Aalborg Universitet



Computer-assisted Translation with a Focus on Revision at TextMinded Danmark A/S a Case Study

Bundgaard, Kristine

Publication date: 2013

Document Version Early version, also known as pre-print

Link to publication from Aalborg University

Citation for published version (APA): Bundgaard, K. (2013). Computer-assisted Translation with a Focus on Revision at TextMinded Danmark A/S - a Case Study.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
 You may freely distribute the URL identifying the publication in the public portal -

Take down policy If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Thesis Proposal

Computer-assisted Translation with a Focus on Revision at TextMinded Danmark A/S – a Case Study

Kristines TM,SDL BeGlobal Community - Translatio	n Results		👻 🕂 🗙	Term Recognition	-	φ ×	P
😼 Project Settings 👔 🚺 🚺 🦓							Weive
Maskinstøttet oversættelse med fokus på revision hos				i No open tembase.			
Maskinstettet oversættelse med fokus på revision hos CM Computer-assisted translation with a focus on revision at							
2 Maskinstøttet oversættelse med fokus på revision hos	AT Maskinstøttet translatio	on with a focus	on audit at				
Kristines TM	07-03-2013 :	21:54:01 MM_M	ac_Win7-Pc\MM_Mac_Win7				
Kristines TM,SDL BeGlobal Community - Translation	on Results Concordance Search Co	omments Mess	ages	Term Recognition Termbase Search			
Til forside.docx.sdlxliff [Translation]*					++	* >	c
Til forside.docx		J	Til forside.docx]
1 Thesis Proposal		AT	Thesis Proposal			P	
2 Maskinstøttet oversættelse med fokus på revision hos		AT	Computer-assisted translation with a focus on revision at			P	
3 TextMinded Danmark A/S – et casestudie		AT 🔄	TextMinded Danmark A/S - a case study P		Ρ		
4 21.03.2013		🕭 AT	21.03.2013 P		P		
Til forside.docx			Til forside.docx				
					_		
				INS 🗋 0.00% 👍 100.00% 🏈 Chars: 5	7		e .

Kristine Bundgaard, PhD fellow Research Group for Translation and Interpreting Department of Business Communication School of Business and Social Sciences Aarhus University

Principal supervisor: Associate Professor Tina Paulsen Christensen, PhD Co-supervisor: Associate Professor Anne Schjoldager, PhD

Table of Contents

1. Introduction	3
2. Motivation	3
3. CAT tools	4
3.1. The concept of revision	5
4. Investigating CAT as an 'extended' activity	7
4.1. Translation as a two-sided process	7
4.2. Situated, Embodied Cognition	7
5. Purpose statement and research questions	8
5.1. TextMinded Danmark A/S	9
5.1.1. Translation modes	10
5.1.1. Translation modes6. Research design	10 11
5.1.1. Translation modes6. Research design6.1. Case study comprising an experiment	10 11 11
 5.1.1. Translation modes	10 11 11 11
 5.1.1. Translation modes 6. Research design 6.1. Case study comprising an experiment 6.2. Mixed methods strategy 6.3. Workflow study 	
 5.1.1. Translation modes 6. Research design 6.1. Case study comprising an experiment 6.2. Mixed methods strategy 6.3. Workflow study 6.4. Revision study 	
 5.1.1. Translation modes 6. Research design 6.1. Case study comprising an experiment 6.2. Mixed methods strategy 6.3. Workflow study 6.4. Revision study 6.4.1. Methods 	
 5.1.1. Translation modes 6. Research design 6.1. Case study comprising an experiment 6.2. Mixed methods strategy 6.3. Workflow study 6.4. Revision study 6.4.1. Methods 7. Conclusion 	

1. Introduction

In this Thesis Proposal, I will focus on describing and discussing the background and design of my PhD project on computer-assisted translation (CAT). The project is a case study which examines internal (cognitive) and external (workflow) translation processes with a focus on revision at TextMinded Danmark A/S, Denmark's second largest language service provider (LSP). The case study includes an experiment, among other things, and as a consequence of this, I am facing a number of methodological challenges.

I have been looking forward to this Thesis Proposal and hope to discuss the research design in general as to whether it will be suitable in relation to the overall purpose of the project and furthermore, I would be thankful to receive input on specific decisions that will have to be made in the near future. These I will mention explicitly at the end of the Thesis Proposal.

The Thesis Proposal is divided into two main parts: in the first part, the background of the project is explained. This background includes an introduction to the central concepts which have guided the overall design of the project and a presentation of the purpose statement and research questions. In the second part, the research design is described and discussed.

2. Motivation

A number of studies have shown that CAT tools change the way translators work (see, for example, Dragsted 2004, 2006; Alves & Liparini Campos 2009) – a point that I personally believe to be true. For example, a translator was previously, when working without CAT tools, met with a blank screen or even a blank piece of paper when s/he had to make a translation. Today, with the integration of translation software, the translator is provided with suggestions for the translation of some or all sentences in the source text, depending on the type of CAT tool s/he is using. It seems that this is bound to change the translation process.

Since the late 1990s, translation memory (TM) systems have been the main CAT tool used in professional translation (O'Hagan 2009). This is supported by a survey that was carried out in 2006, in which 82.5% of translation professionals answered that they used a TM system (Lagoudaki 2006). Today, machine translation (MT) is increasingly being integrated into TM systems resulting in the type of CAT, which is often termed MT-assisted TM translation (Garcia 2012; Pym 2012). In fact, in a global survey carried out by the Translation Automation User Society (TAUS) in 2010, 49.3% of LSPs answered that they provided so-called post-editing of MT to their clients on a regular basis (TAUS 2010), and, in a survey conducted by the online networking forum for translation Proz.com in 2012, 54% of freelance translators reported that they use MT for some tasks (ProZ.com 2012).

In MT-assisted TM translation, the translator receives a suggestion for the translation of every source text sentence. With each suggestion, s/he has to decide whether it can be accepted as it is or whether it has to be revised. Thus, revision becomes a crucial part of translation. Actually, it can be assumed that revision constitutes translation in this translation environment, and if this is true, it can be argued that it redefines the traditional concept of translation. This is the motivation for focusing on the revision processes involved in MT-assisted TM translation in the present study.

This study sees itself as a contribution to Translation Process Research. Translation Process Research is a fairly young field within Translation Studies and the first empirical studies date back to the mid-1980s (Krings

1986; Gerloff 1988; Lörscher 1991). Before these studies, research in Translation Studies almost exclusively focused on translation *products* (Göpferich 2008:3). In Translation Process Research, as the name indicates, the focus is on all processes that lead to translation products (Göpferich 2008:1), but despite the widespread use of CAT tools and the quite obvious influence they have on the translation process, still relatively few studies have investigated the processes involved in CAT empirically (see, however, Dragsted 2004, 2006; O'Brien 2007, 2008; Alves & Liparini Campos 2009; Risku 2009; O'Brien et al. 2010; Christensen & Schjoldager 2011). What is more, almost no field studies have been conducted investigating the translation processes of professional translators working in their usual environments (see, however, Désilets et al. 2009).

To sum up, the motivation behind the study is first and foremost an interest in the impact of CAT tools on the translation process as CAT tools seem to change the activity of translators from mainly translating into revising. Another motivation is the absence of studies specifically investigating these aspects in a professional context.

In the next section, I will have a closer look at CAT and define the concept of revision used in this study.

3. CAT tools

CAT covers TM and MT systems. A TM is basically a collection of source texts and their translations, which are divided into segments (often sentences). By means of *matching*, these segments can be retrieved from the TM when the translator is about to translate a similar sentence. In this way, translations can be stored and reused. Three main types of TM matches are normally distinguished:¹ if the TM contains a source language segment which is identical to the source segment about to be translated a *100% match* is retrieved. If the TM contains a segment which is only similar to the one being translated a *fuzzy match* is presented to the translator, and if the TM does not contain a segment with a sufficient level of similarity, a *no match* is generated. When presented with a 100% or a fuzzy match, the translator needs to make a decision as to whether the match can be used as it is, whether it has to be revised, or whether it has to be rejected, in which case the segment is to be translate from scratch. When presented with a no match in a traditional TM system, the translator has to translate the sentence him/herself.²

In addition to these main match types, two other types of matches can be retrieved, namely *context matches* and *perfect matches*. A context match is a 100% match where the two source text segments are also preceded by exactly the same segment. In that sense, a context match is better than a 100% match. A perfect match is a form of context match, but in the case of a perfect match, the matching is done between an updated source file and a corresponding set of old bilingual documents rather than between a source text and a TM. If the TM system finds matches between the two source texts, the matches are also checked for context, i.e. the surrounding segments are checked to ensure that they are the same. If so, the perfect matches are extracted from the old bilingual documents and transferred to the new translation. Context and perfect matches are expected to require no further revision (SDL 2010).

¹ The description in this section of the functionalities of TM and MT systems is based on the TM software SDL Trados Studio 2011 and the MT software SDL BeGlobal, which are the systems mainly applied at TextMinded and thus, in this study.

 $^{^{2}}$ It is, however, possible for the translator to search the TM for a particular word or phrase using a so-called concordance search. Furthermore, a terminology database can be integrated with the TM in which case the translator will be supplied with the translation of specific terms.

When the TM system searches for matches in the database, it displays the degree of match between the source text segments as a percentage figure. In most TM systems, 100% matches are naturally defined as 100% correspondence between source text segments, and as default, matches with a correspondence between 70 and 99% are classified as fuzzy matches. If the degree of match between the source text segments is less than 70%, they are treated as no matches. The thresholds between the different match types can be defined according to what the LSP or the individual translator wishes. Context and perfect matches are displayed with the abbreviations *CM* and *PM*, respectively.

In MT-assisted TM systems, an MT function is added as an additional source of matches, and the no matches are then replaced by *MT matches*. The MT matches are generated on the basis of a so-called baseline MT engine comprising large amounts of data, which the MT system is able to analyze in order to provide the translator with a suggestion for the translation of the specific segment. Consequently, MT-assisted TM translation differs from traditional TM translation in that the translator no longer encounters empty target text segments, i.e. segments where no suggestion for the translation could be retrieved from the TM database. Instead, s/he is provided with a suggestion for the translation of every segment, either based on a TM or an MT engine, and when a match is generated by an MT engine, the match is displayed with the abbreviation AT for automated translation.³

When translating by means of MT-assisted TM, translators can be presented with the different types of matches in different ways. One way is that matches can occur interactively (on the fly, as it were). When matches occur interactively, the target text segment will stay empty until the translator enters that particular segment. The TM will then search for matches and insert the best match into the target text segment. Another way in which matches can occur is by means of pretranslating (without the translator having to ask for it). When segments are pretranslated, the target text segments are prefilled with the matches before the translator opens the file to be translated. In other words, when matches are pretranslated, the translator can read the proposed matches without actively entering the specific segments, and in effect, s/he does not have to enter them at all if s/he finds that they require no revision. As default, 100%, context and perfect matches are pretranslated whereas fuzzy and MT matches occur interactively. However, an LSP or an individual translator can specify which match types should occur in which way. Furthermore, the LSP or translator can choose to pretranslate the no matches using the MT engine and then turn off the MT engine before all matches are revised in order to reduce the load of the engine on the computer. However, they can also choose to leave the MT engine activated throughout the translation process, in which case the translator will also be able to see a suggestion from the MT engine when revising a match proposed by the TM. The distinction between interactively occurring and pretranslated segments and the activation of the MT engine are important for the present study and therefore, I will return to this in section 5.

3.1. The concept of revision

The activity of revising matches generated by a TM or an MT system is referred to in different ways in the literature. The revision of fuzzy matches is often called 'editing' or 'fixing' (Arenas 2009; Garcia 2010), whereas the revision of MT matches is most often termed 'post-editing' (Mossop 2007b; Garcia 2010; O'Brien 2011; Yamada 2011). However, there is some confusion as the terms 'editing' and 'fixing' are sometimes also used to capture the revision of both fuzzy and MT matches (Arenas 2009; Yamada 2011). In this study, 'revision' will be used to cover the translator's activity of changing all types of TM and MT

³ In the case of SDL BeGlobal, it is also possible for an LSP or a translator to 'train' the baseline MT engine with data from TMs. This means that the quality of MT matches can be improved by combining the data in the MT engine with translations already made by human translators.

matches into acceptable translations of the source text segments in question. Using Krings' (2001) terms, the present study will focus on both online revision and end revision (see also Jakobsen 2003; Asadi & Séguinot 2005). Online revision covers the changes that translators carry out during the 'drafting phase' of translation, i.e. the phase which "runs from the first text production keystroke until the first typing of the final punctuation mark" (Jakobsen 2002:192), whereas end revision covers the revision "done after the first full draft has been completed" (Jakobsen 2003:80) until the translator decides that the translation is finished, i.e. revision undertaken in the 'end revision phase' (Jakobsen 2002:193).

Another distinction between different types of revision is between the concepts 'self-revision' and 'otherrevision' by Mossop (2007a, 2007b). Self-revision refers to "the translator's own check of the draft translation" (Mossop 2007b:116), which seems to correspond to the term 'end revision' introduced above. However, Mossop states elsewhere that "self-revision is intermixed with the drafting process" (Mossop 2007a:12) which signals that self-revision might also encompass online revision. Other-revision is, in contrast, the revision of other translators' work (Mossop 2007a:6). As mentioned above, this study is interested in both online and end revision, i.e. self-revision. However, I find the term of self-revision a bit problematic as the revision of matches proposed by a TM system can be seen as an instance of other-revision as the translator is actually revising translations produced by other translators.⁴ Therefore, I use the term 'revision' to refer to the translator's revision of all matches in his 'own' translation, i.e. both online and end revision or in Mossop's terms, self-revision.

Research into the revision processes involved in CAT is scarce. Dragsted (2004, 2006) has shown that the use of a TM reduces the time spent on end revision and suggests that instead, the translators spend more time on online revision. Alves/Liparini Campos (2009), however, showed that translation technology reduces the need for revisions during the drafting phase. Garcia (2010) has studied the time necessary to revise pretranslated MT matches as opposed to translating directly from the source text and has shown that the time difference is not significant. However, Garcia's data suggested that translators might achieve higher quality when working with MT matches. O'Brien (2007) studied the cognitive load for different match types and showed, among other things, that the cognitive load for MT matches is similar to the one for 80-90% fuzzy matches. Arenas (2009) also compared MT and fuzzy matches and did this with respect to time spent on revision and final quality. The results suggested that translators work faster and achieve higher quality when revising MT matches.

This overview does not claim to be complete, but to my knowledge, no study has specifically investigated how revision is carried out in CAT, i.e. how translators actually turn the matches into an acceptable translation of the source text, and which impact it has on the translation process and the quality of the translation product whether the matches occur interactively or whether they are pretranslated. Furthermore, very few studies have studied the external processes involved in CAT. The present study hopes to help fill these gaps.

⁴ When a baseline MT engine trained with TM data is used, revision of MT matches can also, at least partly, be seen as an instance of other-revision in this sense, as the data used for generating the MT matches also include translations made by human translators.

4. Investigating CAT as an 'extended' activity

As mentioned in the beginning of the Thesis Proposal, the present study investigates both internal (cognitive) and external (workflow) translation processes. The primary reason for this is that a guiding principle for my research design is that I view translation as a so-called 'extended' activity, a notion introduced by Risku and Windhager (2013, forthcoming). In this section, the different types of translation processes will be explained together with an explanation of what is involved in investigating CAT as an extended activity.

4.1. Translation as a two-sided process

Studies in Translation Process Research have generally investigated the translation process in two different ways: either with a focus on internal or external processes (Göpferich 2008; Schubert 2009). According to Schubert (2009:19), the internal translation processes comprise the "mental activity involved in carrying out the translation work with all its steps and decisions", these not being open to direct observation. These internal processes are also often referred to as cognitive processes (Englund Dimitrova 2010). In contrast, external processes are "everything in the translation process which can be observed by another person", i.e. the translation workflow in which the translation process takes place (Schubert 2009:19).

Even though the internal and external processes have been investigated separately, they should not be seen as separate processes. The translator's work with the translation task, i.e. the internal translation processes, takes place in a specific context, in a workflow. This context influences the translator's mental activity and therefore, cannot be ignored when we want to understand what goes on in the translator's mind while translating (Göpferich 2008). In fact, we might think of it in this hypothetical way: if a translator was to translate the same text at the same time in two different places, embedded in different workflows, with different colleagues, the text coming from different customers, with different deadlines, with different standards for other-revision, maybe even using different CAT tools, would we then expect the translator to perform his/her task in the same way? My point being that we have to take the translator's context into account when we want to understand the task s/he is carrying out.

However, existing research in Translation Process Research seems to be characterized by an either/or with regard to internal and external processes. The existing studies have either focused on internal processes, which is the case for the majority of the studies (Göpferich 2008), or on the external processes (see, for example, Koskinen 2008; Risku 2009). No studies seem to have taken a combined approach to internal and external processes (Christensen 2011). In the present study, I will attempt to do this, which is reflected in the research design described in section 6 below.

4.2. Situated, Embodied Cognition

The idea that it is not enough to concentrate on what goes on in the translator's mind is a key aspect of the Situated, Embodied Cognition paradigm which has been introduced in Translation Studies by Risku (2010). In Risku's words: "Due to the major role played by the environment, any attempts to explain translation by describing processes in the mind of an individual alone are bound to fail. We need to find out not only what happens in a translator's mind, but also what happens elsewhere, e.g. in their hands, in their computers, on their desks, in their languages or in their dialogues" (Risku 2010:103).

The Situated, Embodied Cognition paradigm originates from cognitive science and according to this view of cognition, the individual and his/her surrounding environment forms an integral part of the processes of thought and behaviour, and therefore action is described as a *contextual* activity. Thus, the central concern of the Situated, Embodied Cognition paradigm is not the brain itself, but the fact that the brain allows

individuals to interact with the environment, artefacts and other human beings (Risku 2010). Cognition is thus not only embodied and situated, but also distributed between people and artefacts. This distributed nature of cognition is specifically stressed in the Distributed Cognition paradigm (Hutchins 2000). Following the theory of Distributed Cognition, cognitive processes can be distributed across members of a group and across internal and external structures, the external structures being material structures, i.e. artefacts. According to Hutchins (2000), artefacts are things that make humans smarter and that amplify the cognition of the artefact user which seems to correspond to what Wilson refers to as 'offloading' cognitive work onto the environment in order to reduce the cognitive workload (Wilson 2002).

TMs and MT engines are considered artefacts (Risku 2010; Christensen 2011), and it seems to be probable that what translators are doing when using TMs or MT is an instance of off-loading cognitive work onto the environment, i.e. the environment is used in a way that reduces the need to store, search for and process knowledge in the mind (Risku 2010). Actually, in line with the name translation *memories*, translators can be said to use the technology to store their memories and even share their memories with others. In this sense, the task of translating by means of MT-assisted TM can be seen as an activity that is distributed across the translator and other translators via an artefact (Dragsted 2006).

When not only focusing on the mind, but including the situation in the study of cognition, the view on cognition can be said to be 'extended' (Risku & Windhager 2013, forthcoming). When applying this view of cognition to the study of translation processes, Risku uses the term 'extended translation' to describe a research design which includes the situation and the key artefacts of translatory action it contains, e.g. translation technology (Risku & Windhager 2013, forthcoming). My view of CAT is inspired by the above-mentioned approaches to cognition and translation, and this is the rationale behind examining CAT as an extended activity, namely as an activity that comprises not only the translator's internal processes, but also the situation and the workflow s/he is a part of, i.e. the external processes. Furthermore, in this light, it seems most fruitful to carry out the study as a case study taking the actual workplace of the participating translators as the starting point. Carrying out field studies in the usual workplace of translators is something which has been encouraged repeatedly within Translation Process Research (Risku 2002; Göpferich 2008; Christensen 2011; Hubscher-Davidson 2011; Risku 2012) and Göpferich specifically states that, in order to live up to the idea of cognition as situated and in order to achieve ecological validity, future studies should let translators work under realistic conditions with their usual facilities and tools (Göpferich 2008:253). This leads me to the overall purpose of the study and to my research questions.

5. Purpose statement and research questions

This concurrent mixed methods study intends to examine computer-assisted translation (CAT) as an extended activity with a focus on revision at TextMinded Danmark A/S. The investigation comprises two interrelated studies: One part of the study explores external (workflow) translation processes qualitatively using participant observation and interviews. In the other part, a mix of quantitative and qualitative data is used to examine internal (cognitive) translation processes involved in two types of MT-assisted TM translation (modes). This other part of the study examines and compares revision processes of a number of professional translators applying either the interactive translation mode or the pre-translated mode (see below).

Based on the purpose statement, both parts of the study are guided by a central question and associated subquestions:

Workflow study:

Central question:

A. How is the workflow at TextMinded Danmark A/S – who does what, when and how?

In order to answer this question, I find it necessary to answer the following sub-questions:

A1. How does TextMinded Danmark A/S perceive of their workflow?

A2. How do concrete authentic translation assignments travel through TextMinded Danmark A/S and why is this so?

A3. How does TextMinded Danmark A/S perceive of the concepts revision and quality and which role do these concepts play for how the actual workflow is organized?

Revision study:

Central question:

B. How is revision carried out in an interactive and in a pretranslated translation mode?

This question will be explored through an experiment, which aims to answer the following sub-questions:

B1. How much time do the translators spend on the translation tasks?

B2. Do the translators sometimes delete matches and translate the source text segment from scratch and if so, which match types are deleted?

B3. What is the extent (measured in number of words and characters and in percentage of segment) of revision for the different match types (100%, fuzzy and MT matches)?

B4. Which types of revision (related to the selected revision typology) do the translators make in the different match types (100%, fuzzy and MT match)?

B5. Which reasons do the translators give for certain revisions?

B6. Does the translators' tolerance towards MT matches differ in the two modes?

B7. Do the translators sometimes prefer an MT match over a TM match and why?

B8. How is the quality of the translations when the translators decide that the translation is finished, i.e. when the end revision phase is completed?

5.1. TextMinded Danmark A/S

TextMinded Danmark A/S is Denmark's second largest LSP, employing a total of 63 people in its three Danish offices in Aarhus, Vejle and Copenhagen and in its offices in Chile, New Zealand and China. TextMinded Danmark A/S is also a member of the TextMinded Group, which is a group of independent European LSPs. The present study exclusively investigates TextMinded Danmark A/S⁵ and in particular, the main office located in Aarhus. TextMinded employs ten in-house translators, who are all based in Aarhus and apart from these, TextMinded also makes use of the services of a large network of freelance translators. The Aarhus office also counts, among others, Project Managers, Key Account Managers and the daily management. TextMinded's key service is translation, but the LSP also provides copywriting, terminology services and DTP services to its customers. These customers range from private people over small and medium-sized companies to large multinational companies and public sector companies.

⁵ From now on only referred to as 'TextMinded'.

5.1.1. Translation modes

At TextMinded, MT is at the moment being integrated into the TM system primarily used by the translators, SDL Trados Studio 2011. The MT system is the SDL BeGlobal software, which provides a baseline MT engine that can be trained with TextMinded's TMs. Together, these pieces of software constitute an MT-assisted TM system. This system allows TextMinded to predefine whether specific matches should occur interactively or be pretranslated; a decision, which has not yet been made conclusively. However, it is quite certain that 100%, perfect and context matches will continue to be pretranslated as they were in the traditional TM system, and fuzzy matches will continue to occur interactively. The most important choice seems to be whether MT matches should occur interactively during the translation process, or whether these matches should be pretranslated.

Thus, in the experiment in the revision study, two so-called MT-assisted TM 'modes' will be examined and compared with regard to specific revision processes: an interactive mode and a pretranslated mode (see Figure 1). Actually, these names are a bit misleading, as there are interactively occurring and pretranslated segments in both modes; however, the names are used to describe the mode in which *MT matches* appear interactively or are pretranslated, respectively. With regard to the other match types, matches appear in the same way in the two modes. The thresholds between the different match types used in the experiment will be the ones applied at TextMinded at the time of the experiment and in both modes, the MT engine will stay activated throughout the translation process as this allows me to see whether in some cases, the translator might prefer to use (parts of) suggestions from the MT engine when processing fuzzy matches.



Figure 1: The interactive and pretranslated MT-assisted TM translation modes.

6. Research design

In this section, the research design will be explained. First, the overall case study design and the mixed methods strategy will be introduced, and subsequently, the workflow study and the revision study will be described, explaining the methods applied in each of the studies and their relation to the research questions.

6.1. Case study comprising an experiment

The study is a case study in that it examines a contemporary phenomenon, CAT, in depth and within its reallife context over which I as a researcher have little or no control (Yin 2009). The case study design allows for the use of multiple methods and for a mix of quantitative and qualitative data (Yin 2009).

Sometimes, case studies have been equated with ethnography, which is, however, not necessarily true (O'Reilly 2009; Yin 2009). Case studies can employ ethnographic *methods*, but do not need to and are not limited to these (O'Reilly 2009; Yin 2009). Also, case studies do not need to take a lot of time as opposed to ethnographies, which usually require long periods of time, years even, in the field (Maaløe 2002:77; Yin 2009:15).

In the present case study, ethnographically inspired qualitative methods are applied in the workflow study, and in the revision study, another traditional method of social science research, namely an experiment, is conducted utilising both quantitative and qualitative methods, most of them frequently applied in Translation Process Research (see, for example, Dragsted 2006; Alves & Liparini Campos 2009; Mesa-Lao 2011; Teixeira 2011). With regard to the time period in the 'field', besides several meetings as preparation for the study, a week has been spent at TextMinded in February/March 2013 and the main part of May 2013 is planned to be as well.

The combination of ethnographic research methods and the study of cognitive processes is something which has been encouraged within Translation Process Research (Christensen 2011; Hubscher-Davidson 2011). Hubscher-Davidson (2011) has especially advocated this combination in order to enhance our understanding of what underlies the translation process and states, although without specifically discussing her view of cognition, that "translation process researchers can make efficient use of ethnographic research methods to increase their understanding of translators' perspectives and *situated* behaviours" (Hubscher-Davidson 2011:14, my emphasis). Here, again we sense the notion that the activities of the translators are situated and that understanding their cognitive processes therefore presupposes an examination of the situation they work in. To my knowledge, Koskinen (2008) and Risku (2009) are the only Translation Studies scholars who have applied ethnographic methods to the study of external processes in the actual workplaces of translators, and an ethnographically inspired study of external processes still remains to be combined with an examination of internal translation processes.

6.2. Mixed methods strategy

In order to fulfil the purpose of the study, a mixed methods design has been chosen. This choice is based on the notion that "there is more insight to be gained from the combination of both qualitative and quantitative research than either form itself. Their combined use provides an expanded understanding of research problems" (Creswell 2009:203). Also, maybe due to the interdisciplinarity and the young age of the field, combining quantitative and qualitative methods is an accepted way of investigating research issues within Translation Studies (Hansen 2005). In Translation Process Research, mixed methods designs are also encouraged (Hansen 2010; Christensen 2011; Hubscher-Davidson 2011) and are very frequently applied (see, for example, O'Brien 2007; Alves & Liparini Campos 2009; Teixeira 2011).

More specifically, the study employs a concurrent embedded strategy of mixed methods. According to Creswell (2009), when using this strategy, both quantitative and qualitative data are collected simultaneously, and one method (quantitative or qualitative) is embedded within another method (quantitative or qualitative), hence the name of the strategy. When applied to this study, I find that the revision study is logically embedded within the workflow study, cf the discussion in section 4. However, in Creswell's description of the concurrent embedded strategy, he states that this strategy generally has a primary and a secondary method, the secondary being embedded within the primary one. Even though I find that the notion of 'embedding' covers the relation between my two studies quite accurately, I do not consider the revision study to be a secondary method. My intention is that the workflow study will serve as an explanatory framework for my analysis of the data collected in the revision study, and in this sense, the workflow study is secondary. However, when taking into account my view of cognition and thus translation processes as very interrelated, it seems unfruitful to differentiate between primary and secondary. I shall here refrain from determining decisively whether one study is secondary.

In this study, qualitative methods are applied in the workflow study, and in the revision study, a mix of quantitative and qualitative data is used to examine cognitive translation processes. Because of this mix of data in the revision study, the design actually differs from Creswell's strategy where each of the two applied methods seems to be either quantitative or qualitative. However, I adopt the concept of an 'embedded' strategy, because it illustrates my research design well, but in contrast to Creswell, it here refers to a study in which one study applying both qualitative and quantitative methods is embedded in a study only using qualitative methods. Figure 2 shows the suggested research design:



Figure 2: Concurrent embedded mixed methods strategy.

6.3. Workflow study

As stated above, the workflow study aims to answer the following central question and related sub-questions:

A. How is the workflow at TextMinded Danmark A/S – who does what, when and how?
A1. How does TextMinded Danmark A/S perceive of their workflow?
A2. How do concrete authentic translation assignments travel through TextMinded Danmark A/S and why is this so?
A3. How does TextMinded Danmark A/S perceive of the concepts revision and quality and which role do these concepts play for how the actual workflow is organized?

Thus, the primary purpose of this study is to investigate the workflow at TextMinded. Schubert defines a workflow as follows: "A workflow is a chain or sequence of activities which can be described in terms of agents carrying out individual activities and influences controlling the activities" (Schubert 2009:19). These 'controlling influences' can be communication and cooperation with other agents as well as job specifications, resources and researched information, i.e. all elements which have an impact on the work of a specific agent. I take 'resources and researched information' to encompass CAT tools as Schubert states that resources can control the content, the linguistic form, the appearance and the work process, and that researched information has an impact on both the contents and on the linguistic form of the translation (Schubert 2009:24), which applies to CAT tools. In this sense, these 'resources' and this 'researched information' in Schubert's terms also seem to correspond to the notion of artefacts mentioned in section 4.2., namely tools or things that make humans smarter. On a more general level, there also seems to be link between the idea of controlling influences as elements (e.g. communication and cooperation with other agents and the use of artefacts), which impact on the work of a specific agent and the idea of cognition as distributed between people and between people and artefacts. Thus, workflows have to do with activities (what) that are related to each other with regard to the point in time when they occur (when), they can be described in terms of agents (who) carrying them out, in this case for example translators, project managers etc., and in terms of the controlling influences, which impact on or control the way agents work (how).

In order to provide answers to the research questions of the workflow study, I will apply ethnographically inspired methods, namely participant observation and interviews, which are typical data collection methods within ethnography (Hubscher-Davidson 2011). Ethnographic methods can yield empirical data about the lives of people in specific situations (Spradley 1980:16) and "ethnography usually involves the researcher participating, overtly or covertly, in people's daily lives for an extended period of time, watching what happens, listening to what is said, and/or asking questions through informal and formal interviews, collecting documents and artefacts" (Hammersley & Atkinson 2007:3). Thus, ethnographically inspired methods appear to be especially well suited for investigating the research questions of the workflow study in that it is possible to study when people do what and how they do it.

I will use observation, i.e. watching what happens and listening to what is said, by way of "the systematic noting and recording of events, artefacts and behaviours of informants" (Daymon & Holloway 2011:258) in order to capture the workflow at TextMinded, for example through observing how concrete translation assignments travel through the LSP (question A2). These observations will be supplemented by informal and formal interviews in order to answer question A1, i.e. how TextMinded perceives of their workflow and the 'why' part of question A2, namely why the 'journey' of translation assignments is as it is. Interviews and observation will also be triangulated to address TextMinded's perception of the concepts *revision* and *quality* and the role these concepts play for the organization of the workflow (question A3). The investigation of the

LSP's perception of revision and quality is also highly relevant as a framework for the revision study, which investigates exactly these concepts. Formal interviews will be conducted with relevant informants at TextMinded, for example managers, project managers, quality assurance managers and the language tools director.

The 'participant' part of 'participant observation' has to do with the overt/covert dichotomy mentioned above. "Participant observation means that you take part to some extent in the activities of the people you are observing" (Davmon & Holloway 2011:262), and this degree of participation can differ. Spradley developed a typology of participation, ranging from non-participation over passive, moderate and active participation to complete participation (Spradley 1980; K. M. DeWalt & B. R. DeWalt 2011). In the case of nonparticipation, the researcher has no involvement with the people or activities studied, and s/he thus observes phenomena from outside the research setting, for example by watching television. When participating passively, the researcher is present at the scene of action, but does not participate or interact with people. In fact, those being observed might not even know that the researcher is observing them. "Moderate participation occurs when the ethnographer is present at the scene of the action, is identifiable as a researcher, but does not actively participate or only occasionally interacts with people in it" (K. M. DeWalt & B. R. DeWalt 2011:23 about Spradley) whereas active participation occurs when the researcher does almost everything that the people under study are doing. Finally, complete participation covers a situation where a researcher studies a situation in which he is already an ordinary participant. In the present study, the level of my participation is expected to be between moderate and active participation in that I will be identifiable as a researcher at TextMinded, interact with people and engage in activities deemed relevant. This means that in some respects, I will not actively participate in the activities of the people at the LSP, but only observe them and occasionally interact with them and in other respects engage actively in the activities undertaken at the LSP. For example, I will spend time translating using the two MT-assisted TM modes being examined in the revision study in order to become acquainted with the technological reality the translators are faced with. As Hubscher-Davidson points out when arguing for the combination of Translation Process Research with ethnographic research methods: "it might be that process researchers can only fully understand an activity if they undertake it as a participant, and thus experience for themselves the process that they are studying" (Hubscher-Davidson 2011:14). An obvious threat to the validity of my data is the possibility that the people observed react atypically because of my presence (Daymon & Holloway 2011). Hopefully, the fact that I have visited TextMinded regularly, that I will be present at the LSP for several weeks, and that I will participate in certain activities will result in me not being considered a 'spectator' and make the employees at the LSP continue to go about their tasks in their accustomed ways and thus, reduce the so-called 'observer effect' (Daymon & Holloway 2011).

The combination of participant observation and interviews is hoped to provide answers with regard to what the LSP says it does, and what it does (Daymon & Holloway 2011:259), with the aim of understanding the workflow at TextMinded, i.e. answering the main question of the workflow study, question A. Inspired by Kastberg (2009), the data generated when following concrete translation assignments' journeys through TextMinded (question A2) will also form the basis for the modelling of ontogeneses, i.e. life cycles, of translation assignments which will illustrate which agents are involved in which activities and with the help of which artefacts at specific points in time during the workflow. In doing so, the study contributes with necessary knowledge about the genesis of translations in practice (as advocated by Risku & Windhager 2013, forthcoming).

6.4. Revision study

The revision study consists of an experiment investigating cognitive translation processes in CAT. The questions governing this study are:

B. How is revision carried out in an interactive and in a pretranslated translation mode?
B1. How much time do the translators spend on the translation tasks?
B2. Do the translators sometimes delete matches and translate the source text segment from scratch and if so, which match types are deleted?
B3. What is the extent (measured in number of words and characters and in percentage of segment) of revision for the different match types (100%, fuzzy and MT matches)?
B4. Which types of revision (related to the selected revision typology) do the translators make in the different match types (100%, fuzzy and MT match)?
B5. Which reasons do the translators give for certain revisions?
B6. Does the translators ' tolerance towards MT matches differ in the two modes?
B7. Do the translators sometimes prefer an MT match over a TM match and why?
B8. How is the quality of the translations when the translators decide that the translation is finished, i.e. when the end revision phase is completed?

The advantage of carrying out an experiment is that it allows for control of a number of variables which again allows for comparison of other variables (Yin 2009). In order to obtain this control of variables, experiments are often, and have in Translation Process Research almost exclusively been, carried out in laboratory settings. This experiment is carried out in a field setting, i.e. at TextMinded, where the translators are allowed to sit at their usual desks, use their usual computers and their usual CAT tools. This is absolutely essential if one is to take the notion of situated cognition seriously, as the situation is expected to have a great impact on the cognitive processes. Christensen (2011) has specifically advocated this combination of experiments and field studies arguing that it allows for a high degree of ecological validity *and* the control of relevant variables.

In the experiment, I will 'zoom' in on the part of the workflow that comprises the participating translators' actual work with translating specific texts, more specifically on how revision is carried out in the drafting and in the end revision phase. I plan to ask four experienced in-house translators to translate two similar texts (text A and text B), which will be either two different, but similar texts, or two excerpts of the same longer text; one in the interactive and one in the pretranslated mode. The two texts will come from the Danish company Bang & Olufsen, which sells high-end audio, video and multi-media products. The texts will most probably be technical, translation will be from English into Danish, and the two texts will be relatively short to avoid fatigue. In any case, the texts used in the experiment will be authentic translation assignments, i.e. source texts that have been or will be translated for Bang & Olufsen.⁶ More specifically, I am considering whether I should let all four translators translate text A in the pretranslated mode and text B in the interactive mode (see experimental design 1 in Figure 3), or whether I should let two of the translators translate text A in the pretranslated mode and the same for text B (see experimental design 2 in Figure 3). Not surprisingly, each design has its advantages and disadvantages.

⁶ I will naturally make sure that the two source texts have not been translated by translators participating in the experiment, and in both cases, a copy of the TM will be made before any translator starts translating the specific source texts so that all translators are presented with the same matches and not matches produced by the translator(s) before them. Also, the baseline MT engine will not be trained with the TM material.

Design 1 has the clear advantage that it will provide data on four translation processes on the same text in the same mode and thus, I will be able to see four 'shifts' from one translation mode to the other and hopefully see changes in revision behaviour that are attributable to the change in mode and not (only) to the individual working styles of the translators. However, the disadvantage is that the change might be due to the change in texts, even though the texts will be similar. In design 2, the advantage is that the data might still show a change in revision behaviour between the two modes and in this case, I might be able to say, at least cautiously, that the change is neither attributable to the translator's individual working styles *nor* the text, as not all translators translated the same text in the same mode. However, the disadvantage is that this design will merely provide data on two translation processes on the same text in each mode and thus reduce the strength of the design. This is another decision on which I would be very thankful to receive input.



Figure 3: Possible experimental designs.

6.4.1. Methods

In order to examine and compare the revision processes in both modes, I plan to triangulate a number of data collection methods:

- Keystroke logging (software: Inputlog)
- Screen capture (software: BB FlashBack Express)
- SDLXLIFF Compare (software provided by SDL allowing for comparison of 'before' and 'after' translation files highlighting the revisions made)
- 'Track Changes' function in SDL Trados Studio 2011
- Observation
- Retrospective interviews
- Quality evaluation

One of the primary challenges related to the experiment is the collection of data on the translators' keystroke activities when they are revising the suggested target text segments. Keystroke logging seems to be an obvious choice, as such software can capture an extensive amount of data on the process automatically and thus, it provides a large amount of data for analysis, but this is not as straightforward as I had hoped. Most logging systems are limited to logging activities in their own editors (for example Translog), which disqualifies them from being applied here as it is absolutely essential that the translators work with their usual CAT tool. Apparently, the only logging programs that can log events in other environments than their own editors are uLog and Inputlog (WritingPro 2013). Unfortunately, it has not been relevant to test uLog (Noldus 2013) with SDL Trados Studio 2011, as the current version of uLog only runs on computers with Windows XP, and the computers at TextMinded all have Windows 7 as their operating system. I have been in contact with the people behind uLog, who have informed me that a new version for Windows 7 will be released this spring, and if this happens before the data collection in May, I will of course test whether this program is compatible with SDL Trados Studio 2011 and is able to provide me with the data I need to answer the research questions.

I have been able to test Inputlog (Leijten & Van Waes 2006), which has been used successfully in combination with Trados (Torres-Hostench et al. 2010; Lacruz et al. 2012), but in both cases with previous versions of Trados (Trados Translator's Workbench). Mesa-Lao (2011), however, reports that Inputlog did not work properly in combination with Trados Translator's Workbench, but taking Torres-Hostench et al.'s (2010) study into account, this might have been due to the fact that the Workbench was also combined with TagEditor in Mesa-Lao's study. I have tested Inputlog with SDL Trados Studio 2011, both with an expert in Inputlog and with an expert in SDL Trados Studio 2011, and unfortunately, Inputlog logs insufficient information in SDL Trados Studio 2011.⁷ At this point, I am quite convinced that Inputlog cannot provide me with relevant information on the translation processes, but I plan to let Inputlog run if it does not compromise any of the other data by for example slowing down other programs.⁸ This will be tested further.

⁷ The reason for the incompatibility of Inputlog with SDL Trados Studio 2011 might be that this version of Trados applies a so-called side-by-side environment, where translation is carried out in the Trados editor itself. In the previous version, Trados Translator's Workbench, a so-called hybrid translation environment was used, which meant that texts were translated in Word (Christensen 2011). ⁸ Inputlog can log when the translator types or deletes something in SDL Trados Studio, but cannot identify the position of deletions and typings. As a result, the log file cannot tell me anything about the extent of a deletion, i.e. whether the translator deleted one letter or marked a whole word or sentence and then pressed 'delete'. What is more, Inputlog cannot log proposed matches from the TM or the MT engine, which is very important here, nor can it log it if the translator for example copies a term and pastes it into SDL Trados Studio. I have been in contact with the developers of Inputlog about this problem and know that they are now talking to SDL Trados Studio about the possibility of solving this.

As keystroke logging does not seem to be a fruitful method, I have looked into other ways of obtaining the needed data. Here, screen capture appears to be a useful method. Thus, the screen capture software BB FlashBack Express will be used to capture the translators' keystroke activities. BB FlashBack Express is a programme which records the activities on the translator's screen without being visible to the translator. As such, it supports the aim of obtaining ecological validity. Through BB FlashBack Express, I can measure how much time the translators spend on the translation task (question B1), I can see whether the translators choose to delete some matches and translate the source text segment from scratch (question B2) and I can identify match types and the number of words and characters being revised (question B3 and B4). This will also allow me to examine whether the tolerance towards MT matches differs in the two modes (question B6), and I can see whether translators choose to use (parts of) MT matches when processing fuzzy matches (question B7). However, the disadvantage of using screen capture is that the data are provided in a visual way, without the possibility to export them to a file, which can be analysed. Unfortunately, this means that I will have to count everything manually and enter the numbers in a spreadsheet to analyse.

Fortunately, in the case of pretranslated segments, there is an easier way of obtaining some of the data. Through the software SDLXLiff Compare, a comparison of 'before' and 'after' versions of the translation file is possible resulting in a file where revisions are highlighted and the number of changed words and characters are displayed. This is unfortunately only possible with regard to the pretranslated segments. Track Changes in SDL Trados Studio 2011 provides a similar function, but is limited to highlighting revisions. I am considering also activating this function, primarily as a backup source of data. Both methods are invisible to the translator.

As a supplement to screen capture and in order to answer question B4 about the relation between types of revision and different match types (question B4), I will also need to apply a revision typology. For example, I am considering Lindgren and Sullivan's (2006) online revision taxonomy, which they developed for and applied to writing. Here, revisions are categorised according to their position in the text and according to their content. When categorised according to their position, Lindgren and Sullivan distinguish between precontextual and contextual revisions, where pre-contextual revisions are revisions made at the point of inscription (i.e. the revision is only preceded and not followed by text), whereas contextual revisions are made within previously written text. With regard to content, Lindgren and Sullivan divide both precontextual and contextual revisions into form revision and conceptual revision. In the case of contextual revisions, the form revisions are further divided into typography, spelling, grammar, punctuation and format, and meaning-preserving revisions, and conceptual revisions are divided into text-based and balance revisions. With regard to Lindgren and Sullivan's revision taxonomy, I find the distinction between pre-contextual and contextual revisions particularly interesting in relation to MT-assisted TM translation. Intuitively, a revision of a match appears to be contextual as it is made within text already written, but practically, it might be made at the point of inscription in the sense that cognitively the translator has only reached this specific point in the match. Thus, I find the discussion of whether the revisions of a match are pre-contextual or contextual, interesting, but, I suspect, also difficult.

I have also considered using Brunette et al.'s (2005) revision criteria accuracy, readability, appropriateness and linguistic coding, which they apply to other-revision of translations, and I have considered using Mossop's (2007b) revision parameters. Mossop works with the broad parameters transfer, content, language and presentation, and within each of these, Mossop specifies a total of 12 parameters. Thus, Mossop's typology is more detailed, but it still seems applicable. I would be very thankful to receive input on whether one of these revision typologies seems especially well-suited for the present study, whether another typology would be more suitable, or whether it would be more advantageous to apply a hybrid of two or more typologies.

In order to answer question B5 on the reasons translators give for certain revisions and the 'why' part of question B7, I will carry out retrospective interviews when the translators have completed each of the translation tasks. As for example O'Brien 2007, Alves & Liparini Campos 2009 and Teixeira 2011, I choose to combine so-called 'online' data collection methods (Krings 2005), i.e. methods where data are collected *during* the translation process, with 'offline' methods, i.e. methods where data are collected *after* the translation process in order to supplement online quantitative data with qualitative data. In the retrospective interviews, the screen capture video will be used to prompt the translators' reflections on their processes and the translators will be asked to comment on specific revisions (question B5). At the time being, I am considering whether to preselect specific interesting points that would require revision and let the translators comment on these, which would allow me to carry out the interviews immediately after the translation process, or whether to take the translators' individual processes as the starting point and take some time to watch the screen capture videos and identify particularly interesting points before carrying out the interview. In both cases, I will be able to include my observations in the interviews. The choice between immediate and delayed interviews can obviously be seen as a trade-off between obtaining data on the revisions specifically relevant to the individual translator's process and the risk that the translators forget information about their processes (Ericsson & Simon 1984:19; Göpferich 2008; Heine 2012). In both cases, however, I find it important to ask the translators to comment on specific revision processes as several scholars point to the risk that participants might not be able to verbalise their processes if the task has become a routine (Ericsson & Simon 1984:15; Göpferich 2008; Christensen 2011), which must be said to be a relevant risk in this study as the participant translators are professionals who are used to working with CAT tools. However, I would hope that the relatively short translation tasks, the screen capture videos and the direction of attention to specific revisions will help facilitate reflection. This is supported by Ericsson and Simon's recommendation to instruct participants "to only report details that they remember heeding at the time of the original episode" because this may eliminate many people's "tendency to fill in information that they can't remember, but "must" have thought" (Ericsson & Simon 1984:19-20). Also, my intuitive expectation would be that professional translators have a comprehensive knowledge of concepts suitable for describing their processes, which one cannot, as pointed out by Heine (2012), count on in the case of inexperienced and semiprofessional producers of texts such as students.

Finally, with the aim of answering question B8 about the quality of the translations, a quality evaluation model will be applied to investigate the quality of the translations, especially whether a difference can be detected with regard to the mode in which the texts were translated. I am considering applying both an analytical approach focusing on errors in the target texts and a more holistic approach to quality evaluation. However, as the participants in this study are professional translators, I expect the target texts to be of a very high quality, and in this light, a holistic approach might not be able to capture differences in quality as opposed to a more analytical approach. At this point, I am primarily considering using the LISA (Localisation Industry Standards Association) QA model, which has been used by several scholars in Translation Process Research (see, for example, Arenas 2009; O'Brien et al. 2010), and to which I have access through TextMinded. A number of evaluators would then be asked to apply the LISA QA model to the eight translated texts with the aim of hopefully ensuring inter-rater reliability.

On the basis of exploring the different sub-questions through the triangulation of a number of methods, I hope to be able to answer the central question of the revision study, namely how revision is carried out in an interactive mode and in a pretranslated translation mode.

7. Conclusion

The purpose of this Thesis Proposal has been to describe and discuss the background and design of my PhD project. Thus, I hope to have provided the premises for a fruitful discussion at the oral part of the Thesis Proposal. As mentioned in the introduction, the combination of a case study with an experiment poses a number of challenges to the research design, but will hopefully also contribute with new knowledge within Translation Process Research. At the oral part of the Thesis Proposal, I hope to discuss the research design generally as to whether it seems suitable in relation to the overall purpose of the project. Furthermore, it should have become clear that a number of decisions still remain to be made, and I would be very thankful to receive input on these. Below, I have listed a number of questions, which I myself regard as relevant to discuss at this stage. As a concluding comment, I would like to say that it has been challenging, but also very pleasant to explicitly verbalise my considerations about the project and I look forward to discussing them very much.

Proposed points for discussion:

Points for discussion could include, but are certainly not limited to, the following:

- Which experimental design can best contribute to answering the questions of the revision study?
- Does one of the mentioned revision typologies seem specifically suitable for classifying the revisions made in the experiment or should I consider using another typology?
- Does the LISA QA model seem suitable for evaluating the quality of the translations from the experiment, or should I consider using another or an additional quality evaluation model?
- How should I approach the retrospective interviews in the revision study? Should I give priority to the immediate collection of data or should I take time to prepare an interview based on the participant's individual process?

Bibliography

- Alves, F. & Liparini Campos, T., 2009. Translation Technology in Time: Investigating the Impact of Translation Memory Systems and Time Pressure on Types of Internal and External Support. In S. Göpferich, A. L. Jakobsen, & I. M. Mees, eds. *Behind the Mind. Methods, Models and Results in Translation Process Research*. Copenhagen: Samfundslitteratur, pp. 191–218.
- Arenas, A.G., 2009. Productivity and quality in the post-editing of outputs from translation memories and machine translation. *Localisation Focus: The International Journal of Localisation*, 7(1), pp. 11–21.
- Asadi, P. & Séguinot, C., 2005. Shortcuts, Strategies and General Patterns in a Process Study of Nine Professionals. *Meta: Journal des traducteurs*, 50(2), pp. 522–547. Available at: http://id.erudit.org/iderudit/010998ar [Accessed February 7, 2013].
- Brunette, L., Gagnon, C. & Hine, J., 2005. The GREVIS Project: Revise or Court Calamity. Across Languages and Cultures, 6(1), pp. 29–45. Available at: http://www.akademiai.com/openurl.asp?genre=article&id=doi:10.1556/Acr.6.2005.1.3 [Accessed February 7, 2013].
- Christensen, T.P., 2011. Studies on the Mental Processes in Translation Memory-assisted Translation the State of the Art. *trans-kom*, 4(2), pp. 137–160.
- Christensen, T.P. & Schjoldager, A., 2011. The Impact of Translation Memory (TM) Technology on Cognitive Processes: Student-Translators' Retrospective Comments in an Online Questionnaire. In B. Sharp et al., eds. *Human-machine Interaction in Translation*. Copenhagen: Samfundslitteratur, pp. 119–130.
- Creswell, J.W., 2009. *Research design: Qualitative, quantitative and mixed methods approaches*, 3rd ed., Thousand Oaks, California: SAGE Publications.
- Daymon, C. & Holloway, I., 2011. *Qualitative research methods in public relations and marketing communications*, 2nd ed., New York/London: Routledge.
- Désilets, A. et al., 2009. How Translators Use Tools and Resources to Resolve Translation Problems: an Ethnographic Study. *Proceedings of Beyond Translation Memories: New Tools for Translators MT Summit XII*. Available at: http://www.mt-archive.info/MTS-2009-Desilets-2.pdf [Accessed February 15, 2013].
- DeWalt, K.M. & DeWalt, B.R., 2011. *Participant Observation A guide for fieldworkers*, 2nd ed., Plymouth: AltaMira Press.
- Dragsted, B., 2004. Segmentation in Translation and Translation Memory Systems An Empirical Investigation of Cognitive Segmentation and Effects of Integrating a TM System into the Translation Process. Copenhagen Business School.
- Dragsted, B., 2006. Computer-aided translation as a distributed cognitive task. *Pragmatics & Cognition*, 14(2), pp. 443–464. Available at: http://openurl.ingenta.com/content/xref?genre=article&issn=0929-0907&volume=14&issue=2&spage=443 [Accessed February 19, 2013].

- Englund Dimitrova, B., 2010. Translation process. In Y. Gambier & L. van Doorslaer, eds. *Handbook of Translation Studies. Volume 1*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 406–411.
- Ericsson, K.A. & Simon, H.A., 1984. *Protocol Analysis: Verbal Reports as Data*, Cambridge and London: The Massachusetts Institute of Technology.
- Garcia, I., 2010. Is machine translation ready yet? *Target*, 22(1), pp. 7–21. Available at: http://openurl.ingenta.com/content/xref?genre=article&issn=0924-1884&volume=22&issue=1&spage=7 [Accessed February 7, 2013].
- Garcia, I., 2012. Machines, translations and memories: language transfer in the web browser. *Perspectives: Studies in Translatology*, 20(4), pp. 451–461.
- Gerloff, P., 1988. From French to English: a look at the translation process in students, bilinguals, and professional translatord. Harvard University.
- Göpferich, S., 2008. *Translationsprozessforschung: Stand Methoden Perspektiven*, Tübingen: Narr Francke Attempto Verlag.
- Hammersley, M. & Atkinson, P., 2007. Ethnography: principles in practice, 3rd ed., London: Routledge.
- Hansen, G., 2005. *Qualitative and Quantitative Research and Empirical Translation Studies*. Available at: http://www.est-translationstudies.org/resources/research_issues/qualitative and quantitative research.html [Accessed February 12, 2013].
- Hansen, G., 2010. Integrative description of translation processes. In G. M. Shreve & E. Angelone, eds. *Translation and Cognition*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 189– 211.
- Heine, C., 2012. Prozessansatz im traditionell produktorientierten "Academic Writing" Textproduktionsunterricht. In D. Knorr & A. Verhein-Jarren, eds. Schreiben unter den Bedingungen von Mehrsprachigkeit. Frankfurt am Main: Peter Lang, pp. 99–116.
- Hubscher-Davidson, S., 2011. A discussion of ethnographic research methods and their relevance for translation process research. *Across Languages and Cultures*, 12(1), pp. 1–18. Available at: http://www.akademiai.com/openurl.asp?genre=article&id=doi:10.1556/Acr.12.2011.1.1 [Accessed February 11, 2013].
- Hutchins, E., 2000. *Distributed Cognition*. Available at: http://www.artmap-research.com/wp-content/uploads/2009/11/Hutchins DistributedCognition.pdf [Accessed February 11, 2013].
- Jakobsen, A.L., 2002. Translation drafting by professional translators and by translation students. In G. Hansen, ed. *Empirical Translation Studies: Process and Product*. Copenhagen: Samfundslitteratur, pp. 191–204.
- Jakobsen, A.L., 2003. Effects of think aloud on translation speed, revision, and segmentation. In F. Alves, ed. *Triangulating translation: perspectives in process oriented research*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 69–95.

- Kastberg, P., 2009. Zur Ontogenese einer Instruktion. Erarbeitung eines Methodengerüsts und erste Erfahrungen. In M. G. Ditlevsen, P. Kastberg, & C. Andersen, eds. Sind Gebrauchsanleitungen zu gebrauchen? Linguistische und kommunikativ-pragmatische Studien zu skandinavischen und deutschen Instruktionstexten. Tübingen: Narr Francke Attempto Verlag, pp. 38–68.
- Koskinen, K., 2008. *Translating Institutions: An Ethnographic Study of EU Translation*, Manchester: St. Jerome Publishing.
- Krings, H.P., 1986. Was in den Köpfen von Übersetzern vorgeht: eine empirische Untersuchung zur Struktur des Übersetzungsprozesses an fortgeschrittenen Französischlernern, Tübingen: Gunter Narr Verlag.
- Krings, H.P., 2001. *Repairing texts: empirical investigations of machine translation post-editing processes*, Kent, Ohio: Kent State University Press.
- Krings, H.P., 2005. Wege ins Labyrinth Fragestellungen und Methoden der Übersetzungsprozessforschung im Überblick. *Meta: Journal des traducteurs*, 50(2), pp. 342–358.
- Lacruz, I., Shreve, G.M. & Angelone, E., 2012. Average Pause Ratio as an Indicator of Cognitive Effort in Post-Editing: A Case Study. In AMTA 2012 Workshop on Post-Editing Technology and Practice (WPTP 2012). San Diego, USA, pp. 21–30.
- Lagoudaki, E., 2006. Translation Memories Survey 2006. Translation Memory Systems: Enlightening Users' Perspective, London.
- Leijten, M. & Van Waes, L., 2006. Inputlog: New Perspectives on the Logging of On-Line Writing Processes in a Windows Environment. In K. P. H. Sullivan & E. Lindgren, eds. *Studies in Writing: Vol.* 18. Computer Key-Stroke Logging and Writing: Methods and Applications. Oxford: Elsevier, pp. 73– 94.
- Lindgren, E. & Sullivan, K.P.H., 2006. Analysing online revision. In E. Lindgren & K. P. H. Sullivan, eds. Computer Keystroke Logging and Writing: Methods and Applications. Oxford: Elsevier, pp. 157–188.
- Lörscher, W., 1991. Translation performance, translation process, and translation strategies. A psycholinguistic investigation, Tübingen: Gunter Narr Verlag.
- Maaløe, E., 2002. Casestudier af og om mennesker i organisationer : forberedelse, feltarbejde, generering, tolkning og sammendrag af data for eksplorativ integration, test og udvikling af teori, Copenhagen: Akademisk Forlag.
- Mesa-Lao, Bartolomé, 2011. Explicitation in translation memory-mediated environments. Methodological conclusions from a pilot study. *The International Journal for Translation & Interpreting Research*, 3(1), pp. 44–57.
- Mossop, B., 2007a. Empirical studies of revision: what we know and need to know. *The Journal of Specialised Translation*, (8), pp. 5–20.
- Mossop, B., 2007b. Revising and editing for translators, 2nd ed., Manchester: St. Jerome Publishing.
- Noldus, 2013. *uLog*. Available at: http://www.noldus.com/human-behavior-research/products/ulog [Accessed February 19, 2013].

- O'Brien, S., 2007. Eye-Tracking and Translation Memory Matches. *Perspectives: Studies in Translatology*, 14(3), pp. 185–205.
- O'Brien, S., 2008. Processing fuzzy matches in translation memory tools: an eye-tracking analysis. In S. Göpferich, A. L. Jakobsen, & I. M. Mees, eds. *Looking at eyes. Eye-tracking studies of reading and translation processing*. Copenhagen: Samfundslitteratur, pp. 79–102.
- O'Brien, S., 2011. Towards predicting post-editing productivity. *Machine Translation*, 25(3), pp. 197–215. Available at: http://www.springerlink.com/index/10.1007/s10590-011-9096-7 [Accessed March 12, 2013].
- O'Brien, S., O'Hagan, M. & Flanagan, M., 2010. Keeping an eye on the UI design of Translation Memory: How do translators use the "Concordance" feature? In *European Conference on Cognitive Ergonomics*. Delft, pp. 1–4.
- O'Hagan, M., 2009. Computer-aided translation (CAT). In M. Baker & G. Saldanha, eds. *Routledge Encyclopedia of Translation Studies*. London/New York: Routledge, pp. 48–51.
- O'Reilly, K., 2009. Key Concepts in Ethnography, Thousand Oaks, California: SAGE Publications.
- ProZ.com, 2012. *State of the industry: freelance translators in 2012*. Available at: http://www.youtube.com/watch?v=zHHVAoUMv_g&list=UUq-uq3OdWl-4oZB-njIzDfw&index=6 [Accessed February 4, 2013].
- Pym, A., 2012. Translation skill-sets in a machine-translation age. Available at: http://usuaris.tinet.cat/apym/on-line/training/2012_competence_pym.pdf [Accessed February 4, 2013].
- Risku, H., 2002. Situatedness in translation studies. Cognitive Systems Research, 3(3), pp. 523–533.
- Risku, H., 2009. *Translationsmanagement. Interkulturelle Fachkommunikation im Informationszeitalter* 2. ed., Tübingen: Narr Francke Attempto Verlag.
- Risku, H., 2010. A cognitive scientific view on technical communication and translation: Do embodiment and situatedness really make a difference? *Target*, 22(1), pp. 94–111. Available at: http://openurl.ingenta.com/content/xref?genre=article&issn=0924-1884&volume=22&issue=1&spage=94 [Accessed January 12, 2013].
- Risku, H., 2012. Review of Gregory M. Shreve & Erik Angelone, eds. Translation and Cognition. *Target*, 24(1), pp. 198–202.
- Risku, H. & Windhager, F., 2013, forthcoming. Extended Translation: A Sociocognitive Research Agenda. *Target*, 25(1), pp. 34–46.
- Schubert, K., 2009. Positioning Translation in Technical Communication Studies. *The Journal of Specialised Translation*, (11), pp. 17–30.
- SDL, 2010. *About Translation Memory Matches*. Available at: http://producthelp.sdl.com/SDL Trados Studio/client_en/Edit_View/TMs/EVWorkingwithTMsAbout_Translation_Memory_Matches.htm [Accessed February 4, 2013].
- Spradley, J.P., 1980. Participant observation, Belmont, California: Wadsworth.

TAUS, 2010. *Postediting in Practice*. Available at: http://www.translationautomation.com/reports/postediting-in-practice [Accessed February 4, 2013].

- Teixeira, C., 2011. Knowledge of Provenance and its Effects on Translation Performance in an Integrated TM/MT Environment. In B. Sharp et al., eds. *Proceedings of the 8th International NLPCS Workshop Special theme: Human-Machine Interaction in Translation*. Copenhagen: Samfundslitteratur, pp. 107–118.
- Torres-Hostench, O. et al., 2010. TRACE: Measuring the Impact of CAT Tools on Translated Texts. In M. L. Gea-Valor, I. García-Izquirdo, & M.-J. Esteve, eds. *Linguistic and Translation Studies in Scientic Communication*. Bern: Peter Lang, pp. 255–276.
- Wilson, M., 2002. Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), pp. 625–636. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12613670 [Accessed February 7, 2013].
- WritingPro, 2013. *Logging Programs*. Available at: http://www.writingpro.eu/logging_programs.php [Accessed February 19, 2013].
- Yamada, M., 2011. *Revising text: An empirical investigation of revision and the effects of integrating a TM and MT system into the translation process.* Rikkyo University.
- Yin, R.K., 2009. *Case study research. Design and methods*, 4th ed., Thousand Oaks, California: SAGE Publications.