to the main research task. If participants are expected to produce rich concurrent or retrospective protocol data, then it is advisable to train them in this skill in advance as it is not reasonable to expect them to produce a rich protocol if they have not done so before.

2.3 Ethics

Serious consideration should be given to the ethical dimension of the research and to the potential implications for the participants, the researcher and the sponsor (if there is one).

2.4 Data explosion

Research “teams” are still an unusual concept in Translation Studies, but this might be the solution to the data explosion resulting from the triangulation of methods. Several researchers could analyse data produced using different methods or, indeed, the same data could be analysed by different researchers from different angles (as is the case with the TransComp project currently ongoing at the Karl Franzens Universität in Graz, Austria⁴).

Statistical analysis of eye-tracking data is usually desirable. For example, reporting that two averages are different is not very informative unless the statistical significance of that difference is calculated. Unfortunately, training in statistics is not common in Translation Studies and researchers often find themselves struggling with this requirement. If feasible, consult with a statistician during the study design and after data collection. If the research is funded, putting funds aside for consulting (or even hiring) a statistician is advisable.

2.5 Validity

Some research questions can be adequately investigated at the sentence or sub-sentential level. However, where the object of enquiry is the text, then

⁴ <http://gams.uni-graz.at/fedora/get/container:tc/bdef:Container/get> [Last accessed 15/05/09].

valid examples of “texts” should be used in order to meet validity requirements.

The appropriateness of the text length should be considered. Since this is particularly problematic for studies involving eye tracking, triangulation could be considered as a way of compensating for the scrolling and time limitations imposed by eye-tracking equipment.

If text type and/or complexity is a factor in the research design, consideration must again be given to how representative the selected texts are of particular text types and what level of complexity they represent. Texts could be selected on the basis of research which profiles the linguistic make-up of specific text types using large corpora (e.g. Biber *et al.*'s work from 1998). Text complexity, on the other hand, should be examined according to strict and reliable criteria (the reliability of readability formulae for this task is open to debate and rhetorical structure theory (Taboada and Mann 2006) has been proposed as an adequate way of measuring text complexity by Alves *et al.* (this volume). Also, Jensen (2009) provides a helpful distinction between text *difficulty* and *complexity* and discusses the relevance of readability indices in measuring text complexity in translation process research, pointing out that readability indices only take surface structure into account and ignore semantics.

As with all research design, the priming effects caused by the sequence of text presentation on participants must also be taken into account.

3. Summary and conclusions

The aim of this paper was to address methodological challenges faced by translation process researchers especially when they wish to embrace eye tracking as one of the measurement methods. We have highlighted challenges pertaining to the research environment, research participants, ethics, data and validity and we have attempted to suggest some ways of tackling these challenges. We hope that this discussion will be of use to those already engaged in eye-tracking research but, more specifically, that those who have not yet employed this method but are hoping to do so in the future will find the discussion useful and will be in a position to better manage some of the challenges highlighted. While eye tracking does not

reveal all there is to know about how humans translate, it certainly adds a very rich dimension to the tools and methods we have for investigating this activity, and the challenges involved in implementing it, while not insignificant, can be overcome. The more the research community embraces new methods of investigation, the more mature the research will become and this will be to the advantage of all who are interested in this field.

References

- Alves, F. (ed). 2003. *Triangulating Translation: Perspectives in Process Oriented Research*. Amsterdam: John Benjamins.
- Biber, D. Conrad, S. and Reppen, R. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge and New York: Cambridge University Press.
- Frey, L. R., Botan, C. and Kreps, G. 1991. *Investigating Communication – An Introduction to Research Methods*. Englewood Cliffs, N. J.: Prentice Hall.
- Göperich, S. Jakobsen, A.L. and Mees, I. (eds). 2008. *Looking at eyes: Eye-tracking Studies of Reading and Translation Processing*. (Copenhagen Studies in Language 36) Copenhagen: Samfundslitteratur.
- Jääskeläinen, R. 1987. *What Happens in a Translation Process – Think-Aloud Protocols of Translation*, Unpublished pro gradu study, Savonlinna: University of Joensuu, Savonlinna School of Translation Studies.
- Jensen, K.T.H. 2009. *Indicators of Text Complexity*, In Göpferich *et al.* (eds), *Behind the Mind*. (Copenhagen Studies in Language 37) Copenhagen: Samfundslitteratur, 61–81.
- Krings, H. P. 2001. *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes*. Kent, Ohio: The Kent State University Press, edited/translated by G.S. Koby.
- O'Brien, S. 2005. Methodologies for measuring the correlations between post-editing effort and machine translatability. *Machine Translation* 19 (1): 37–58.
- Tirkkonen-Condit, S. 1989. Professional vs. non-professional translation: a think-aloud protocol study. In C. Séguinot (ed.). *The Translation Process*. Toronto: H.G. Publications, School of Translation, York University. 73–85.
- Taboada, M. and Mann, W. 2006. Rhetorical Structure Theory: looking back and moving ahead. *Discourse Studies* 8(3): 423–459.