artesis VT working papers
in translation studies

2011-1

TRANSEDITING

IN STUDENTS' TRANSLATION PROCESSES

Iris Schrijver[1], Leona Van Vaerenbergh[1], Luuk Van Waes[2]

Artesis University College Antwerp[1]
University of Antwerp[2]

CONTACT

Iris Schrijver
Artesis University College Antwerp
Department of Translators and Interpreters
Schildersstraat 41
B-2000 Antwerp
iris.schrijver@artesis.be

artesis
vertalers en tolken

ABSTRACT

Editing and translating are interconnecting concepts with fuzzy borderlines. In 1989, Stetting coined the term *transediting* to refer to the overlap of both activities in the translation task. This article evaluates a methodology for studying transediting in the translation processes of translation students with different degrees of declarative and procedural knowledge. We examined the transediting carried out by four MA translation students, who were asked to translate an American patient information leaflet for a Dutch-speaking audience in accordance with the valid European Medicines Agency (EMA) directive. Of the four participants, two participants possessed only declarative knowledge of both the EMA standards and the text type. The other two participants not only possessed declarative knowledge of the EMA standards and the text type but also had some experience with translating patient information leaflets. Data on the translation processes were collected through thinking-aloud protocols and computer keystroke logging. By triangulating the data, we found not only variance in the degree of transediting carried out by the participants but also divergence in phase allocation of transediting in the translation processes.

KEYWORDS

transediting, translation-process research, segmentation, declarative knowledge, procedural knowledge

1. Introduction

In the 1980s, translation-process research emerged as a new research paradigm (within Translation Studies) which placed the main research focus on the different mental steps that translators take to produce target texts (TTs). The 'translation process' that these steps make up comprises a vast array of aspects such as reading and understanding the source text (ST) and the translation brief, documentation and research activities and the actual production of the target text. As a result of this new tradition of empirical process research, many insights have been gained into what takes place in the translator's 'black box'. With regards to the writing or drafting of TTs – in other words, the translator's text production – aspects such as the use of translation strategies and the segmentation of the writing phase in the translation process have been examined. However, a phenomenon which has received relatively little attention in translation-process research is transediting, the combination of translation and editing.

2. Investigating transediting

To provide an adequate background for the current study, we first review and discuss the literature on transediting and the related research methods, thinking-aloud protocols (TAPs) and computer keystroke logging (CKL).

Although translation may not be considered by some as writing per se, writing and/or text production do form central components of the translator's work. In various empirical process-oriented studies, the actual writing phase of the translation process has been examined with special attention devoted to segmentation, pause analysis, the use of translation strategies, units of translation and the working profiles of translators with different levels of expertise. However, the text production investigated in these studies primarily concerns the written transfer of content and style of the ST (in language A) into a TT (in language B). Yet, on some occasions translation goes beyond the mere reproduction of the ST in another language. Because of different text conventions, cultural divergence or even poor ST quality, translators are often forced to make minor and/or major textual changes. In this context, the term *rewriting*, which is sometimes used to describe translation, acquires a new meaning. Such rewriting – especially of poorly composed STs – is a professional reality for translators, which, to our knowledge, has not been examined extensively in translation-process studies.

In the late 1980s, Stetting (1989) introduced the term *transediting*, which she used as a composite term of *translating* and *editing* to refer to the "combination of both tasks" (p. 374). To illustrate the content of this term, Stetting (1989) lists three forms of transediting: (1) cleaning-up transediting, (2) situational transediting and (3) cultural transediting. These forms of transediting all imply acts of ST rewriting, and sometimes of ST re-ordering, which is essentially both communication-oriented and receiver-oriented. According to Stetting (1989), translators may opt to carry out transediting for three reasons: to adapt the ST to (1) a standard of efficiency in expression in the target language, (2) the intended function of the translated text in the target context, and (3) the needs and conventions of the target culture. Stetting (1989) does not enter into the discussion about free versus literal translation but simply presents transediting on a spectrum "at one end [the free end] followed by dynamic translation" (p. 379). She also stresses that transediting will predominate in the translation of non-fictional, practical, everyday texts, for which adequacy is sought.

Despite its innovativeness, Stetting's concept of transediting poses a number of problems. First, it is difficult to distinguish between translating and transediting. Since Stetting does not provide a clear definition of the term *transediting* (in contrast with translation proper), the two concepts remain fuzzy. This definitional vagueness is perhaps intrinsically characteristic of the concept of transediting since it can be argued that all translation tasks – and especially those in technical translation – involve translating,[1] writing[2] and forms of editing[3] (Risku, personal communication, 18 January 2010). Second, it is complicated to differentiate between transediting and other translation strategies such as adaptation.[4] This is not only due to Stetting's vague description of the concept but also because she lists cultural rendering as an example of transediting, which can actually be seen as a microlevel form of adaptation. Chesterman (1997) has tried to delimit transediting from adaptation and other translation strategies. He lists transediting among his classification of pragmatic translation strategies,[5] which he defines as "strategies which primarily have to do with the selection of information in the TT,… governed by the translator's knowledge of the prospective readership of the translation" (1997, p. 107). In Chesterman's (1997) classification, transediting is described as "the sometimes radical re-editing that translators have to do on badly written original texts: it includes drastic re-ordering, rewriting, at a more general level than the kinds of changes covered by the other [pragmatic] strategies" (p. 112). Unfortunately, Chesterman does not provide the reader with any examples of transediting to illustrate his description. When comparing Chesterman's description with Stetting's understanding of transediting, Chesterman's interpretation of transediting seems to cover only the cleaning-up form of transediting introduced by Stetting's.

[1] In this publication, we consider translating to be the linguistic, cultural and content transfer of an ST into a TT.

[2] In this publication, writing is understood to be the physical drafting of a text, without taking into consideration whether the unit being produced is transferred from a particular ST or whether the unit constitutes a 'new' item created by the person who is drafting the text.

[3] Mossop (2001) defines editing as follows: "Finding problems in a text which is not a translation, and then correcting or improving it, with particular attention to making the text suitable for its future readers and for the use to which they will put it" (p. iii). In the context of translation, editing can be interpreted as the monolingual revision of the TT, without taking into account the ST. Corrections and changes made to the TT may lead to clear deviations in formal correspondence from the ST (e.g., additions, omissions, optional class shifts).

[4] The term *translation strategy* is used in this publication to refer to the cognitive procedure(s) used for the solution of a translation problem.

[5] Chesterman (1997) lists the following pragmatic translation strategies: cultural filtering, explicitness change, information change, interpersonal change, illocutionary change, coherence change, partial translation, visibility change, transediting, and other pragmatic changes.

At this point it is also warranted to look at the observations made by Mossop in his 2001 book *Revising and Editing for Translators*. Mossop (2001) states that translators often engage in (mental) stylistic, structural and content editing, as well as copy-editing, as a means to tailor the TT to the target recipients. Whereas Mossop's stylistic tailoring can be considered part of the semantic and syntactic translation strategies proposed by Chesterman, his structural and content editing appears to correspond to Chesterman's view of transediting. Given Stetting's (1989) broad description of transediting, nearly all of Mossop's editing forms could be considered forms of transediting even though Mossop does not use or refer to this particular concept.

In her 1989 article, Stetting also presented several proposals on how to study the phenomenon of transediting among professional and novice translators. Although reality for many translators, transediting has unfortunately not been subject to extensive empirical research to date. The study of transediting has been restricted to news translation (e.g., Hursti, 2001; Vuorinen, 1996), in which particularly forms of gatekeeping (i.e., the control and selection of news flow) have been examined. To our knowledge, transediting has not been studied in the translation of other text types although it is most definitely present in intergeneric translation.[6] In this publication, we will present the results of a small-scale, empirical study on transediting in the translation processes of four MA translation students. Given that Stetting does not provide a clear definition of transediting, in which she delimits transediting from other translation strategies, we will use the following operational definition of transediting: Transediting is the manipulation of the ST content and ST structure in the TT, including omission, addition, restructuring and substitution of ST units.

3. Research questions

The experiment reported on in this publication was drawn up as a pilot study in the larger framework of a more detailed PhD research project on text production in the translation process. Four Dutch-speaking MA translation students were asked to translate an American patient information leaflet (PIL) for a Dutch-speaking audience in accordance with valid European legislation.

For this exploratory study, we addressed four research questions, which we formulated as follows:

1. What is the time spent on the different phases in which the translation processes of four MA translation students are divided?
2. How are the participants' translation processes segmented?
3. Do the participants use transediting in their translation processes? And if so, (a) Which transediting strategies can be observed?, and (b) In what phase of the translation process are transediting strategies used?
4. Can triangulation of TAPs and CKL shed light on the use of transediting in the translation process?

[6] *Intergeneric translation*, which implies the transfer between genres, is a term proposed by Askehave and Korning Zethsen (2002).

4. Method

4.1. Materials: Source text and translation brief

In order to be able to study transediting, the participants were given a translation brief, which can be considered an extreme case of translation since the TT has to comply with other textual norms than those found in the ST. The TT has to be in accordance to the valid European legislation and guidelines.[7] These norms, established by the European Medicines Agency (EMA), stipulate that the package inserts of medicinal products must be "legible, clear and easy to use"[8]. They must also adhere to a certain structure. These standards are outlined in a template drafted by the Quality Review of Documents (QRD) group of the EMA.[9] Important features are standard phrases in each of the six sections of the PIL, as well as the active way of addressing the reader and the explanations of difficult medical terms to improve PIL readability and usability.

Because the study focused on transediting, the ST which was used had to meet certain criteria concerning structure, content and terminology. The selection of a British PIL would not have been an option since medicine commercialized in the UK must contain a PIL formulated in compliance with EMA norms. Consequently, little to no transediting would have to be carried out when translating such a text into Dutch. Therefore, we chose to use patient information about a medicine not yet commercialized in the European Union. We selected the patient information of Geodon® from the American website RXList.[10] We decided to shorten this text to approximately 900 words without any alterations to the essence of the text. Repetitions and irrelevant sections were deleted. In addition, we also deleted some explanations of medical terms for laypeople. The ST layout was changed slightly, for example, by removing bullet points to create an enumeration in one long phrase. The logic behind these manipulations was our interest in the translators' focus on TT readability.

4.2. Participants

The selection of the participants was based on a number of parameters. They all had to be native speakers of Dutch with English as one of the two foreign language studied in the master programme. The participants had completed a BA programme

[7] Legislation concerning patient information leaflets are set out in Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use, which was amended by Directive 2004/27/EC of the European Parliament and of the Council of 31 March 2004.
[8] Title V, article 59, §3 of Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use (Consolidated version: October 5, 2009).
[9] It is important to note that the experiment presented in this publication had been carried out before the Dutch QRD template was revised in January 2010 and versions 7.3. and 7.3.1. were issued.
[10] Full text available from http://www.rxlist.com/geodon-drug.htm (last consulted on 21 September 2010).

in applied language studies with a specialization in translation at our department. A third requirement was that they had taken the introductory course on technical and scientific translation. In this course, detailed information had been given on the EMA directive as well as on PIL readability and usability. By selecting participants who met these requirements, we were able to assume that they all shared important characteristics concerning language and translation competence. More important, all of the four participants had declarative knowledge of the EMA standards and the patient inserts text type. However, two of the selected participants also had some experience – albeit limited – with translating Spanish and British PILs into Dutch. In other words, they also had some procedural knowledge.

4.3. Design and procedure

Each participant did the experiment individually in an empty classroom. One of the authors of this article was present in the same room but did not intervene unless the participant stopped verbalizing her thoughts for more than two minutes. The same laptop was provided for every session. All four participants were allowed to use dictionaries and the Internet. They had at their disposal two electronic monolingual Dutch dictionaries (one generic, one medical) as well as two bilingual dictionaries English–Dutch/Dutch–English (one generic in electronic version, one medical in paper copy). Moreover, a paper copy of the Dutch QRD template was supplied. No time restrictions were imposed for the translation task.

The participants were asked to say everything that came to mind while translating, without trying to explain or justify these verbalizations. The verbalizations were audio-taped using a digital voice recorder. Since the participants were not familiar with this method of concurrent thinking aloud, a practice session was organised for each participant individually before starting the pilot study. The two participants working with *Translog* were shown the program beforehand to familiarize them with the computer software.[11] The participants who were monitored using *Inputlog* were not familiarized with this program since they were simply asked to work in an MS Word environment.

4.4. Data-collection methods

Data on the translation process were collected using two online methods, that is, during the translation process itself (Krings, 2005): (1) think-aloud protocols (TAPs) and (2) computer keystroke logging (CKL). Each data-collection method is discussed below as well as their merits and limitations. Let us first look at TAPs. TAPs are a research method developed in the field of cognitive psychology by Ericsson and Simon (1984, 1993). However, for the last two decades it has also been widely used in translation-process studies (see Jääskeläinen, 2002, for an overview). By having translators verbalize their thoughts when translating, we can tap into information heeded in their short-term memory (or working memory). TAPs have provided insights into the problem-solving and decision-making proc-

[11] *Inputlog* was used to track CB's and SA's translation processes, whereas RN's and KL's translation processes were registered using *Translog*.

esses of translators, both novices and professionals. Although this research method has proved extremely useful to unravel some of the mysteries surrounding translators' cognitive processes, it does have some limitations. The main points of criticism levelled at TAPs concern interference and low levels of both validity and reliability (see Dam-Jensen & Heine, 2009, and Göpferich, 2008 for an extensive overview). Verbalizations may disturb the problem-solving cognitive processes of translating since translating and speaking are performed simultaneously in the same domain. Another consequence of this interference is a slowing-down effect: thinking aloud while translating takes up more time than translating 'silently'. This claim has been empirically validated by Krings (2001) and Jakobsen (2003), who report average delays of 33% and 25% respectively. Several studies have also found that TAPs alter the sequencing of cognitive processes, for example, the segmentation of the translation process (Jakobsen, 2003). However, further research is needed to assess the effect of TAPs on task performance. Objections have also been raised against the validity of TAPs. Verbalizations provide only indirect data since they reflect only conscious, non-automated cognitive processes. Only information that is in the person's focus of attention, that is, heeded in short-term memory, will be verbalized. Furthermore, participants' willingness to verbalize thoughts may vary considerably. The reliability of TAPs has also been questioned since it might be possible that inconsistencies occur between the mental processes involved and the verbalizations uttered.

In this study, we chose to use a second data-collection method (computer keystroke logging (CKL)), with the objective of complementing the verbal report data *and* compensating some of the shortcomings mentioned above. CKL provides a quantitative account of all the movements on the computer (keystrokes and mouse movements such as clicks, deleting, scrolling and cursor navigation) during a task. This research instrument has long been used in the field of writing research but since the development of *Translog* in 1997 CKL has also found its way into translation-process research. Several keystroke logging tools are currently available (e.g., *Inputlog*, *Jedit*, *Proxy*, *ScriptLog*, *Trace-it*, *Translog*), each with its own characteristics. For the pilot study, two experiments were logged using *Inputlog* version *3.0 Beta*, developed at the University of Antwerp by Van Waes and Leijten (for a presentation of the program, see Leijten and Van Waes, 2006). For the other two experiments, we opted for *Translog 2006 Academic Edition*, designed by Jakobsen and Schou at the Copenhagen Business School (see Jakobsen and Schou, 1999). Although *Inputlog* was initially designed for tracking writing processes and *Translog* for recording translation processes, both programs essentially log the same data. There are indeed some differences to be observed but in the interest of brevity we will state the most striking difference: the configuration of both logging tools. *Inputlog* records the keystroke and mouse movements in all Windows programs (e.g. MS Word), whereas *Translog* has its own in-house developed word processor.

4.5. Data analysis

The data-collection methods provided us with a vast amount of qualitative and quantitative data. In order to manage the amount of data efficiently, it was essential that we selected a transcribing and segmenting protocol which would enable us to align and compare the verbalizations with the log data easily and in a structured manner. Widely used transcript models for verbal report data are the GAT system (Gesprächsanalytisches Transkriptionssystem) and the TEI (Text Encoding Initiative) guidelines (for discussions of both methods, see Göpferich, 2008). We chose to use a protocol which was a slightly adapted version of the GAT system. All verbalizations, the lengths of pauses and paralinguistic signals were transcribed.

The output data were extracted from the keystroke logging tools. *Inputlog* saves the data in IDF files, from which MS Excel or XML files can be generated. In *Translog*, the *TranslogSupervisor* interface generates linear representations of the output data. Although both logging tools use different symbols in their notation systems, the resulting linear representations are easy to read. Table 1 shows an excerpt of the output data generated by *Inputlog*. Table 2 provides similar information for *Translog*. For main symbols and conventions in *Inputlog* and *Translog* output files, see respectively Leijten and Van Waes, 2006, and Jakobsen and Schou, 1999.

Table 1. Inputlog excerpt

Interval	output
0:00:00	{4437}[Movement]{12203}[Movement]]{37500}Lees·de·hele·bijs
0:01:00	luiter·zorgvuldig·{2610}door·voordat·u·start·met·het·gebruik·van·dit·ge neemiddel{3000}.[ENTER]{5875}- ·Bewaar·deze·bijsluiter.·Het·kan·nodig·zijn·{2360}om·deze·bogmaals· {2109}[BS9]nogmaals·{7203}[BS6]·eens·door·te·lezen
0:02:00	.[ENTER]{7000}Heeft·u·nog·vragen, ·rond[BS3]{3812}aadpleeg·dan· {6312}u·[BS]w·arts·of·apotheker.[ENTER]{6313}Dit geneemiddel·is·aan· u·persoonlijk·{2422}voorgeschreven.·Geef·dit·geneesmiddel·n

Table 2. Translog excerpt

```
[Start][✱04:19.691][🖱]✱✱✱✱✱✱✱[🖱]B✱ijsluiter:◆informatie◆voor◆d
e◆ghe⌫⌫ebruik(sters)↲✱✱G✱✱✱edodon⌫⌫⌫⌫⌫odonn↲([🖱][Paste]✱
✱✱[🖱]✱✱✱[🖱]✱[🖱]✱✱⌫✱⌫⌫↲(✱ziprasio⌫done◆HCI)↲↲✱✱✱✱✱✱✱1
.◆✱✱✱✱✱WAT◆✱IS◆GEODON◆✱✱✱EN◆✱WAARVOOR◆WORDT◆HET◆GEBRUIKT✱↲✱✱✱
GEOD⌫⌫⌫eodon◆is◆een◆[✱43.646]
```

The transcripts of both the verbal data and the logging data were divided into segments, which are also called language bursts (Chenoweth & Hayes, 2001). These segments are strings of words, characters or mouse clicks preceded and followed by pauses longer than the defined threshold. We decided to segment the transcripts since we believed that continuous stretches of TAP placed next to linear representations of logging data would not have provided clear representations of the translation process. Minimum pause length has frequently been used as a segmentation parameter since pauses in verbalizations are often considered as indicators of problems or cognitive processing. We used the following operational definition of a segment: A segment comprises either all keystrokes and/or mouse movements that follow a pause of ≥2 seconds and that are followed by a pause of ≥2 seconds, or a pause of ≥2 seconds. This definition was formulated in accordance with features relevant to the pilot study *and* on the basis of previous research conducted by other scholars (e.g., Englund Dimitrova, 2005; Krings, 2001). For instance, Englund Dimitrova (2005) states that "very short pauses, up to about 2 seconds, are often followed in the process by the correction of typographical errors, and thus seem to be used for monitoring processes that are not directly related to the process of translating as such" (p. 96). The central research questions that we strived to answer in the pilot study did not address the correction of typographical errors in fluent text production.

As a final step in the preparation of the data for analysis, we placed the segments of both data sources next to each other in an MS Excel file. Subsequently, each TAP segment was coded as a means to reconstruct the processes taking place in the participants' heads. Each segment could receive more than one coding label. Then we analysed the log data to see if these confirmed our initial coding of the TAP segments or provided further information. In the event of additional information, we introduced another coding label. The coding labels used in the pilot study were inspired by the classification of coding classes proposed by Krings (2001), which consists of the following main classes:

- GLOBTASK: global task-related processes
- SOURCE: source text-related processes
- TARGET/PROD: target text production-related processes
- TARGET/RED: transediting in target text production-related processes
- TARGET/MON: target text monitoring-related processes
- REFBOOK: reference work related-processes

The classification above is an adapted version of Krings's classification since categories have been omitted and added to the original classification in light of the formulated research questions. For example, the coding class WRITE was omitted because physical writing processes were not investigated in the pilot study. The coding class TARGET/RED is not featured in Krings's classification. It was introduced in the pilot study to focus on the primary research item of the study, transediting. This main coding class was divided into several subcodes, which differentiated between various trans-editing strategies.

In empirical research, triangulation has often been restricted to the moment of data collection (Dörnyei, 2007). To ensure that triangulation would also be implemented at the moment of data analysis, we created a working document in which the TAP and log-data segments were placed next to each other. Table 3 shows a sample excerpt of this document.

Table 3. Excerpt of triangulation file of coded TAP log-file segments

Start	TAP[1]	End	Code	Code	Start	Inputlog[2]	End
10:35	is een soort psychotropische psychische psy psychische uitwerking komma uh ook bekend als	11:02	TARGET/PROD/CONCRETE/LETT	TARGET/PROD/CONCRETE/VARIANT	10:35	geneesmiddel·met·een·psu[BS]ychotro[Movement][LeftButton][Movement][LeftButton][Movement][BS5]hische·uitwerking,·	10:55
						{3578}	10:57
						ook·bekend·als·	11:01
11:02	(4.0)	11:06			11:01	{2937}	11:03

Note 1. Translation of Dutch verbalization: *is a sort of psychothropic, physic, psy, physic effect comma uh also known as*
Note 2. Translation of keystroke log data: *medicine with a psy[BS]ychothro[Movement][LeftButton][Movement] [LeftButton][Movement][BS5]ical effect*

5. Results and discussion

This section reports the findings of the current study. Given the exploratory character of the data analysis, the results will be presented alongside critical reflections.

5.1. Times spent on different phases of the translation process

Let us first take a look at the total amount of time needed to complete the translation task, i.e. the processing time (see Table 4). At a glance, there are no major inter-individual differences. Participants with both declarative knowledge of the text type and procedural knowledge of translating PILs (i.e., participants CB and SA) completed the task in less time than participants with only declarative knowledge of the text type (i.e., participants KL and RN). The participant with the highest amount of procedural knowledge (CB) needed considerably less time than SA. This may seem self-evident but studies have shown that more experienced translators do not necessarily complete translations in less time than novice translators. As Krings (2001) explains, "the speed increase due to automation of certain processes is often compensated by a greater awareness of problems and a greater depth of processing of other sections" (p. 279). It should be remembered, however, that the participants in the pilot study were all MA translation students and were not considered to be experienced translators. The differences between participants are mainly based on their declarative and procedural knowledge.

In addition, in order to further refine our observations, we took into account the total number of words as to calculate their production speed.[12] Important to note is that whereas CB pondered every ST sentence in her translation process and omitted very few ST elements, SA forgot to translate three ST paragraphs. This information sheds light on the total time which SA spent on the translation task. A similar observation can be made about RN, who forgot to translate the paragraph on how to take Geodon. Another factor that helps put processing time into perspective is the verbalization volumes since verbalizations will increase the times that are needed to complete tasks. This parameter can be operationalized as "all strings between two spaces in the transcript" (Krings, 2001, p. 286) and measured in words. Verbalization effort is the verbalization volume per processed amount of TT words.

Table 4. Processing times and production speeds, verbalization volumes and efforts

participant	processing time[1]	TT word count	production speed[2]	verbalization volume[3]	verbalization effort[2]
CB	02:07:45	1115	8.73	4937	4.43
SA[13]	02:16:16	1039	7.62	3388	3.26
RN	02:23:09	978	6.83	6001	6.13
KL	02:24:32	1024	7.09	2720	2.66

Note 1. Expressed in hours:minutes:seconds.
Note 2. Expressed in words per minute.
Note 3. Expressed in words.

In the analysis of the distribution of the overall translation task into different phases, we first have to take a look at the different methods that can be used to di-

[12] Krings (2001) refers to this parameter as *processing speed* and operationalizes it as the amount of ST processed per unit time. In our pilot study, some ST units were not even to be present in the TTs. For this reason, we decided to focus on the total numbers of words produced per unit time (i.e., per minute)in the final versions of the TTs.
[13] The verbalization volume and verbalization effort of participant SA were calculated in such a way that we could take into account the verbalizations that were technically inaudible and therefore not transcribable. 474 segments with clearly audible verbalizations consisted of 2595 words, which comes down to 5.47 words per segment. This led to the extrapolation that the remaining 145 segments with inaudible verbalizations might consist of roughly 793 words. This would amount to a calculated verbalization volume of 3388 words, and subsequently a verbalization effort of 3.26 words per minute.

vide the translation process into chunks. In empirical translation-process studies (e.g., Englund Dimitrova, 2005; Jääskeläinen, 1999; Jakobsen, 2002), the translation process has generally been divided into three major phases: (1) the pre-writing phase, (2) the writing phase and (3) the post-writing phase. Englund Dimitrova (2005, p. 86) provides the following definitions of the three phases:

1. Pre-writing phase: begins when the participant has received the ST and the oral information about the translation brief, and finishes when the participant starts to write down the TT as an integral text. Making notes about word meanings, etc. while reading the ST for the first time is thus not considered as a start of the writing phase.
2. Writing phase: begins when the participant starts to write down the TT and finishes when she has written down an integral version of it.
3. Post-writing phase: begins immediately after the writing phase and finishes when the participant declares that she is finished with the translation task.

This method of dividing the translation process will be referred to as Method 1. Table 5 shows the temporal data (absolute and relative values) of the task performances for all four participants.

Table 5. Times spent on the translation task and its three phases, expressed in hours:minutes:seconds (in absolute and relative values according to Method 1)

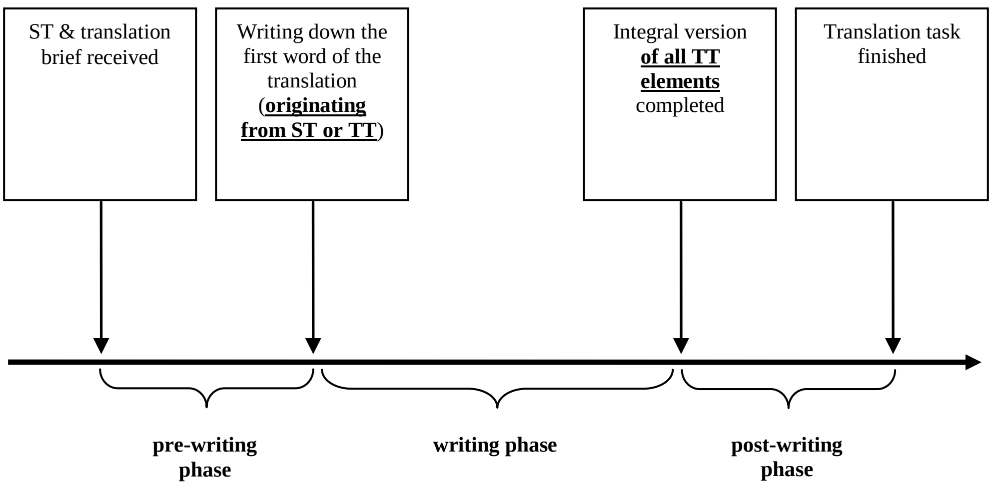
participant	translation task	pre-writing phase	writing phase	post-writing phase
CB	02:07:45 (100)	00:00:55 (0.72)	01:59:02 (93.18)	00:07:48 (6.11)
SA	02:16:16 (100)	00:00:38 (0.46)	01:59:02 (87.35)	00:16:35 (12.17)
RN	02:23:09 (100)	00:04:38 (3.21)	02:15:37 (94.74)	00:02:54 (2.03)
KL	02:24:32 (100)	00:00:16 (0.19)	02:23:23 (99.20)	00:00:53 (0.61)

All four participants had short pre-writing phases. The data of the writing phases do not present major inter-individual differences in relative terms, with the exception of SA. The post-writing phase presents two striking cases which suggests a difference in working styles: KL appeared to prefer to solve translation problems and revise while drafting the TT, whereas SA gave priority to translating fluently and returning to challenging items at the end of the translation process.

Several observations have to be made about the delimitation of the three phases. First, the notion of an integral version of the TT has not been defined clearly. Does an integral version refer to a TT that consists of the target-language equivalents of *all* the ST units? Or of all the *necessary* ST units? Has an integral version been produced when the participant has declared that she has finished a draft translation and will subsequently look for missing parts? Or must these forgotten words and phrases be inserted into the TT to be able to speak of an integral version? Thus, a detailed definition of integral version of the TT is fundamental for establishing the end point of the writing phase and, consequently, the starting point of the post-writing phase. Of course, these points may be easy to pinpoint in a straightforward translation task, but in case of text types which vary considerably among languages and/or cultures, no clear three-phase translation process may be discerned. Especially for *intergeneric* translation tasks, the translator will have to produce TTs that may not contain all ST units. Jakobsen (2002) states that the end of the writing phase can be identified unambiguously by the typing of the final punctuation mark since this is "typically followed by a great deal of cursor movement … indicating that the target text is being monitored and reworked" (p. 193). However, in the data collected during the pilot study this boundary was not clear-cut. This may have been the result of transediting, which required extensive restructuring and therefore a vast amount of cursor movement throughout the translation process. However, even in a normal translation situation, it can be challenging to pinpoint the end of the writing phase. For example, a translator can decide to leave a number of particularly difficult translation problems open and only come back to it after having revised all the already translated units. A second observation is that in the current three-phase division it is unclear how many other activities are present in one single phase other than the dominant activity (reading, writing or revising respectively).

In the pilot study, the translation task may lead to the use of transediting. The first TT word written down did not necessarily have to stem from the ST. To delimit the writing phase from instances of text generation not related to the ST, we decided to integrate this parameter into the definitions, which led to Method 2. The TAP data and log data also confirmed this need to differentiate. Figure 1 contains an overview of the main characteristics of both methods in which the differences between both methods have been underlined.

Method 1



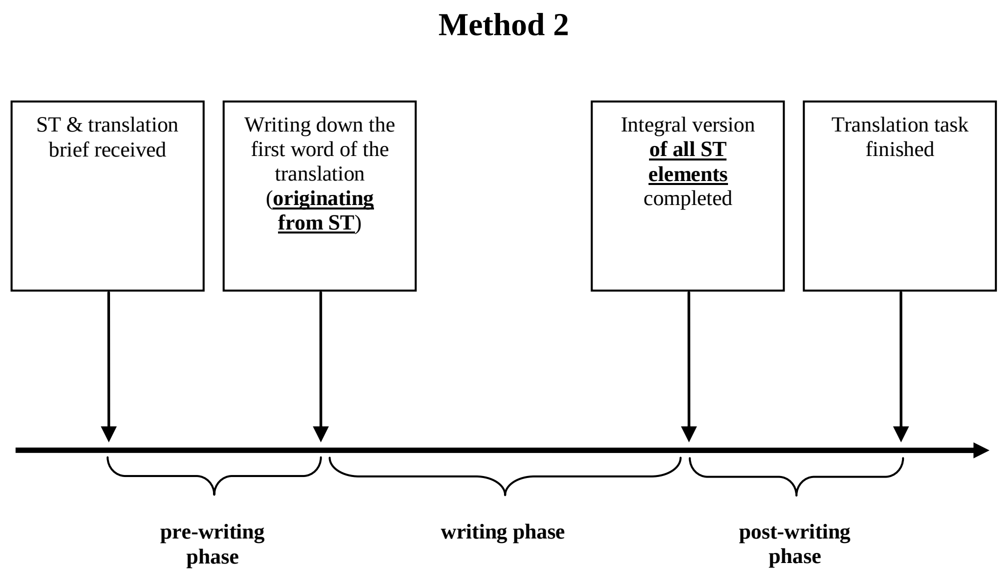


Figure 1. Characteristics of Methods 1 and 2 (differences underlined).

Table 6 shows the temporal data (absolute and relative values) of the task performances for all four participants, including the times spent on each of the three phases of the participants' translation processes.

Table 6. Times spent on the translation task and its three phases, expressed in hours:minutes:seconds (in absolute and relative values according to Method 2)

participant	translation task	pre-writing phase	writing phase	post-writing phase
CB	02:07:45 (100)	00:10:18 (8.06)	01:46:32 (83.39)	00:10:55 (8.55)
SA	02:16:16 (100)	00:12:39 (9.28)	01:47:09 (78.63)	00:16:28 (12.08)
RN	02:23:09 (100)	00:05:58 (4.17)	01:42:18 (71.46)	00:34:53 (24.37)
KL	02:24:32 (100)	00:14:08 (9.78)	02:09:31 (89.61)	00:00:53 (0.61)

The inter-individual differences in time spent on the different phases are small and appear to be results of the four participants' different working styles (see section 5.3). Activities such as reading, inserting the introductory table and table of contents into the TTs, restructuring the ST according to the TT norms, introducing the standard phrases into the TTs, transferring the ST information into Dutch by translating, adding and omitting ST information, and revising the translation draft were carried out at different points in the participants' translation processes. RN almost immediately started translating the ST, whereas CB, SA and KL already carried out many transediting activities before the actual translation of ST units. With regards to the duration of the writing and post-writing phases, slightly greater inter-individual differences can be observed which also appear to be the results of different working profiles. Both CB and SA integrated the EMA norms while they were translating the ST. They used their post-writing phases to reread the TTs. However, RN has a much longer post-writing phase, since she first translated the whole ST and then started transediting the TT. KL spent very little time on the post-writing phase. This may seem surprising but it is partly due to the definition used for the post-writing phase. KL translated the ST by taking the QRD template as a starting point and revised while drafting the TT.

By using Method 2 to divide the translation process into phases, more attention was drawn to the pre-writing and post-writing phases, as can be seen in the graph in Figure 2.

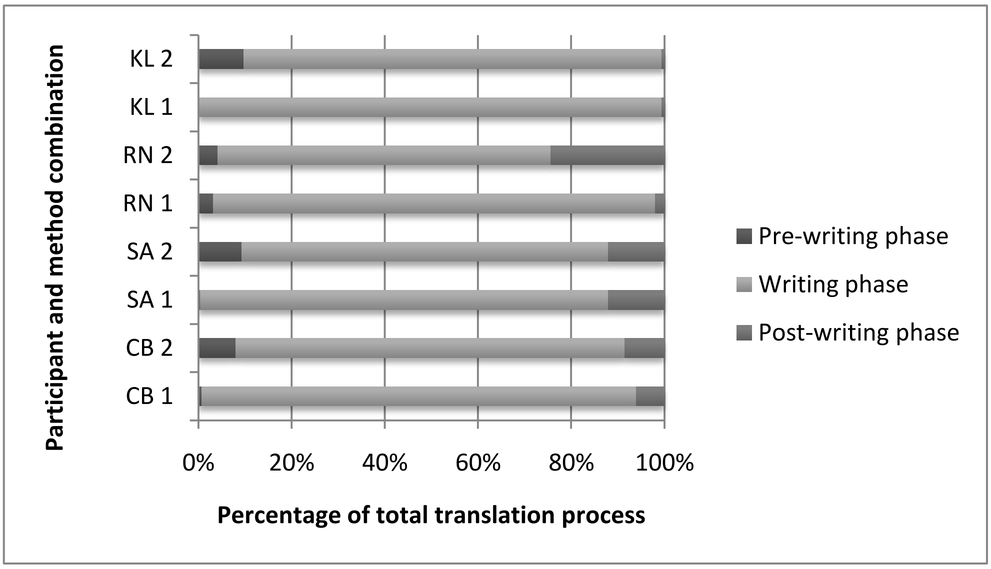


Figure 2. Division of translation process into three phases (according to Methods 1 and 2, with KL1 = participant KL according to Method 1, KL2 = participant KL according to Method 2, etc.).

It can be argued that Method 2 provides a clearer view of ST-based text production. However, Method 2 still does not take into account the recursive nature of the translation process.[14] Thus, in further research it might be helpful to divide the translation process into a number of phases, without labelling them according to the predominant cognitive activity. In this respect, several approaches can be taken into consideration. For example, looking at the process on the basis of TT paragraph level or leaving text units together and focusing instead on time by dividing the translation process into a particular number of chunks. It would also be interesting to focus on which and how cognitive activities follow one another to examine the recursive nature of the translation process and the different functions of one particular cognitive activity such as revision or writing (1-to-1 rendering of the ST or transediting) in the translation process, as suggested by Breedveld (2002).

5.2. Segmentation of the translation process

Another interesting aspect constitutes the segments. According to Dragsted (2005), the segmentation of the translation process is an indicator of cognitive processing because "the size and nature of the unit processed cognitively […] is governed by the limited capacity of working memory" (pp. 49–50). Table 7 provides an overview of important segmentation data (on the basis of Method 2):

Table 7. Total numbers of segments and segmentation of the different phases of the translation process

participant	total number of segments	pre-writing phase	writing phase	post-writing phase
CB	968	83	807	78
SA	1105	80	848	177
RN	1024	26	675	323
KL	932	104	827	1

Given CB's experience with the translation of PILs and her declarative knowledge of the text type under investigation, it could be argued that the translation task did not require the same amount of cognitive effort for her as was possibly the case for the other participants with only declarative knowledge of PILs. This would result in fewer pauses and, consequently, in fewer segments. However, this hypothesis does not appear to be valid for SA. She produced the TT with the highest number of segments. Surprisingly, KL's translation process consisted of only 932 segments whereas she spent the longest on the translation task and her translation contained the least number of characters (see table 8). In terms of segment distribution across the three phases, we can observe that while RN had the shortest pre-writing phase

[14] Note that although the recursive nature of the process can be observed in both TAPs (a phenomenon also present in writing processes, as discussed by Van den Bergh and Rijlaarsdam, 1996) and log data, some instances of the recursive nature of the translation process may not be visible, such as internal revision.

she also needed relatively fewer segments for this phase than the other participants did. This can be explained by looking at RN's working style.

In writing research, segmentation is also considered an indicator of fluency defined by rate of production (Chenoweth & Hayes, 2001). This line of reasoning can also be applied to the translation process. More fluent or more experienced translators use fewer segments, which are larger in size, to produce TTs, as demonstrated by Jakobsen (2003). This may be the result of knowledge already proceduralized (Chenoweth & Hayes, 2001 paraphrasing Towell, Hawkins, & Bazergui, 1996). Fluency can be defined as the average burst length by which the TT is produced. To assess fluency, we cannot turn to production speed only, but we also have to look at the number of keystrokes – including text-elimination activities – that were needed to produce the final TTs. A second important factor is the number of segments with which the TTs are produced[15]. In Table 8 data are presented on text-production keystrokes and text-production segments.

Table 8. Text-production keystrokes and segments, fluency, processing and production speeds, process/product ratios across the pilot-study participants

translation-process feature	**participants**			
	CB	**SA**	**RN**	**KL**
total number of text-production characters[1]	8906	8479	9262	7527
total number of TT characters[2]	7327	6798	6293	6472
total number of text-production segments	454	539	515	466
fluency[3]	19.62	15.73	17.98	16.15
processing speed[4]	69.71	62.22	64.70	52.08
production speed[5]	57.35	49.88	43.96	44.78
process/product ratio[6]	1.22	1.25	1.47	1.16

Note 1. Includes spaces and text elimination

Note 2. Includes spaces.

Note 3. Expressed in characters per text-production segment.

Note 4. Expressed in text-production characters per minute.

Note 5. Expressed in TT characters per minute.

Note 6. Expressed in characters.

The amount of text-production characters that RN produced is the highest. This implies that on the one hand RN's processing speed was relatively high, while on the other hand she had to delete a large amount to ensure that her translation complied with TT norms, which resulted in lower production speed. A similar – albeit

[15] The total number of segments, shown in Table 7, included pause segments. It could be argued that no physical text production takes place in pause segments and that, consequently, they should not be taken into consideration when calculating fluency.

opposite – trend can be observed for KL. KL translated only those ST segments which were deemed necessary in light of the TT norms. KL's degree of text elimination (and, consequently, the product/process ratio) was therefore relatively low. RN used 49 segments more than KL, but her segments were longer than KL's, which makes RN the more fluent translator of the two. When looking at production speeds, it could be argued that KL was slightly more efficient than RN.

CB and SA had comparable process/product ratios. Given their somewhat similar working profiles, these ratios seem to suggest that they both revised and corrected more or less the same amount of text while drafting their translations. However, they differ in other respects. CB produced the TT with the highest number of characters but she did so with the smallest number of segments. In other words, CB was the most fluent translator of all four participants. She outperformed KL, who required 466 segments to produce her TT, which was 855 characters (i.e., 91 words) shorter than CB's TT. This explains KL's lower processing and production speeds. Thus, if we look at only the segment data, the overall picture becomes quite distorted. By taking into consideration fluency, processing and production speeds, the differences between CB and KL become clearer. It is also warranted to say that TT quality should be taken into account in further research to make more detailed and nuanced observations.

It may not come as a surprise that CB was the most fluent translator since she possessed declarative and (the highest degree of) procedural knowledge. However, in terms of fluency the difference between CB and SA was striking. Consequently, no clear distinction could be made between the two participant groups. SA needed less time than RN and KL to produce a (larger) TT. Nevertheless, the numbers of segments with which she produced the TT was larger, which amounts to a shorter average burst length and thus to less fluency.

5.3. Transediting strategies

After careful consideration of the TT norms, we believed that the main transediting strategies to be used by the participants would be macrolevel and a microlevel[16] restructuring, substitution and omission of ST units as well as addition of information in the TT. This belief proved to be correct when we analysed and coded the TAP and log data. In this section we describe and discuss the participants' different working profile in detail. Our main focus is on the transediting strategies. In addition, we will also highlight how the participants applied transediting in the three phases of their translation processes (see Method 2).

Participant 1 (CB)

CB's translation process was characterized by the integration of various transediting strategies in all three phases and the verbalization of text knowledge. CB started the translation process *not* by reading the ST but by writing the introductory

[16] *Microlevel restructuring* of ST information in the TT refers to the restructuring at the level of the phrase and the sentence whereas *macrolevel restructuring* takes place above the level of the sentence and consists of mainly the restructuring of paragraphs.

table and index obligatory in EMA PILs. Then, she proceeded with the translation of the ST. Before starting to translate a particular paragraph, she verbalized to which TT section the paragraph belonged and placed the translation in the right TT section, introducing the obligatory Dutch TT section header and standard phrases. When confronted with an ST paragraph that did not belong to the TT section which she had been drafting, CB immediately placed it into the right TT section.[17] The ST macrolevel restructuring took place according to the EMA directive and did not appear to be problematic for CB, with the exception of the information about alcoholic beverages, which CB did not place into the correct TT section. Noteworthy is that CB did not verbalize these restructuring strategies. However, the log data showed that she did so without any hesitation. This may indicate automatized action. In the pre-writing and writing phases, CB adjusted the layout of the TT in accordance with the EMA directive, using bold, underlining, etc. Having translated all the ST information, CB read through the entire TT and corrected a few sentences.

CB often verbalized declarative and procedural knowledge. She remarked that the ST did not contain information that would normally feature in the TT, such as data on the excipients, the marketing authorization holder and manufacturer and the date of last approval of the PIL. She introduced these phrases into the TT but left some open space for the client to fill in the missing information. This declarative knowledge could be inferred from the QRD template but the restructuring of the TT, as verbalized in the TAP, presented indicators of knowledge which could not have been deduced from the template. For example, CB placed the information on the active ingredient and excipients of the medicine after the index. This structure was not mentioned in the QRD template provided to the participant. Although this particular information on the active ingredient and excipients should be placed in the last section of the information leaflet, it is sometimes mentioned after the index in Dutch EMA-approved PILs. Therefore, it suggests that CB used her knowledge of the text type and perhaps procedural knowledge of translating PILs for this particular action. Moreover, the ST did not explicitly state that the active ingredient of Geodon was ziprasidone but CB deduced this from the fourth ST paragraph. The other participants, except for KL, did not appear to notice this.

CB also reorganized information at the microlevel for reasons of readability and usability, mostly in the writing phase. For example, CB did not hesitate to break down long enumerations of heart diseases or possible side effects by using bullet points. However, she did not do so in the enumeration of symptoms of schizophrenia and bipolar disorder in the second ST paragraph. Nevertheless, the focus on readability and usability was also present in CB's translation of ST illocutions. Some of the ST instructions are expressed by a *it is important to + infinitive*-construction, which CB translated into Dutch using an imperative form to "make the instruction more direct". For example, she translated *It is important to be patient* with *Wees geduldig* (Dutch for *Be patient*). Surprisingly, CB used the same

[17] The exception to this general working method were ST paragraph 2 and 3 which she skipped. CB recognized the title of the fourth ST paragraph to be similar to the second section of EMA PILs and proceeded with its translation. After having translated the final sentence of the ST, she translated the two previously skipped paragraphs.

argument for translating several sentences starting with *it is best to + infinitive*. She first translated these recommendations quite literally but then substituted them with imperative forms. Making the instructions more direct or explicit was also a motive for CB to change the sentence order of these illocutions. A few ST instructions were accompanied by statements as to why the patient had to act in a particular way. When these explanations were placed before the instructions in the ST, CB decided to change the sentence order. This stylistic choice of microlevel restructuring is not mentioned on the QRD template but is advocated by the Dutch regulatory commission of PILs and constitutes a common feature in Dutch PILs.

A third aspect to which CB paid special attention was consistency in word choice. She systematically used *geneesmiddelen'* (Dutch for *drugs*) instead of *medicatie* (Dutch for *medication*), and *arts* (Dutch for *physician*) instead of *dokter* (Dutch for *doctor*), a common feature in Dutch PILs. This precise word choice, which resulted in substitution of ST units, can also be observed in the formulation of the translation. Several studies (e.g., Tirkkonen-Condit, Mäkisalo, & Immonen, 2008) have suggested that translators verbalize a literal translation as the first step in the text production. However, often CB immediately opted for specific words in the target language, which were not literal translations of ST words but were indeed common in Dutch PILs.

Addition of information was also observable. While translating, CB instantly recognized immediately at first glance that several ST phrases were similar to QRD standard phrases and opted for those standard phrases instead of ST-oriented translations. For example, *Take Geodon only as directed by your doctor* was translated with *Volg bij het innemen van Geodon nauwgezet het advies van uw arts* (Dutch for *Always take Geodon exactly as your doctor has told you*). If the ST contained more information than the Dutch standard phrase, CB added particular elements to standard phrases. CB applied the same strategy in other instances too, using phrases that are frequent in Dutch PILs. For example, she translated *Do not stop taking your medicine without your doctor's approval* with *Stop de behandeling met Geodon niet zonder uw arts te raadplegen* (Dutch for *Do not stop the treatment with Geodon without consulting your physician*). Another example of addition was the explanation of medical terms in the TT, a strategy that CB used more frequently than the other participants. Difficult terms such as *long QT syndrome*, *heart failure*, *anti-arrhythmics*, *hyperglycemia* and *diabetes* were consistently followed by descriptions for laypeople. CB verbalized that she would explain those terms but did not add why these descriptions were needed.

Omission was observable in the TAP and log data. CB consciously left out ST information in the translation to avoid redundancy. For example, she did not translate the third ST paragraph about the risk of dangerous changes in heart rhythm because this information had already been included in previously drafted TT sections. The log data also showed several (not verbalized) cases of omission. Some of these could be regarded as cases of redundancy (e.g., the omission of temperatures in Fahrenheit when also temperatures in Celsius were mentioned) but there were also a few units which CB first left out because she was unable to find the right equivalents. However, she subsequently forgot to translate those units (e.g., *Class Ia*, when referring to the type of anti-arrhythmics). Nonetheless, in comparison with the other participants, CB omitted fewer ST words in total. Based on her verbalizations, it could be argued she used a relatively conservative approach to omission in comparison to SA and KL.

Participant 2 (SA)

In terms of planning the overall translation process, SA clearly manifested another working method than CB. She started her translation process by drafting the introductory table but then turned to the ST. Subsequently, she read through the entire ST, jotting down which TT sections each ST paragraphs belonged to. Then, she started drafting the TT, first typing the index and then translating the ST paragraphs that she had assigned to TT Sections 1 through to 5. Interestingly, SA had erroneously assigned the third ST paragraph to TT Section 3 but while translating the paragraph, she placed it correctly in TT Section 2 without verbalizing why. Thus, macrolevel restructuring was part of SA's pre-writing and writing phases. After having drafted all 5 TT sections, SA checked the "gaps in the text", referring to the problematic elements that she had previously skipped and the ST sentences that she had not yet included in the TT. Surprisingly, SA left the sixth ST paragraph completely out of the TT, as well as the paragraphs on NMS and hyperglycemia, without verbalizing why. She mentioned the aspects which were not featured in the ST and would normally be stated in the TT but she did not express the desire to contact the client or express those aspects in the translation. In the post-writing phase, SA read through the TT, correcting some orthographical errors and adding a few words. The TT was structured correctly in accordance with the EMA directive. The adaptation of the TT layout was part of the pre-writing and writing phases.

Focus on readability and usability – which would lead to microlevel restructuring – was not a consistent feature in SA's translation process. The only reference to readability present in the TAP is the use of bullet points for enumerations. However, the TAP and logging data did not manifest any concern for an active, target audience-oriented style in the translation. The instructions on how to take Geodon were translated with infinitives in Dutch, and not with more logical imperative forms (e.g., *Swallow the capsules whole* was translated with *De capsules in hun geheel inslikken*). When an instruction was followed by an explanation, this structure was maintained in the TT. Generally, SA substituted *arts* (physician) for *dokter* (doctor) in her translation and the log data indicated that she corrected herself if her initial translation contained *dokter* (doctor). However, one occurrence of *dokter* was still present in the translation. Important to note is that SA's final TT contained several grammatical errors.

SA added the necessary QRD headers and obligatory standard phrases to her translation. However, no explanations of difficult medical terms were added, except for an explanation of the term *long QT syndrome*. However, it can be argued that the explanation given (*komt voor bij het syndroom van Jervell-Lange-Nielson en het syndroom van Romano Ward*, Dutch for *does occur with Jervell and Lange-Nielson syndrome and Romano-Ward syndrome*) did not provide any clarification of the medical term. The TAP and logging data also revealed that SA omitted a relatively large number of ST elements. For example, in the section on pregnancy, she added the standard phrase *Vraag uw arts of apotheker om advies voordat u een genees-*

middel inneemt (Ask your doctor or pharmacist for advice before taking any medicine), but omitted the ST phrase *Before taking Geodon, tell your doctor if you are pregnant or plan on becoming pregnant*. Other omissions of information were those of paragraph 6, the information on NMS and hyperglycemia, and – more importantly – the warning of increased risk of death for patients with dementia-related psychosis. The latter was translated by SA with *een verhoogd risico* (an increased risk), with the omission of *of death*. SA did not provide any explanations for these omissions nor did she refer to them in the TAP.

Participant 3 (RN)

RN started her translation process by reading the QRD template and the source text. She remarked that the ST differed considerably from the structure proposed by EMA. After approximately six minutes, RN started translating the entire ST without changing its structure. As she herself verbalized, she "would focus first on the content". However, on some occasions she did take into consideration the information provided in the QRD template. For example, she used the EMA section headers when translating the title of the first and fourth ST paragraphs as well as the information about pregnancy and breast feeding. Moreover, when translating the sentence *While on Geodon, check with your doctor before starting any new prescription or over-the-counter medications, including natural/herbal remedies*, RN decided to use a sentence from the QRD template although she confirmed that it was not completely equivalent to the ST sentence. RN began to use transediting strategies only during the last minute of the writing phase when she translated the final two ST sentences. At this point, she restructured the translation according to the QRD template, introduced the obligatory introductory table, index and standard phrases, adapted the TT layout and omitted redundant words and sentences. The macrolevel restructuring of the translation was generally in accordance with the EMA directive but there were a few exceptions. In Section 1, RN kept the subtitle *Wie moet Geodon innemen?* (Who should take Geodon?), which is not common in Dutch PILs. Section 2 was not very coherent since not all the segments on who should not take Geodon were formulated addressing the reader directly. First, RN uses the *you* construction *als u aan dementie gerelateerde psychose lijdt* but then proceeds in the same enumeration with a general description of elderly patients (*bij oudere patiënten die lijden aan een …*). Moreover, the QRD segment title *Wees extra voorzichtig met* (Be extra careful if) was not added to Section 2.

With regards to microlevel restructuring, RN applied the following strategies: bullet points for all enumerations, Dutch imperatives for instructions even when expressed by *it is best* in the ST, consistent placement of the instruction first and then the explanation. For the latter strategy, RN explicitly tapped into her declarative knowledge by saying "that is how we learned it should be". She also verbalized the need for consistency in word choice (substitution). She checked if was using the same words for, for example, heart rhythm abnormality and medication. RN applied all these strategies in the writing phase of the translation process.

RN only added information to the TT in the post-writing phase. Noteworthy is that she did not explain a single medical term by adding descriptions. She did add the obligatory section titles and standard phrases while restructuring the translation in the post-writing phase. The TAP and logging data did also reveal cases of conscious and unconscious omissions. Like CB and SA, RN consciously omitted the temperatures in Fahrenheit - at the end of the writing phase- because she believed that only the temperatures in Celsius were relevant for the target audience. As mentioned above, RN sometimes used standard phrases as translations of ST sentences but this text production strategy resulted in the omission of some ST elements (see the pregnancy example in SA's working profile). RN also made several remarks on the occurrence of TT redundancy as a result of high degrees of similarity between the QRD standard phrases and the translations of other ST sentences. She omitted the most obvious redundancies but when in doubt she decided not to omit redundancies. In the QRD template no mention is made of a poison control centre, so RN decided to omit this term. In view of these redundancies or ambiguities, RN stressed the need to contact the client in case of translation problems or ambiguities. We could also observe some unconscious omissions in the log data, e.g. the ST paragraph on how to use Geodon. Given RN's working method of first translating the ST and then editing the translation, it is surprising that RN completely forgot to translate this paragraph. However, TAP data indicated that she did not register this mistake, not even when drafting Section 3 of the TT. This unconscious omission demonstrates that not all unconscious actions (that is, activities that do not appear in the TAP) are automatized, but that they are sometimes due to lack of concentration, overload of the working memory or other events (Gile, 2005).

Participant 4 (KL)

In the overall translation process, KL's working profile greatly resembled CB's working profile. KL drafted the introductory table and index before translating the first ST paragraph for TT Section 1. Subsequently, she skipped the second and third ST paragraphs and proceeded with the remaining ST paragraphs. However, in contrast to CB, KL followed the TT structure in her translation process. She took the QRD template as a starting point and filled in the TT sections with the translations of the corresponding ST paragraphs. She adjusted the TT layout while drafting the TT. At the end of her translation process, KL translated the second ST paragraph and reread the final TT. Neither the TAP nor the logging data gave any indications as to whether KL checked whether all the ST information was included in the TT. Microlevel restructuring was hardly observed in the TAP and the logging data. In terms of taking into account TT usability and readability, KL only used bullet points to make enumerations more legible (e.g., in Sections 2 and 4).

Addition was present in KL's pre-writing and writing phases. She added the introductory table, and index and standard phrases from the QRD template. However, some verbalizations and logging data revealed that KL did not fully comprehend the content of QRD template. In the template provided to the participants, obligatory standard phrases appeared alongside exemplary sentences which are not mandatory. KL included all the sentences mentioned in the QRD template. For example, the sentence *Geneesmiddelen dienen niet weggegooid te worden via het afvalwater of met huishoudelijk afval. Vraag uw apotheker wat u met medicijnen moet doen die niet meer nodig zijn. Deze maatregelen zullen helpen bij de bescherming van het milieu* (Dutch for *Medicines should not be disposed of via* *wastewater or household waste. Ask your pharmacist how to dispose of medicines no longer required. These measures will help to protect the environment*). This is a standard QRD-template phrase, which is often omitted in Dutch PILs and which the other participants did not include in the TT. KL did not explain the difficult medical terms in her translation. What she did add, was information on the excipients, the form and content of the capsules, marketing authorization holder, manufacturer and the date of last approval of the PIL. However, the translation brief stated that some information elements necessary for the TT were not included in the ST but that the translator should not try to find this information elsewhere. The other three participants did not include this information nor did they try to locate it on the Internet. Another difference is that KL correctly placed information into TT Section 2 that CB and SA did not. She correctly added to segment *Gebruik Geodon niet* (Dutch for *Do not use Geodon*) that people younger than 18 years old should not use *Geodon*. CB and SA added this information to Section 4 and the end of Section 2 respectively. RN also included this warning in the first segment of Section 2 but added a Dutch translation of the explanation given in the ST. In contrast to the other three participants, KL correctly added *als u diabetes heeft* (if you have diabetes) to the enumeration of segment *Wees extra voorzichtig met Geodon* (Take special care with Geodon) in Section 2. In this case too she omitted the explanation as to why patients with diabetes have to take special care when using *Geodon*.

Furthermore, KL generously omitted ST units for considering it "irrelevant information". For example, she - rather dubiously - decided to omit the sentence *Elderly patients with a diagnosis of psychosis related to dementia treated with antipsychotics are at an increased risk of death when compared to patients who are treated with placebo*. For the sake of brevity, she also summarized the warning for patients who have certain heart diseases such as long QT syndrome, a recent heart attack, severe heart failure or certain irregularities of heart rhythm with the following translation: *als u een hartziekte of hartritmestoornissen hebt* (if you have a heart disease or heart rhythm irregularities). This rather free approach can be observed in various other places in the TT. Sometimes KL gave her own – often incorrect or partly correct – interpretations of the ST. An example is the translation of *[if you] have had any problem with fainting or dizziness* with *als u problemen hebt ondervonden met uw bloeddruk* (if you have had any problem with your blood pressure). KL also frequently opted for sentences mentioned in the QRD template as translations of ST sentences, just like CB did. However, unlike CB, KL often left out elements which were present in the original ST sentences. According to the TAP data, KL did not appear to notice these inconsistencies or omissions.

Overall, the conscious use of several transediting strategies was observed in all of the four participants' translation processes although they used the strategies to different extents (see Table 9).

When we look at these data in absolute terms, the participants with both declarative and procedural knowledge seem to refer more frequently to the need for transediting than the participants with only declarative knowledge. Nevertheless, these absolute numbers may – for various reasons – give a distorted picture. First, the absolute numbers should be related to the total numbers of TAP segments since it is likely that a participant with a high degree of verbalization will refer more frequently to transediting too. The relative numbers paint a different picture: of the participants' total verbalization volumes, 30.63% of KL's volume and 25.18% of NR's volume refer to transediting. In contrast, 21.55% of CB's volume and 17.15% of SA's volume refer to transediting. A possible explanation for this – at least for CB – is that the participants with both declarative and procedural knowledge might have internalized and automated this knowledge to a certain degree. Therefore there might be fewer traces of transediting strategies in their verbalizations. The numbers presented in Table 9 do not take into account all the actual occurrences of transediting strategies in the translation process. They reflect only the degree of conscious awareness of (the need for) transediting. Second, these data do not state whether those strategies are used correctly or not. The analysis of the quality of the transediting strategies used – as well as the quality of the overall TT – was beyond the scope of this study. However, TT quality and the effect of transediting strategies certainly ought to be taken into consideration in further research.

Table 9. Types and numbers of transediting strategies

Transediting strategy	**participants**			
	CB	**SA**	**RN**	**KL**
restructuring				
macrolevel	17	16	17	9
microlevel	3	0	3	0
substitution	3	0	3	2
addition				
standard phrases	61	46	40	25
explanation medical terms	8	10	0	0
omission	4	3	5	2
total	**96**	**75**	**68**	**38**

Despite the reservations described above, the absolute numbers do shed light on the difference in focus. Especially CB consciously took into account a variety of transediting strategies, with a frequency clearly above the average. The awareness to introduce the obligatory standard phrases into the TT and the restructuring of the ST are striking among all participants. No clear-cut differences between the two participants groups can be seen with regards to the use of microlevel restructuring. However, it is interesting to see that the participants with only declarative knowledge of PILs did not pay any attention explaining difficult medical terms, which suggests that they did not take into account the knowledge level of the TT recipients.

In terms of phase allocation of transediting strategies, three of the four participants referred most frequently to the transediting strategies during the writing phase. Only RN discussed and used far more transediting strategies in her post-writing phase than the other participants. KL resorts to the use of transediting strategies almost exclusively during the writing phase. These inter-individual differences are indicative of the different working styles.

5.4. Methodological considerations and further research

Triangulation proved to compensate the shortcomings of the research methods individually and complemented each other. The qualitative findings helped elaborate the quantitative results, because as Jakobsen (1999) pointed out, keystroke logging data "do not speak until we begin to theorise about them" (p. 15). From the log data in Table 10 we could theorize that the participant made the correction of *ge-*, which she replaces with *inneemt* (take), based on stylistic preferences. However, the TAP provides us with an explanation that the participant herself verbalized: "Since *Geodon* is a capsule, the verb *innemen* (take) would be a more logical choice than the verb *gebruiken* (use).

Table 10. Example of how TAP can complement keystroke logging data

Start	TAP	End	Start	Inputlog	End
05:44	2 wat u moet weten (----) voordat u (----) geodon capsules (----) in (--) gebruikt	05:57	05:43	{3422}	05:46
05:57	(2.0)	05:59	05:46	Wat·u·moet·weten·voordat·u·Geo don·capsules·ge	05:58
05:59	hm (----) en het kan inneemt zijn (----) omdat het een capsule is (----) hoe wordt	06:11	05:58	{4407}	06:01
			06:01	[BS2]inneemt[ENTER]	06:06
			06:06	{3047}	06:09

Although TAP creates extra pauses and stretches them, it proved useful to understand the nature of pauses in text production. TAP data indicated, for example, if the participant was consulting reference works for an equivalent word in the target language. This was particularly helpful for the analysis of the *Translog* log data since no keystrokes or mouse movements outside the *Translog* program were recorded and it was not clear from the log data alone why a participant stopped generating text. *Inputlog* reflects all keystroke activity (also outside the MS Word document) but without providing information about the software program that the participant was working in.[18] It was therefore impossible to track which particular dictionary or website that the participant was consulting. This can be remedied in further research, by using screen recording software such as *Camtasia*. This tool can provide valuable information about text production strategies too. On several occasions, the TAPs suggested that the participants used phrases encountered on the Internet as the translations for ST sentences. This can be easily checked if screen images are available. In addition, we would suggest using a video camera to register the participants' movements and expressions. This may be helpful as an additional source when processes are unclear from the other data sources alone. For instance, for the restructuring of the ST it was sometimes unclear whether the participant made her decision by looking at the QRD template or by tapping into her knowledge of the text type.

In the pilot study we chose to use two keystroke logging programs to discover which program would be better suited for this kind of studies. Both programs are easy to use and provide invaluable insights into the translation process. Translog has been designed with the main purpose of logging translation processes. It does this by creating an environment in which both (fragments of) the ST and the emerging TT are presented. However, we have come to prefer *Inputlog* for several reasons. First, we found *Inputlog* to be less intrusive since it allowed participants to work in, for instance, an MS Office environment using MS Word, Internet Explorer and online dictionaries just like they would normally do. This, of course, increases the ecological validity of the experiment. Second, *Inputlog* registers in which software program (in the Windows environment) the participant is working whereas *Translog* logs only internal activities. Third, the statistical data provided by the two programs vary. *Translog* provides information about user events, numbers of text-production and text-elimination keystrokes, cursor navigation, mouse events and miscellaneous events, as well as information about total duration, user events per minute and text production per minute. *Inputlog*, on the other hand, yields those statistical data[19] and more, such as information about the numbers of segments, pause time, average writing time and word length.

Although triangulation provided rich data on transediting it was disappointing that in many instances the TAP did not provide further insights into why the participants translated ST segments the way that they did. This was probably caused by cognitive overload and fatigue. In order to avoid these problems, the method of immediate (prompted) retrospection might be more suitable than concurrent TAP. With retrospective protocols, the risk of cognitive overload during the translation process will be reduced. Moreover, retrospection does not affect the duration, the segmentation and the nature of the translation process itself. Furthermore, the contemporary logging programs have solved the presumed lack of reliability of retrospection (i.e., participants tend to forget what they have done and give dis-

[18] The latest versions of *Inputlog* (from 4.1 onwards) do represent this information.

[19] *Inputlog* does not provide explicit data on total user events or text elimination. However, in *Inputlog 4.1.* text elimination can be investigated more profoundly using S-notation. It is also important to note that there is a slight difference between *Translog* and *Inputlog* in the treatment of text-production activity. *Translog* considers all keystrokes – including spaces, punctuation marks and hard enters – as instances of text production whereas *Inputlog* – in analogy with MS Word – does not take into consideration hard enters.

torted views of their translation processes). They offer the possibility to replay the translation process which serves as stimulated recall. It also manifests discrepancies between the participants' verbalizations and the actual translation processes logged by the program. Nevertheless, it might still be possible that the participant gives extremely general comments about the underlying motivation to use specific text-production strategies in general and transediting strategies in particular. Hansen (2006) has conducted a study using prompted retrospective dialogue. She not only received clarification but was also able to provide the participants with feedback to increase their insights into the phenomena of interest even more. In experiments that do not have a didactic purpose, retrospective dialogues may be substituted for retrospective interviews. In so doing, the researcher can still focus on particularly interesting points in the translation process (see Leijten, 2007). Nevertheless, the advantages and disadvantages of retrospection for the specific experimental design have to be ascertained: Does retrospection actually yield more relevant data? Do participants have trouble remembering what they did despite the stimulated recall by replaying the log file?, etc. The length of the ST used in the experiment will also have to be decreased if retrospection is used.

The pilot study also highlighted the difficulty of triangulating TAP and computer keystroke logging, particularly in terms of data analysis. Publications on mixed-method experiments in empirical translation-process research rarely give information about how triangulation was carried out in the analysis of the various data sources and about which method was taken as a starting point in the analyses of data. With this paper, we have designed a working document in which TAP data are placed alongside log data. This allowed us to analyse the two data sources comparatively (as shown in Table 11).

The juxtaposition of TAP data and log data is a first step to improving triangulation in the data-analysis phase but in the current working document, the TAP segments and log segments are not completely aligned. This has two major implications: (1) it complicates the coding of the segments and (2) the total number of segments of the two data sources vary, which makes it difficult to establish the quantitative interindividual incidences of coding labels. The number of segments with which the TT is produced has, of course, been calculated according to the log file but the coding of the cognitive processes has been based primarily on the TAP segments. Consequently, statistical analysis is problematic. Given the limited number of participants[20] and the exploratory and descriptive nature of this study, statistical analysis was not an issue. However, the segmentation of the working document is a challenging factor which could be remedied in further research. It might be argued that TAP-based coding is sufficient. However, in the log data we encountered some examples that represent the use of transediting strategies which were not mentioned in the TAP, i.e. automatic use of the transediting strategy. A solution to improving the working document and its possible statistical uses would be to split the existing TAP and log-file segments and to align them fully. However, this would imply re-

[20] This use of a small sample of participants appears to be a critical point in the majority of translation-process studies. Having few participants is nevertheless logical, because using both TAP and computer keystroke logging provides an ample amount of data, which are almost impossible to handle in case of dozens of participants. However, in order to be able to generalize findings larger populations are necessary.

specting the pauses in the TAP while segmenting the log data and vice versa. This would make the transcription and segmentation of the TAP data and the log data even more time-consuming. Therefore, it might be sensible to adopt a more pragmatic approach especially for experiments with a larger sample of participants. In that case, researchers could design a working document, using the method proposed above, but only zooming in on critical incidents.

Table 11. Working document for comparative analysis for TAP and log data

Start	TAP	End	Code	Code	Start	Inputlog	End
21:50	hm een psychose zoals geestes	21:58	TARGET/PROD/CONCRETE/LETT	TARGET/RED/EMEA/INHOUD/UIT-LEG/INTMIS	21:50	een·psychose·zoals·geestesw[BS]zkt[BS2]wakte·werd·vastgesteld {	22:05
21:58	(5.0)	22:03					
22:03	zwakte werd vast-gesteld	22:06	TARGET/PROD/CON-CRETE/LETT				
22:06	(3.0)	22:09			22:05	{3563}	22:08
22:09	en be-handeld werd met (----) hm	22:13	TARGET/PROD/CON-CRETE/LETT		22:08	·en·behandele·d[BS3]d·werden·met·	22:15
22:13	(3.0)	22:16					

Although the pilot study has provided important insights into transediting, more research is needed from a variety of perspectives. For example, it would be interesting to take into account various text types. The degree of (permissible) transediting will vary considerably according to the text type. On the basis of interviews with several professional translators specialized in the translation of PILs, we have found that the degree of transediting displayed in the pilot study would not be feasible in professional environments. Pharmaceutical companies, for example, give professional translators very little room to improve TTs or even make them correspond with the EMA directives. Another focal point would be translations students' transeditorial consciousness, as suggested by Stetting (1989). This is certainly an important aspect in light of translation pedagogy, but also considering students' future as professional translators. Unfortunately, they will almost certainly be confronted with poorly composed STs and should thus be adequately prepared to transedit to avoid a 'garbage in, garbage out' situation.

Lastly, the quality of the TT should be taken into consideration when investigating translation processes. This will make observations about segmentation and working profiles more complete and profound. It might also yield interesting data which could be used in translation pedagogy. The evaluation of translation quality is a thorny subject among scholars. Nevertheless, we feel that by using several expert raters to evaluate the TT according to a well-established model of evaluation criteria, subjectivity can be controlled if interrater reliability is high.

6. Conclusion

The overall trend was that participants with both declarative and procedural knowledge of PILs spent less time on the translation process than participants with only declarative knowledge of the text type. However, it has been argued that several parameters have to be taken into account when analysing and discussing processing times and the distributions of times spent on the different phases of the translation process, especially in the event of a translation that goes beyond mere transfer of a ST. The amount of transediting used in the translation process as well as the number of transediting strategies applied varied considerably among participants. However, no clear correlation could be established between the use of transediting strategies and the possession of declarative and procedural knowledge. In terms of the methodology chosen for the pilot study, several suggestions were made to optimize the methodological framework for future research.

REFERENCES

Asadi, P., & Séguinot, C. (2005). Shortcuts, strategies and general patterns in a process study of nine professionals. *Meta, 50*(2), 522–547.

Askehave, I., & Korning Zethsen, K. (2002). Translating for laymen. *Perspectives: Studies in Translatology, 10*(1), 15–29.

Breedveld, H. (2002). Writing and revising processes in professional translation. *Across Languages and Cultures, 3*(1), 91–100.

Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication, 18*(1), 80–98.

Chesterman, A. (1997). *Memes of translation*. Amsterdam: John Benjamins.

Dam-Jensen, H., & Heine, C. (2009). Process research methods and their application in the didactics of text production and translation. *trans-kom: Zeitschrift für Translationswissenschaft und Fachkommunikation, 2*(1), 1–25.

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.

Dragsted, B. (2005). Segmentation in translation: Difference across levels of expertise and difficulty. *Target, 17*(1), 49–70.

Englund Dimitrova, B. (2005). *Expertise and explicitation in the translation process*. Amsterdam: John Benjamins.

Ericsson, K. A., & Simon, H. (1984 [1993]). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: MIT Press.

Göpferich, S. (2008). *Translationsprozessforschung: Stand, Methoden, Perspektiven.* (Translationswissenschaft 4). Tübingen: Narr.

Hansen, G. (2006). Retrospection methods in translator training and translation research. *JoSTrans, 5*, 2–41.

House, J. (2000). Consciousness and the strategic use of aids in translation. In S. Tirkkonen-Condit & R. Jääskeläinen (Eds.), *Tapping and mapping the processes of translation and interpreting* (pp. 149–162). Amsterdam: John Benjamins.

Hursti, K. (2001). An insider's view on transformation and transfer in international news communication: An English–Finnish perspective. *The Electronic Journal of the Department of English at the University of Helsinki, 1*. Retrieved from http://blogs.helsinki.fi/hes-eng/volumes/volume-1-special-issue-on-translation-studies/an-insiders-view-on-transformation-and-transfer-in-international-news-communication-an-english-finnish-perspective-kristian-hursti/

Jääskeläinen, R. (1999). *Tapping the process: An explorative study of the cognitive and affective factors involved in translating*. Joensuu: Joensuun Yliopisto.

Jääskeläinen, R. (2002). Think-aloud protocol studies into translation: An annotated bibliography. *Target, 14*(1), 107–136.

Jakobsen, A. L. (1999). Logging target text production with Translog. In G. Hansen (Ed.), *Probing the process in translation: Methods and results* (pp. 9–20). Copenhagen: Samfundslitteratur.

Jakobsen, A. L. (2002). Translation drafting by professional translators and by translation students. In G. Hansen (Ed.), *Empirical translation studies: Process and product* (pp. 191–204). Copenhagen: Samfundslitteratur.

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* (pp. 69–95). Amsterdam: John Benjamins.

Jakobsen, A. L., & Schou, L. (1999). Translog documentation. In G. Hansen (Ed.), *Probing the process in translation: Methods and results* (pp. 151–186). Copenhagen: Samfundslitteratur.

Krings, H. (2001). *Repairing texts: Empirical investigations of machine translation post-editing processes*. Kent, OH: The Kent State University Press.

Krings, H. (2005). Wege ins Labyrinth: Fragestellungen und Methoden der Übersetzungsprozessforschung. *Meta*, *50*(2), 342–358.

Künzli, A. (2003). *Quelques stratégies et principes en traduction technique français–allemand et françois–suédois*. Unpublished doctoral thesis, Stockholms universitet, Stockholm.

Künzli, A. (2004). Risk taking: Trainee translators vs professional translators: A case study. *JoSTrans*, *2*, 34-49.

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 keystroke logging and writing: Methods and applications*, pp. 73-94). Oxford: Elsevier.

Leijten, M. (2007). *Writing and speech recognition: Observing error correction strategies of professional writers*. Utrecht: LOT.

Lorenzo, M. P. (1999). Apuntes para una discusión sobre métodos de estudio del proceso de traducción. In G. Hansen (Ed.), *Probing the process in translation: Methods and results* (pp. 21–42). Copenhagen: Samfundslitteratur.

Lörscher, W. (1993). Translation process analysis. In Y. Gambier & J. Tommola (Eds.), *Translation and knowledge. SSOTT IV* (pp. 195–211). Turku: University of Turku.

Mossop, B. (2001). *Revising and editing for translators*. Manchester: St. Jerome.

Stetting, K. (1989). Transediting: A new term for coping with the grey area between editing and translating. In G. Caie, K. Haastrup, A. L. Jakobsen, J. E. Nielsen, J. Sevaldsen, H. Specht et al. (Eds.), *Proceedings from the fourth Nordic conference for English studies* (pp. 371–382). Copenhagen: University of Copenhagen.

Van den Bergh, H., & Rijlaarsdam, G. (1996). The dynamics of composing: Modelling writing process data. In C. M. Levy & S. E. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 207–232). Mahwah, NJ: Lawrence Erlbaum.

Vuorinen, E. (1996). *Crossing cultural boundaries in international news transmission: Translational approach*. Tampere: University of Tampere.

Wengelin, A. (2006). Examining pauses in writing: Theory, methods and empirical data. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer keystroke logging: Methods and applications* (pp. 107–130). Oxford: Elsevier.

artesis ***VT*** *working papers*

in translation studies

2010-1 **Luyckx B., Delbeke T., Van Waes L., Leijten M. & Remael A.**
Text Reduction in Live Subtitling with Speech Recognition.
Causes and consequences of text reduction.

2010-2 **Van Rietvelde S., Eyckmans J. & Bauwens D.**
Ahead of the game.
Phraseological competence and linguistic anticipation in the interpreting performance.

2011-1 **Schrijver I., Van Vaerenbergh L. & Van Waes L.**
Transediting in students' translation processes.

2011-2 [march 2011] **Reviers N. & Remael A.**
Audio Description and Translation Studies.
An attempt at defining AD in terms of text types.
[provisional title]

Artesis VT working papers is dedicated to research in progress / in press within the broad field of translation and interpreting studies.

The series intends to be a platform for scholars young and old and reaches beyond institutional boundaries to further key concepts of translation and interpreting studies and to boost dialogue between researchers.

Please send proposals as abstracts (300-500 words) to geraldine.devisscher@artesis.be.

Artesis VT working papers in translation studies
are published on-line at

www.artesis.be/vertalertolk/onderzoek/artesis-vt-working-papers-in-translation-studies.htm