



Master

2011

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Translation Technology for the Federal Chancellery - The Usefulness of a Translation Memory System for the German Section of the Central Language Services

Walpen, Nina

How to cite

WALPEN, Nina. Translation Technology for the Federal Chancellery - The Usefulness of a Translation Memory System for the German Section of the Central Language Services. Master, 2011.

This publication URL: <https://archive-ouverte.unige.ch/unige:18409>



UNIVERSITÉ
DE GENÈVE

ÉCOLE DE TRADUCTION
ET D'INTERPRÉTATION

MA Thesis

TRANSLATION TECHNOLOGY FOR THE FEDERAL CHANCELLERY

THE USEFULNESS OF A TRANSLATION MEMORY SYSTEM
FOR THE GERMAN SECTION OF THE CENTRAL LANGUAGE SERVICES

ETI, Université de Genève

Département de traitement informatique multilingue

Thesis Director: Prof. Pierrette Bouillon

Thesis Jury: Marianne Starlander

Nina Walpen

ni.w@gmx.ch

June 2011

Table of Contents

List of Figures.....	3
List of Tables.....	3
1 Introduction.....	4
2 Translation Memory Systems – an Introduction.....	8
2.1 The Translator's Working Environment.....	8
2.2 Translation Memory Technology.....	11
2.2.1 TM – a brief history.....	12
2.2.2 Basic Functionalities.....	14
2.2.3 Impact on translators and translation.....	19
3 Context: Central Language Services.....	23
3.1 ZSD-D.....	24
3.2 Environment.....	25
3.3 Texts.....	27
3.4 Content analysis of Parliamentary Interventions.....	29
4 The tests.....	32
4.1 Evaluation and methodology.....	32
4.1.1 Translation tool evaluation.....	33
4.1.2 Methodology.....	36
4.2 The tests.....	39
4.2.1 Overall arrangement.....	40
4.2.2 Alignment tests.....	43
4.2.3 TM size.....	59
4.2.4 The scenario test.....	62
4.3 Other factors.....	73
5 Conclusion.....	79
6 Bibliography.....	82
7 Annexes	87
Annex A – The interviews.....	87
Annex B – Statistics ZSD-D.....	97
Annex C – Overview tests.....	98
Annex D – The texts.....	100
Annex E – Scenario test task list.....	101
Annex F – Scenario test questionnaire.....	102

List of Figures

Fig. 1: Human and machine translation (Hutchins & Somers, 1992: 148).....	10
Fig. 2: Database Model (Zerfass, 2002b).....	15
Fig. 3: Reference Model (Zerfass, 2002b).....	16
Fig. 4: Statistics 2009.....	27
Fig. 5: Statistics other translations (besides Parliamentary Interventions) 2008.....	27
Fig. 6: Repeated Text, threshold 0%.....	30
Fig. 7: Repeated Text, threshold 75%.....	30
Fig. 8: Statistics test corpus.....	42
Fig. 9: WinAlign alignment error (99.3309).....	54
Fig. 10: TextBase alignment error (99.2209).....	55

List of Tables

Table 1: Summary Compliance.....	45
Table 2: Texts for Alignment (5).....	46
Table 3: Summary Efficiency.....	47
Table 4: Summary Reliability.....	47
Table 5: Segmentation error.....	48
Table 6: Alignment error.....	48
Table 7: Correct segmentation and alignment.....	48
Table 8: Summary Maintainability.....	50
Table 9: Summary Customizability.....	52
Table 10: Summary Accessory Information.....	52
Table 11: Results Compliance.....	53
Table 12: Results Efficiency.....	54
Table 13: Results Reliability.....	54
Table 14: Results Maintainability.....	56
Table 15: Results Customizability.....	57
Table 16: Results Accessory Information.....	57
Table 17: Alignment quality (improved).....	59
Table 18: Search terms SDL Trados (Context search).....	60
Table 19: Source texts for the scenario test.....	64
Table 20: Task sequence.....	64
Table 21: Result presentation.....	65
Table 22: Results time.....	67
Table 23: General appreciation (ranking).....	71
Table 24: Overall ranking.....	71

1 Introduction

Perhaps more than other professionals, translators are feeling the long-term changes brought about by the information age. The snowballing acceleration of available information, the increase in intercultural encounters, and the continuing virtualization of private and business life have resulted in drastic and lasting changes in the way translators work.¹

In 2009, I was invited to do a six weeks' translation internship at the Swiss Federal Chancellery in Berne, in the German language section of the Central language services (ZSD-D). At that time I had only just completed the first year of my MA in Translation at the École de traduction et d'interprétation in Geneva. The internship was therefore a very good opportunity for me to gain some work experience and see if I would really like working as a translator in a full time job. During the internship I helped the service translate French written parliamentary interventions into German. However, during my time there, together with a colleague, I was also asked to briefly present translation memory technology to the members of the service. Up to that moment, the service did not work with any translation tool and indeed, most of the translators there had never worked with one. Markus Nussbaumer, head of the ZSD-D, wanted to know more about these tools and how exactly they could be beneficial for his service, and in discussion with him we agreed that I would look more closely into the question, as part of my MA thesis. The official consent came in December 2009, just before the Christmas break. The basic questions underlying this project were clear: Would it be useful for the service to acquire such a tool? Would the translators and the service benefit from working with it? And what are the elements which need to be considered in case an acquisition is decided, what measures need to be taken?

These questions are legitimate of course, if only because translation technology is becoming more and more important for translation professionals everywhere. Many authors note that the 'business' of translation is changing because our world is becoming more and

1 Austerlühl, 2001: 1.

more globalised, and this brings with it an increased demand for translation.² For example, the Translation Centre for the Bodies of the European Union translated 199'118 pages in 1998. Five years later, in 2003, the number had more than doubled to 238'399 pages, and three years ago, in 2008, this number had jumped to 747'416 pages.³ These numbers, if not necessarily representative of the translation 'business' in general, are at least symptomatic. At the same time "there appears to be a demand by translation commissioners and employers for significantly increased speed in completing translation jobs" (Mossop, 2006: 790). Time is money, and fast translations are therefore less costly, even though a high quality output needs to be maintained. This leads to a dilemma for translators who, on the one hand, are supposed to deliver high quality target texts, but on the other hand find themselves under increasing time pressure. According to Lynne Bowker (2002: 12-13), translation technology thus becomes an instrument in the search for balance in this situation: It helps translators to work faster while it keeps the output quality high, for example through terminology resources or reuse of previously translated material. This development is reflected by the fact that translation technology is increasingly adopted in the curricula of translation degrees.

In the private sector, translation technology and particularly translation memory (TM) systems are already deeply rooted and even freelance translators are often expected to work with one when they collaborate with translation agencies or other providers. According to a survey carried out in 2006, 82.5 percent of respondents were working with a TM system.⁴ Interestingly, this rate is already high for freelancers (81%), but even higher for company employees (84%) and company owners (92%). According to the author of the survey, Elina Lagoudaki, company owners were more open to adopting TM technology, because they are "convinced about the cost savings and productivity gains deriving from the use of TM systems" (Lagoudaki, 2006: 15). These rates seem relatively high and could be due to the fact that a large part of the respondents are working with technical texts, for which this technology is generally more suitable than for others. For translators of general or literary texts, this rate was indeed much lower. On the other hand, five years in technology progress is a long time and the rate could also have gone up since the survey was carried out. Overall it therefore seems clear that TM systems are already an integral part of the translation business

2 See e.g. Austermühl (2001), Bowker (2002), Mossop (2006), Quah (2006).

3 Among other factors, this extreme increase is most certainly due to the enlargement of the EU over the last decade and the new languages into which documents therefore need to be translated. 12 new states have joined the EU in 2004 and 2007, but numbers have also increased in times without enlargements. The statistics are available on <http://www.cdt.europa.eu/DE/documentation/Pages/Translation-Statistics.aspx> (19.04.2011). See also Austermühl, 2001: 3.

4 Translation Memories Survey 2006, carried out by Elina Lagoudaki for the Imperial College London, the survey is available at <http://www3.imperial.ac.uk/pls/portallive/docs/1/7307707.PDF> (24.04.2011).

today and consequently it is justified for any translation service to inquire if it would benefit from the introduction of such a tool.

Generally, the benefits of a TM system for translation professionals concern the translation speed and therefore productivity and cost savings, but also the quality of the translation output. In this they really do respond to the above mentioned translators' dilemma. The participants of the TM-survey also explained their reasons for working with this technology (Lagoudaki, 2006: 22. Multiple answers were possible): it saves time (86%), the consistency in terminology is improved (83%), and the quality of the translation output is improved (70%). Other reasons mentioned were cost savings (34%) and that it represents the best way to exchange resources such as glossaries or TM databases (31%). Also, 71 percent of non-users showed themselves willing to try or even buy a TM system in the near future (Lagoudaki, 2006: 17). On the other hand, when interrogated about the reasons for which they did not adopt such a system, only 8 percent mentioned that they had tested or used one in the past and found that it was not useful to them. It seems therefore that the tools are generally well received by translators, freelancers as much as company employees and owners, and that they would not go back to a state before buying the tool.

Generalisations are always dangerous of course and it still cannot be said that the TM systems are suitable and pertinent to all environments. Even in the above survey, almost a third of non-users stated that the tool was not suitable for their work (Lagoudaki, 2006: 17). Each context, then, needs to be studied and analysed before a definite statement can be made, and this is exactly what this paper intends to do for the ZSD-D. As we shall see, the ZSD-D is not just a governmental translation service, it has many other duties and tasks to accomplish, which do not directly involve translation but rather writing, editing or revising. In that the situation of this service is quite specific, and any tests or conclusions need to take these elements into account. Knowing the texts translated in the service from my internship, it is clear that they are not very repetitive. TM systems, however, normally work best with texts showing high repetition. This paper therefore not only investigates the specific context of the ZSD-D, but also a more general one: Would a TM tool still be useful in a service translating texts without high repetition and would it still deliver the benefits mentioned generally with TM systems. According to Lagoudaki (2006: 16) this is possible, probably because the systems today are less dependent on retrieval of entire sentences or passages which had already been translated, but can offer other solutions as well, for example terminology management. This paper then sets out to investigate if this is the case also for such translation services in general.

Furthermore, this paper will be limited to Translation Memory Systems and not test Machine Translation (MT) or other tools from the vast domain of translation technology. The

reason for this choice lies on the one hand in the texts translated in the Federal Chancellery (FC), which are very varied and for which it would be very difficult to specialise an MT tool. On the other hand, the fact that the translators in the service have little experience with translation tool means that at least some resistance is to be expected. It seems probable that a tool meant to assist them in their work will be better received than one which they fear will replace them.⁵ For these reasons I have decided to test only Translation Memory technology for the context of the ZSD-D.

Chapter two will explain what TM tools really are and how they work. This is important in order to understand how this technology can impact a translation environment and where exactly the promised benefits come from. It will also treat the question of how such tools are generally received by translators and the impact they can have on them and on the translation process. Chapter three then goes on to explore the context of the Central language services and the ZSD-D in particular. As we shall see, some sections inside the Central language services already work with a TM system. The chapter also features a detailed analysis of the texts translated in this service, in order to find out more about the extent of repetition in the texts. Chapter four contains the practical part of this paper: some basic information on the evaluation of translation tools and methodology, as well as the tests that have been implemented for this study. These tests are divided into three parts, each one featuring the setup, the results and their impact. At the end of chapter four, the conclusions which can be drawn from the tests and the factors which have to be considered before a decision on the acquisition of a tool is taken are presented. Finally, the conclusions then present some recommendations in case the ZSD-D decides to adopt TM technology for their service, as well as answers to the more general questions guiding this study.

5 The issue of translators' reception of tools and resistance will be treated in chapter 2.2.3, p.19.

2 Translation Memory Systems – an Introduction

"Never translate the same sentence twice."⁶

This chapter specifically aims at providing basic background information on Translation Memories (TM) with respect to terms and concepts used later on in this study. Also, certain concepts need to be understood to be able to set up appropriate tests for our purpose. In the first section, we will define some important terms and situate the tools examined here within the wide field of translators' aids. We will then move on to the main focus of this study, namely the Translation Memory System, tracing the history of its developments, describing functionalities of and attitudes towards it, as well as the impact this technology has or could have on translators and the translation process. This chapter is mainly based on Bowker (2002) and Quah (2006), where other sources were used they are mentioned specifically.

2.1 The Translator's Working Environment

The times when translators were working with pen and paper, as well as the resources of an extensive library, seem very far away today. Indeed, there are many electronic resources available to a present-day translator which could hardly have been imagined by anyone only a few decades ago, including the translation memories under study here. Nowadays, a translator's working environment, generally called workstation or workbench, includes not only text processing and electronic dictionaries and glossaries, e-mail and online telephone services, but also bilingual corpora with concordancers, translation memories or even machine translation programmes. All these tools can be classified according to different criteria. I will briefly present two classifications here which are helpful in order to situate the

⁶ Publicity slogan for a TM System, quoted by Chama, 2009: 48.

translation memories in this wide field of translation aids. The first one is the functional classification proposed by Alan Melby as early as 1982. The second one, put forward by John Hutchins and Harold Somers (1992), is slightly more recent and classifies the tools according to their degree of automatisation (Austermühl, 2001: 9).

Already in the 1980s, Alan Melby (1982) found that translator's aids can be divided into three groups, representing three consecutive levels. This was at a time when a fully automatic translation of general texts with high quality output was "not on the visible horizon" (215). Melby noticed that many translators were not satisfied with their role as post-editors for machine-translated texts and he was convinced that "an unhappy translator is a serious problem" (216). His solution was, then, to place the translator in control, a point in which he agreed with Martin Kay, author of a memorandum which is today considered to be "one of the most decisive moments in the development of the future translator's workstation" (Hutchins 1998: 8)⁷. Melby therefore suggested to classify the available tools in three levels between which the translator could freely switch while always staying in control in that s/he always decides on how to proceed with a specific text or part of text.

On the first level Melby situated text processing with independent or integrated terminology databases, as well as telecommunications. These could be used even if the source text was not in electronic form. For level two this would have to be the case, however, as it included automatic text processing of the source text, such as automatic lookup of words, which Melby called the "suggestion box option" (218). Level three, finally, included an integrated "full-blown MT system" (218), which the translator could resort to, but only if and when s/he wanted to. This first version of the classification did not include bilingual concordancers or translation memories yet, but Melby later refined his classification (Melby, 1992), adding text analysis as well as synchronized bilingual text retrieval to level two. The latter, according to his description (163), corresponds more or less to today's translation memories. In this classification, then, the systems examined in this study are somewhere in-between basic text processing and machine translation. The source text already needs to exist in electronic form, but the translator still remains in charge of the translation process, deciding whether or not, and at which moment to resort to machine translation.

The second classification, put forward by John Hutchins and Harold Somers (1992), is probably the one used most often today. It situates the tools included in Melby's classification in a wider context and groups them according to whether the machine or the person is carrying out the translation, in other words according to the degree of automatisation (see fig. Fehler: Referenz nicht gefunden). At the automatised end, FAHQT stands for Fully Automatic

⁷ See Kay (1980), as well as section 2.2.1, p.12.

High Quality (machine) Translation or simply Machine Translation (MT).⁸ This corresponds to the original ideal in the 1950s and 60s of a translation system with no human involvement whatsoever. The opposite extreme would be a human translation where no electronic resources are used. In their extreme version, both of these are not very realistic today. Working without any electronic tool, not even electronic dictionaries or the internet, cannot be fast enough in a world where time pressure on translators is ever increasing. Also, the idea of a universal translation machine with no human involvement is generally understood to be unrealistic today, except for very specific contexts (as for the Canadian METEO System⁹) or if concessions can be made on the quality of the target text, for example if it does not necessarily have to be publishable or if only the general meaning of the text needs to be understandable. As Quah (2006: 13) notes, today "the main aim of machine translation is still to generate translation automatically, but it is no longer required that the output quality is high, rather that it is fit-for-purpose."

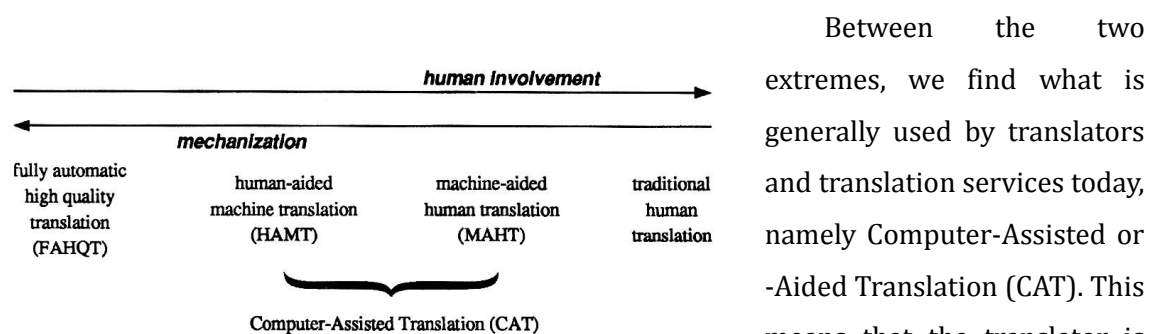


Fig. 1: Human and machine translation (Hutchins & Somers, 1992: 148)

Between the two extremes, we find what is generally used by translators and translation services today, namely Computer-Assisted or -Aided Translation (CAT). This means that the translator is still involved, to various degrees, in the translation process. This group of tools can be divided again into HAMT or Human-Assisted or -Aided Machine Translation and MAHT or Machine-Assisted or -Aided Human Translation. In the first, "the machine carries out most of the work but it might need human assistance either at the text-preparation stage or the output stage" (Quah, 2006: 11). This means that the actual translation process is carried out by the machine translation system, while the translator is responsible for pre- or post-editing the texts. Often, these systems work with a limited number of source text types, with restricted grammar and vocabulary or even controlled language. Also, they are often used for a very limited domain, e.g. technical texts such as legal briefs, manuals or laboratory reports.

With MAHT, on the other hand, the translator remains in charge of the translation process, but s/he is using a number of tools to support him/her. These tools include spell- and grammar-checkers, electronic glossaries and dictionaries, terminology databases and management systems as well as collections of previously translated texts and their source

⁸ There is some ambiguity as to the human involvement in MT, see Quah 2001, p. 7.

⁹ See Hutchins 1995, p.437.

texts, i.e. bilingual corpora and translation memories. Often, these tools are combined in integrated systems, so-called 'workbenches' or 'workstations'. Because of this integration and the fact that the tools become increasingly multifunctional, the distinction between MAHT and HMT becomes more and more difficult to maintain. Certain tools, for example, combine the approaches of Machine Translation and Translation Memory into one tool, thus bridging the gap between the two. They closely resemble what Melby had in mind with his 3-level model. The classification thus cannot entirely account for the situation today, but it is nevertheless a good instrument when it comes to situating electronic translation tools.

With both classifications, the tools examined in this study can therefore be situated somewhere in the middle. In the latter, it could be found in the category MAHT, in the first, it would belong to level two, where the texts already need to be in electronic form but the translator remains in charge. The tools we are interested here represent a kind of compromise, between the use of technology and a need for control, thus responding to a common fear of being replaced by a machine. We will look at this aspect in more detail in section 2.2.3 (p. 19) below. The next step now is to find out more about how exactly this compromise works.

2.2 Translation Memory Technology

As we have seen above, translation memories are today most often integrated into a workbench or workstation, including the translation memory, alignment tools, tag filters, electronic dictionaries, terminology management systems and databases, spell- and grammar-checkers, file converters and so on. But what exactly is a translation memory, and how does it work? The Expert Advisory Group on Language Engineering Standards (EAGLES) defines it as a "multilingual text archive containing (segmented, aligned, parsed and classified) multilingual texts, allowing storage and retrieval of aligned multilingual text segments against various search conditions"(EAGLES, 1996: 140).¹⁰ The basic idea behind it is that previously translated texts can be reused for new translations, so that no sentence is translated twice. This process of reusing older material is generally called "leveraging." Indeed, with all the translations done in the world, the idea of reusing previous translations seems obvious, and the probability that the same or a similar sentence has already been translated somewhere is very high. However, this process needs to be efficient, a translator today cannot afford to spend hours looking through documents in search of one sentence, and this is exactly where the TM Systems help.

¹⁰ For more information on EAGLES see chapter 4.1, p.32.

There are two basic components to a translation memory: the text archive, i.e. the database of previously translated texts, and the software and interface to search it. However, there is no linguistic component which could actually translate text. When working on a new source text, the software compares the new sentence to the texts and sentences already stored in the archive and, if a match is found, presents the translator with the previous translation. The translator can then decide to confirm this translation or modify it for the new target text. Of course, the new sentence can also be translated from scratch. Accordingly, and as we have seen through the classifications above, the translator remains in charge, nothing is translated unless s/he decides. The specific functionalities of the tool will be described in section 2.2 further on, we will now first look at historical developments.

2.2.1 TM – a brief history

The idea of using previous translations as a resource is not very recent. Indeed, even though the number of words and possible sentences in a language is infinite, in reality there are combinations which are more likely to occur and which actually do occur more often than others. Therefore, when encountering a new sentence, the chances that the exact same or at least a similar sentence has already been written and translated somewhere are very high. The archives of previous translations or other texts are a natural consequence of this fact. However, it is only in the second half of the 20th century that an electronic implementation of the idea has become possible thanks to progress in information technology. Serious research into the idea of automatic translation can be traced back to the 1950s, while other tools, including what would later become the translation memory, received more attention from the 1960s (Hutchins, 1995: 434). This was also due to an important report published by the Automatic Language Processing Advisory Committee (ALPAC), commissioned by the US government to evaluate the progress in computational linguistics in general and machine translation in particular. Before the publication of this report in 1966, researchers were convinced that FAHQT was possible, and soon, so that general expectations were very high. However, the conclusions of the ALPAC report stated that nothing developed until that day was as efficient as human translation. Specifically, there was no hope of any useful machine translation system immediately or in the near future (Hutchins, 1995: 436). This report was very disenchanting, of course, and led to many budget cuts in the research of machine translation. However, as Hutchins (1998: 3) notes: “ALPAC was not entirely negative; it encouraged support for basic computational linguistics and, in addition (today often forgotten), the development of computer-based aids for translators.”

The basic idea of the translation memory thus goes back to the 1960s, one of the first tools in this area being developed by the European Coal and Steel Community (ECSC) in Luxemburg. This tool introduced automatic dictionary lookup, but it was essentially a first version of a retrieval tool with a KWIC-display (Reinke, 2004: 36).¹¹ A next step in the development of translation memories was initiated by Friedrich Krollmann in the 1970s. He developed the LEXIS-system, another automatic dictionary lookup-system, at the German Army Translation Services (today Bundessprachenamt). Krollmann imagined that the terminological database could be completed by a 'linguistic databank' including a translation archive (Hutchins, 1998: 5). However, Krollmann was speaking of retrieval only in terms of identical matches, an idea which was finally expanded by Peter Arthern towards the end of the 1970s. He then essentially suggested what is known as a translation memory today. He imagined "translation by text-retrieval," where source and target texts were stored and "any parts of any text" could be retrieved and inserted into the new document (Hutchins, 1998: 7). However, even though the idea was there at the time, computer technology simply was not advanced enough yet to allow for it to be implemented immediately.

In the meantime, Martin Kay's memorandum deserves mention here (Kay, 1980). In agreement with the ALPAC report and Alan Melby, he criticised the idea that machine translation research would eventually lead to a system which would replace the human translator. Rather, he wished for the development of a translator's *amanuensis*, i.e. a workstation where the translator is in charge but has professional tools available for support. This memo not only confirmed general tendencies at the time but also provided an impetus for the development of CAT tools rather than MT. The technical situation was indeed about to change dramatically with the appearance of the first personal computers in the 1980s, while around the same time the statistical means required for text alignment were also improved (Hutchins, 1998: 11-14). The first commercialised translation memory systems were thus released in the early 90s: Transit System (STAR) in 1991, IBM Translation Manager/2 and the Translator's Workbench (Trados) in 1992, ATRIL's Déjà Vu in 1993 and, finally, Eurolang Optimizer in 1994 (Freigang, 2009: 12-22). Others systems followed over the next years, and of course today the list of available systems is much longer, also including some open source systems.¹² Progress has been possible in many areas, most systems have changed considerably from their first version, and there seems to be a general tendency towards standardisation and exchange, thanks to institutions such as the Localization Industry

11 KWIC, "key word in context".

12 A list can be found on Wikipedia: CAT http://en.wikipedia.org/wiki/Computer-assisted_translation (26.01.2011); or in John Hutchins' Compendium of Translation Software (15th ed. January 2009), which includes not only translation memories but also other translation aids: <http://www.hutchinsweb.me.uk/Compendium-15.pdf> (23.02.2011).

Standards Association (LISA). We will not present any of these commercialised tools here, but the next section will present the basic functionalities which are common to most of them.

2.2.2 Basic Functionalities

As we have seen, there are many translation memory tools on the market today, and while they differ with respect to some functionalities, they resemble each other in some basic aspects. In this section we will look at some of these basic functionalities in order to understand how these tools actually work. As we have seen above, a translation memory system basically consists of two components: the database and a software for retrieval, and it can then be integrated into a workstation or workbench, where it is complemented with other tools. The positive impact of the system thus greatly depends on the quality of the database content, but also on the performance of the algorithms used to search it in order to retrieve previously translated text. We will first look at the database, its structure and how it can be filled, before moving on to retrieval. A third section will look at what kind of texts these tools are helpful with.

With most commercialised tools, the content of the memory must be provided by the user, i.e. the translator or the company. The quantity of texts needed for the system to become efficient depends on the context, of course. Sometimes one or more small memories yield better results than an extensive one. However, often the database must be quite extensive to increase the likelihood of finding a corresponding text segment. When acquiring or introducing the tool, one possibility to fill the memory is **interactive translation**. This means that the tool is used when translating and the memory is filled as you go along. With this method, the quality of the memory content is generally very high, but the process can take a long time before the tool becomes useful for new translations. Of course, the process can be sped up with a server-based memory where several translators contribute. Another possibility to speed it up would be to purchase and import a TM from elsewhere, but there is always the risk of poor quality, since the origin of the translations cannot necessarily be known and verified, or they do not necessarily respect company terminology.

If interactive translation takes too much time, another method is **alignment** of previous translations. Alignment is "the process of comparing a source text and its translation, matching the corresponding segments, and binding them together as translation units in a TM" (Bowker, 2002: 109). This process can be broken down into two steps, first the segmentation, i.e. the "cutting" of source and target texts into segments, and then the actual alignment, where the corresponding source and target language segments are stored together

in the memory. A set of one source and one corresponding target language segment is called a translation unit. This procedure can be used not only to create a translation memory but also to add to an existing one, and most systems provide a separate tool for it.

For the **segmentation**, the basic unit is generally the sentence, even though there are other categories such as headings, lists, or table cells. Most tools use punctuation as an indicator for segment boundaries, which is an obvious choice but this can also have its drawbacks. Periods, for example, are not only used to end a sentence, but also in abbreviations such as *e.g.* In some languages, such as German, the period is also used for ordinals, for example in dates: *14. November 2010*. Embedded sentences or the use of colons and semi-colons can also cause difficulties in this process. The segmentation rules can therefore be adapted with most commercialised systems, with the help of stop and skip rules or abbreviation lists. Finally, segmentation can also be difficult when working with some Asian languages, because of structural differences.¹³

Once the texts are segmented, the system generally proposes an alignment which can then be checked and, if necessary, corrected. For the alignment, the tool can rely on many different elements. It can proceed by hierarchy, where it relies on text formatting, first aligning big chunks of the text and gradually moving down in granularity (Macdonald, 2001: 2). The alignment can also be based on segment length, and of course a combination of approaches is also possible. Within the segment, the alignment tool can also rely on punctuation, cognates, names and numbers, acronyms and abbreviations. Here again, the result depends on the suitability of the texts: most importantly they need to have an identical or at least similar structure. Alignment errors are caused in particular by inverted sentences, missing passages in the target text, or conversely, additions. Some incorrectly aligned segments are almost inevitable with automatic alignment, and the user most often has to check and correct the proposed alignment in order to maintain a high standard of quality.

The translation units created through alignment are then stored in the memory. The database, however, is not a simple text file: "The databases have to be highly structured, with indexes to facilitate efficient retrieval of examples" (Somers, 2003: 34). Often, the translation

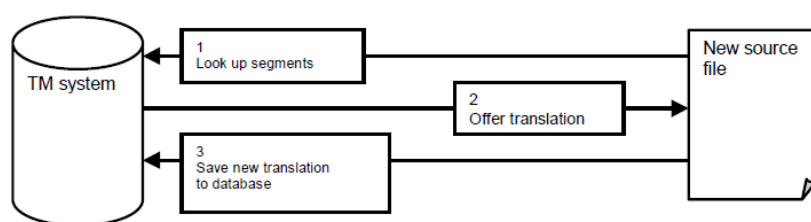


Fig. 2: Database Model (Zerfass, 2002b)

units can be annotated, additional information such as text source, date, code or name of the translator can be helpful for future

¹³ For more information on segmentation see Mikheev 2004.

work. The information can be stored in different ways and this has an impact on how the memory can be used afterwards (Zerfass, 2002a: 11). The most common architecture for the archive is the **database model** (see fig. 2). Most commercialised tools use this model nowadays, including SDL Trados, Wordfast, and others.¹⁴ In this model, the translation unit is stored independently, without its context or any information on its position in the document. This is very efficient for segment retrieval, but once the unit is stored in the memory, it is impossible to reconstruct the original text. When a match is proposed, therefore, the translator cannot look at the sentence before or after the segment, in order to find out more about the context.

With the **reference model**, on the other hand, this is still possible (see fig. 3). This system only references the position of the translation units inside the document, so that the document can be consulted at any moment if context information is necessary. Of course, both models have their advantages and disadvantages. The first one is completely oriented towards the segment, while the second could theoretically be used simply as a reference corpus. There are only few tools working with this model on the market, one example would be MultiTrans.

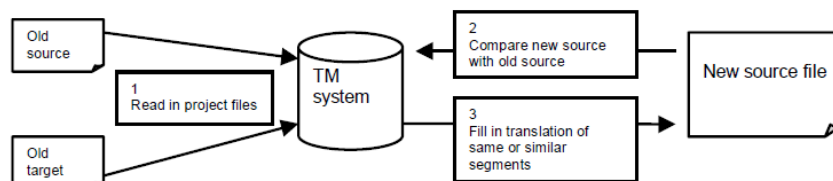


Fig. 3: Reference Model (Zerfass, 2002b)

Once the translation units are stored in this database, the software needs to access them efficiently when the translator works on a new source text. Indeed, if there is no useful information in the database, the translator will have to translate the segment from scratch. The new translation unit will then be stored in the memory and can be retrieved in the future. However, if pertinent information is stored in the memory, there are different approaches to retrieving it. Depending on whether an identical segment or just a similar one exists already, we speak of exact matches or fuzzy matches. An **exact** or **perfect match** is a segment that is "100% identical to the segment that the translator is currently translating, both linguistically and in terms of formatting" (Bowker, 2002: 96). This means that every detail down to punctuation is identical. If exact matches exist for the text, the translator can choose to analyse the text before working on it, in order to find out the percentage of text he will be able to substitute directly. S/he can then run a pre-translation, also called batch translation, where these identical segments are replaced automatically with their translation before the translator starts working on the text. Sometimes, there is no exact match but a **full match**, which means that the segments only differ for so-called "placeables." These are variable

¹⁴ A list with websites of commercialised tools can be found in the bibliography.

elements such as dates, numbers, times, currencies, measurements or proper names. Exact matches normally require little editing, some adjustments might be necessary for the coherence of the entire text or if the context is different.

Fuzzy matching is a more complex issue than exact matches. A **fuzzy match** is a segment which is not exactly identical but similar to a certain degree and might therefore still be of some use. If a fuzzy match is found and suggested by the tool, the differences between the new and the retrieved source language segment are normally highlighted and the translator can then edit and integrate the translation. The degree of similarity is generally expressed in percentage and calculated by a matching algorithm which compares the strings of characters. Most systems have a default threshold for fuzzy matching, but the translator can normally adapt it according to the texts s/he translates or even simply according to personal preference. Setting this threshold is not insignificant: If the threshold is too high, some matches which might still be useful may not be retrieved, a phenomenon called "silence". If, on the other hand, the threshold is set too low, there will be matches with very little resemblance to the new segment and it might take more time to adapt them than to translate the segment from scratch. This phenomenon is called "noise," and is of course as unwanted as the opposite, "silence."

When no match is found by the tool, it is still possible to work with a term match, if the tool is linked to a termbase. This procedure is very similar to automatic dictionary lookup, but often translation memory systems have integrated or linked terminology databases. Some systems now also offer sub-segment matches which differ from fuzzy matches in that not the entire sentence is compared, for which the matching percentage would be very low, but smaller chunks of sentences. In this case, the fuzzy matching threshold can be kept high in order to avoid too much noise. However, this is not generalised for commercialised systems, and the biggest problem with matching algorithms still persists. Indeed, they do not take into account inflection or derivation when comparing the new segment with the translation memory, so that potentially useful matches might not be retrieved because the differences are considered too big, even though it could be just a difference of person (e.g. *he was going...* against *they were going...*).¹⁵ Generally speaking, fuzzy matches can be very useful but the matching threshold needs to be carefully adjusted, and they also require careful proofreading and editing.

Of course, this kind of system does not work equally well with all types of text. Imagine, for example, literary texts and fiction: Even if the translation memory is very extensive, the chances of finding an exact match are tiny, simply because there is not enough repetition in these texts. On the other hand, translation memory systems are useful when

¹⁵ For more examples see Bowker 2002, p. 106-7.

translating texts which contain much internal repetition, for example. Such a highly repetitive content can be found most often in technical, scientific, legal or other specialised language texts. In these fields it is also quite common to reuse parts of older texts, a situation in which a translation memory is obviously helpful. They are also useful with revisions or texts undergoing frequent updates. In the case of user manuals or web pages, the source text is often changed while the translation is still in progress. Especially in the localisation industry this becomes an important factor, because the use of translation memory technology allows for the translation to begin before the source text is even finalised, thus making a step towards the sales-promotional simship.¹⁶ There are other, more formal factors which can influence the suitability of a text for work with a TM system. First of all, the texts have to be available in electronic form but formatting and elements such as foot- and endnotes can also be problematic in certain cases. Additionally, the diversity of file formats can cause some compatibility issues, requiring file conversion before the translation can begin. Today, many systems include a file format converter.

Compatibility can also be an issue if the database has to be distributed, for example to freelance translators who may work with a different translation memory system at home. The database file should be compatible, otherwise conversion is necessary, which always involves a risk of information loss with every import and export. Today, the Translation Memory eXchange format .tmx is widely accepted to exchange these memories between different tools. It was created in 1998 by OSCAR (Open Standards for Container/Content Allowing Reuse)¹⁷ with the goal of allowing "easier exchange of translation memory data between tools and/or translation vendors with little or no loss of critical data during the process." Sharing a translation memory inside a service or a company can also be done with server-based or even internet-based solutions, where all translators have access to the same memories for their translations.¹⁸

Overall, it is generally agreed that the use of a translation memory, especially in a large service or for long documents, improves consistency in terminology. This can be true even for the entire company documentation, if all translators work with the same memories, containing the same terminology. However, as Bowker notes (2002: 116): "a prerequisite for obtaining a high-quality result from a TM is that the translations stored there by humans must be correct in the first place." Most companies or language services have internal revision processes which normally ensure the quality of the translations. The translation memory

16 "An abbreviation of "simultaneous shipment", which refers to the practice of releasing multiple language versions of a product at the same time (or at least as close to the same time as is possible)." (Bowker, 2002: 152).

17 OSCAR is a committee attached to the LISA (Localisation Industry Standards Association). For more information see <http://www.lisa.org/OSCAR-LISA-s-Standards-Committee.79.0.html> (26.02.2011).

18 For more information, see Levitt 2003, or the developers' websites.

systems can support this process in different ways, one of which is to grant different privileges to different collaborators in the service. For example, some translators can only consult the memory but not add any new translation units. Or if they can, the new units can be put on hold and will be verified and corrected by an administrator or senior translator before they are added to the memory for general use. Sometimes, however, it is not the translator's work which decreases the quality of the database. Indeed, translations and terminology can simply become inaccurate over time or because of a change in company or language policy. Constant maintenance of the memory is therefore essential for long term quality assurance.

One often cited advantage of translation memory systems is a gain of time, i.e. an increase of productivity and therefore money. According to Somers (2003: 42) this is often the first thing a potential buyer wants to know. With respect to these expectations, the author mentions that in certain circumstances a 60 percent productivity increase can indeed occur, but generally "30% may be a more reasonable average expectation." Also, this increase will most often be felt only in the long term. When introducing a new tool in the service, the translators first have to adapt to the system and learn how to use it. This learning curve also concerns customisation, such as adapting the matching threshold or segmentation rules to the specific context. Additionally, in some cases a lot of time is needed for the preparation of the TM, for file conversions and text preparation, as well as alignment. Once the memory is up and running, time is also required for maintenance and support, for updates and training. The increase of productivity may therefore not be very noticeable in the initial phase. One last factor to be considered is of course the financial investment for the system, which can be quite important, again, not only for the initial acquisition but also for the follow-up and quality assurance.

2.2.3 Impact on translators and translation

Un décideur averti aura évidemment essayé de tenir compte du fait que des êtres humains vont interagir avec le système (...), aucun système ne peut être efficace si personne n'accepte de l'utiliser.¹⁹

As we have seen above, in the early 1950s and 60s, and especially before the ALPAC report was published, the goal of MT research was essentially a machine which would transform a source text into a perfect target text without any human intervention whatsoever. At this prospect, translators understandably felt threatened, and this fear is still haunts some of them

¹⁹ King, 1993: 265.

today. According to Bowker (2002: 120), they are afraid of being “replaced by a computer or (...) reduced to someone who merely has to click on the “ok” button to accept a ready-made translation.” A translator quoted by Quah (2006: 17) also mentions concerns about the impact of the tools on the translation process: she is afraid that it could “dull” her creativity. These fears are legitimate, of course, but at least partly unfounded. The fear of being replaced, for example, was mainly caused by the high expectations with respect to MT systems, but the ALPAC report is still right today in that there will be no FAHQT in the near future, except for some very specific contexts, and thus translators will not be put out of work just yet.

Also, while changes in a well-set routine are always somewhat unsettling, especially when caused by new technology and introduced in a brief time-span, the negative attitude of translators towards translation memories is mostly based on a lack of knowledge, as Sarah Dillon and Janet Fraser (2006) have found out. They realised that there is some reluctance among professional translators to adopt a TM system and they wanted to know more about the cause of this negative attitude with the help of a questionnaire survey among professional translators. Their results are very revealing: a first hypothesis confirmed that a high number of years of experience have a negative influence on the attitude towards TM systems. On the other hand, “newly qualified translators tend to be more open to the idea of adopting TM” (Dillon & Fraser, 2006: 76). This could also be due to the fact that information on CAT tools increasingly becomes part of the translation curriculum because, as a second hypothesis confirms, the more experience the translators have with TM, the more positive is their attitude. Finally, a third hypothesis showed that “translators with strong IT skills were far more likely to perceive TM as being highly compatible with their working style” (76). It thus seems that not only knowledge of TM systems is important, but also an individual translator's perception of their IT skills. A solution would therefore be to increase training for general IT skills as well as TM systems in particular, during the translation curriculum but also for translators having to adopt a TM system for their work. Clearly, with more understanding some of the translators' concerns could be dispelled.

This still leaves us with the question whether translating with a TM system is different from translating without it. In other words, does this technology really have an impact on the translation process itself? The answer here is not easy to find, particularly because there are few studies carried out on this subject, and the existing ones are rarely very conclusive for statistical reasons. Generally speaking, the study of the translation process is much younger than the study of the product, serious research on what exactly happens during this process only dates back to the 1980s (Schnell & Aranda, 2007: 34). However, if the authors of the few available studies do not necessarily agree on the exact nature of technology's influence, they

do generally agree that recent developments in translation technology and in translation as a business do have an impact on the way translators work today.

The translation process is often broken down into three phases, the orientation, the translation and the revision phase. A general comparison of these three phases when working with or without TM can already show some basic changes (Schnell & Aranda, 2007: 35): During the orientation phase, the translator normally reads the text and decides on a macrostrategy for the target text. This step normally falls away when working with a TM because the text is presented in segments and often already pre-translated, so that it is doubtful whether the translator even has the possibility to read the entire text before starting to translate. In the translation phase, when a previous translation is proposed for the segment, it is possible that certain cognitive processes normally present in translation are not activated at all. On the other hand, the translator would take more time revising the proposed segment and adapting it if necessary. Finally, in the revision phase, the process stays the same, but is normally shorter than in manual translation, either because the revision was already done when translating the segment, or because revising the text in the TM programme can require more effort than doing so on paper.

In order to confirm these general tendencies, Bettina Schnell and Marcos Aranda have carried out a questionnaire survey with users and non-users as well as a scenario test with professional translators, which have allowed them to draw several conclusions.²⁰ The questionnaire first of all confirmed that revision indeed takes place predominantly during the translation phase when working with a TM system. It also confirmed the fact that the impact on creativity is a concern: 81 percent of non-users fear it, while still 46 percent of users attest to it. This is especially interesting since other fears of non-users seem to be unfounded, as mentioned above. Another point raised by the questionnaire and confirmed by the scenario test is an impact on segmentation of the source text, noticed by 63 percent of the users in the survey. In the scenario test, the researchers observed a change in the distribution of pauses. When working with a TM system, the translators paused before every sentence, while there were notably less pauses when working without it. This suggests that translators would use larger chunks of text as "segments" when translating without a TM system. The use of a TM system thus seems to fragment the meaningful units for translators.

As the above study shows, there clearly is some impact on the translation process, even if it is still not easy to locate and quantify. However, this does not mean that the impact is necessarily negative. In any case, no general tendency towards a decreasing quality of translations could be detected since the introduction of translation memory systems. Also,

²⁰ Again, the set of data is very limited, 176 participants for the questionnaire survey and only 2 participants for the scenario test.

these changes need to be seen in the light of more profound changes in the profession and business of translation. As Brian Mossop (2006: 790) notes, there is an increasing demand for speed – because time is money – which leads to a change in techniques, for example to chunking and 'collage' translations. This relatively new procedure is characterised by division of labour, as well as reuse and revision of older material. As already suggested by Schnell and Aranda, this process leads to changes in the mental process. However, Mossop focuses not so much on individual CAT tool, but on translation 'as a business': "the collage method of producing translations, while certainly *enabled* by information technology, is being *driven* by business pressures" (Mossop, 2006: 790). He goes on to conclude that "Technologies are being adopted to serve business purposes, and an offshoot of this, perhaps, is change in the mental process of translation" (792). It could be added here that the technology is not only adopted for business reasons, but also developed and advertised for efficiency and cost-effectiveness.²¹ It seems then that as the world is changing, becoming more and more globalised, translation changes with it. Certain compromises will probably have to be accepted, because "[t]ranslators can no longer afford to be afraid of CAT" (Carter-Sigglow, 2004: 16). In any case it is necessary to monitor these changes and on the impact they have on the work of translators through continuing research.

21 See for example the SDL Trados website (Enterprise customers): "improve process efficiency and increase the quality and volume of localized content – all whilst reducing cost"
<http://www.trados.com/en/enterprise-customers/default.asp> (02.03.2011).

3 Context: Central Language Services

There is no such thing as a best system, but a best system for a particular situation.²²

When looking for answers, one first has to ask the right questions, so to find the best possible solution for the German language service, we first have to know the context into which a potential tool would have to fit. It is therefore crucial to find out more about the Central Language Services before setting up the different tests. Some of the information in this chapter is taken from the Federal Chancellery's website,²³ but most was collected through several directive interviews²⁴ with employees of the different services of the Central Language Services, conducted during the first half of 2010. An overview of the interviews and the different questions and answers can be found in Annex A, p. 87.

The Central Language Services (Zentrale Sprachdienste, ZSD) are part of the Federal Chancellery of Switzerland (FC), where they are attached to the sector Federal Council. They consist of a German language section, a French language section, an Italian language section and a terminology section. An English translation service is attached to the terminology section. The following sections will describe the German language section, for which this study is carried out, the environment, i.e. the other sections of the ZSD, as well as the texts which are translated at the German language section and which will therefore constitute the primary texts for this study. They will be specifically analysed for their content and repetitivity in the last section.

²² Rico, 2001: 2.

²³ Federal Chancellery: <http://www.bk.admin.ch>

²⁴ The interviews followed a basic questionnaire, leaving room for follow-up questions (Fenneteau, 2007: 14-15).

3.1 ZSD-D

As mentioned above, the German language section, the ZSD-D, is one of several language services inside the Federal Chancellery. It is located in Berne and under the direction of Mr. Markus Nussbaumer, who is seconded by Mrs. Isabel Kamber. Their responsibility is specifically to ensure that the texts produced by the FC are precise, clear, free of contradiction and simply formulated at any stage of the process. This is especially important because a great part of their work involves legal texts. According to the website, the section has several responsibilities, translation being only one of them. The first and most important task is "Koredaktion" in French and German, namely the simultaneous production of texts in those two languages, in collaboration with the French language section and the authors from the offices concerned. Another task is "Redaktion", i.e. the production of administrative texts for the FC, including messages (Botschaft), reports (Bericht), statements (Stellungnahme), online text (for ch.ch), etc. Since they are generally experts on language and writing, they also edit manuals and guidelines, and answer language related questions coming from other departments of the Federal Administration.

The translation proper is also part of the section's tasks, but it accounts for only about 20 percent of the total work load. Also, the translation volume varies greatly, it is always high during the Parliamentary Sessions, and much lower between them. Generally speaking, the section translates a wide range of texts into German, mostly texts written inside the Federal Chancellery or by the president of the Federal Council, as well as the Parliamentary Interventions (Parlamentarische Vorstösse). The content of the texts will be further examined in section 3 below.

There are 11 people working in the section at the moment, sharing a total of 8 full-time positions. There is one secretary and one editor of LeGes, a triannual paper on legal texts, the other nine are "Redaktoren" ("writers"). They are all formed translators and writers with several years of experience, their experience with translation tools is limited though. During the Parliamentary Sessions, which take place four times a year for three weeks,²⁵ and for three weeks following the sessions, the section employs a translation intern whose main responsibility is the translation of the Parliamentary Interventions. Depending on the language combination of the intern and on the amount and length of the interventions, the intern will be able to translate all of them, but if this is not the case, others from the section will have to translate some too. When there is no intern, the translations are distributed to all employees according to their current work load and availabilities. With respect to revision, all

²⁵ "Sondersessionen", i.e. special sessions, are possible, but no translation intern is present at those times.

translations are corrected and discussed with at least one colleague in the section. This elaborate system is also used for the translations produced by the intern.

3.2 Environment

As mentioned above, there are several other language sections in Central Language Services. However, their responsibilities and tasks are not always identical. The French language section, which is slightly bigger than the German one, with 12 people for 10.5 full-time positions, resembles it closest. Here too, translation is not the main task, representing about 30 percent of the total work load, while the main tasks consist of "Koredaktion" and revision. Concerning the texts, they translate more or less the same texts as the German language section, i.e. texts produced in the FC as well as Parliamentary Interventions.²⁶ Interestingly, they have examined the question of introducing a translation tool a few years ago but decided that it would not be worth the investment because there is not enough repetition in the texts they translate. Also, there seems to be considerable resistance to the basic approach of Translation Memory systems.²⁷

The Italian language section is the biggest section of the Central Language Services. It is located on two sites, not only in Berne but also in Bellinzona, employing 32 people for approximately 27 full-time positions. Here, translation is the main task of the section, making up 60 to 70 percent of the total work load. This percentage is completed by revision, terminology, documentation, meetings, etc. On average, the section translates some 19'000 to 21'000 norm pages per year. The texts are mostly legal texts, federal acts and ordinances, but also messages, reports, press releases, etc. They also translate Parliamentary Interventions, but only when they concern the FC, otherwise they are translated in the offices concerned.

The English language section, which is attached to the terminology section, is much smaller than the other sections, with only 5 people for a total of 3.2 full-time positions. Here, too, translation is the main task, with some terminology, and the service translates some 2300 norm pages per year. This is not counting some larger projects of translating basic legal texts such as the Civil Code, which run in parallel. The texts they translate are very diverse, as they translate mostly on demand by the FC and the different offices of the Federal Administration. However, they do not translate any Parliamentary Interventions.

²⁶ Unfortunately, more detailed information on text content and volume could not be obtained from this section.

²⁷ See the interview with F. Bertagnollo (Annex A, p.89).

The English and Italian language sections are already using a tool, namely Multitrans, for their translations, and they seem quite happy about this. The initiative came in 2005 from the English language section, which started the initial testing, while the Italian language section followed soon after in 2006. They went through extensive testing of several tools, SDL Trados, Multitrans, Star and MetaRead, and finally chose Multitrans for several reasons. They were looking for a tool which would work even if the texts are not very repetitive and with which they would not have to prepare and align texts for weeks before being able to use the tool. Price was of course also a question – SDL Trados was considered too expensive at the time. The tests with Trados were rather brief, the tool was soon disqualified because there was not enough repetition in the texts. Also, they feared more general resistance against a purely sentence-based tool, while Multitrans is more flexible in its use.

After starting off with a Multitrans server solution and 5 or 6 licences, the two services together count 27 licences today. This is sufficient for them, the floating licences accomodating part-time employment. Some components of the tool are used more often than others, the Translation Agent, the TM interface with automatic segment retrieval, is used very rarely. More often, the TextBase module is simply used as a reference tool in the form of a bilingual corpus. While in the English language section only one database is kept, in the Italian language section there are two: one for legal texts, i.e. acts and ordonances, and one for all other texts. In general, they appreciate the flexibility of Multitrans, as well as the fact that alignment is very easy. For the database of legal texts, for example, realignment is done every three months, but it can be scheduled and carried out during the night, when no one is working. Uploading the 14'000 texts of Swiss legislation takes about 30 minutes. In the English language section, new texts are added every two weeks, again overnight through scheduling.

There was some resistance from translators initially, at least in the Italian language section, where translators were afraid of job cuts because of the tools. Today, as the users see it, the introduction of the tool has led to more consistency, especially in terminology, and also to a slight increase of speed, even if this is difficult to measure. The information obtained from the two sections using Multitrans could be taken as an indicator for our work for the ZSD-D: no need to test Trados, then, because there is not enough repetition in most of the texts. However, this would be oversimplifying the situation. The texts translated in the ZSD-D are very different from those translated in the Italian and English language sections, and so is the overall importance of translation for the section. Stephen Frost, head of the English language section, also clearly sees a difference in the responsibilities and translation volume between his and the Italian language section on the one hand, and the German and French language sections on the other. I therefore believe that, even though the information and experience

from the Italian and English language sections are very interesting and helpful, I cannot base my study on the results from their testing because there are significant differences in the context. These differences concern the texts translated as well as the share of translation in the overall work load. The next step is therefore to examine the texts translated in the German language section more closely, in order to select the right primary material for the tests further on.

3.3 Texts

Finding out more about translation volume and text content is important because the information will allow me to select the texts which matter the most for those who would eventually be using a tool in the section. The statistics in figure 4 show the number of pages translated in 2009.²⁸ They show that during that year, the service has translated 801 pages. This is not much, even compared to the smallest, i.e. the English language section (2300 pages). The Italian language section, as we have seen above, translates roughly 24 to 25 times as much in one year. Generally speaking, then, the text volume is rather small for a language service, which is normal considering the fact that translation is not their main task. Most of the pages are translated from French (703 pages, i.e. 87.8%). The second language is Italian (83 pages, i.e. 10.4%) while only few pages are translated from English (11 pages, i.e. 1.4%). Also, the statistics show that the Parliamentary Interventions constitute the biggest group of texts in the service, namely 58% of all translations done within the service. This picture can yet be refined with the statistics of the other translations (besides Parliamentary Interventions) from 2008 (see fig. 5).

Statistics 2009	
Interventions	463.92
Others	337.1
Total (pages)	801.02
FR total	703.36
IT total	83.47
EN total	11.2

Fig. 4: Statistics 2009

In terms of quantity, the second group corresponds to texts translated for the Federal Chancellery (FC), and amounts to 235.8 pages, which roughly corresponds to half the number of pages of Parliamentary Interventions translated during the subsequent year. Other groups are much smaller. The section translated about 53 pages for the President of the Confederation, another 19 for Parliamentary

Total Pages 2008	337.1
Total FC (incl. FDPIC)	235.8
Total FC President	52.7
Total FDF	5.3
Total Parliamentary Services	19
Total other Administration	24.2
Total FR	288.9
Total IT	47.8
Total EN	0.4

Fig. 5: Statistics other translations (besides Parliamentary Interventions) 2008

²⁸ Statistics communicated by e-mail, I. Kamber (21.06.2010). See also Annex B, p. 98.

Services. Some 5 pages were translated for the Federal Department of Finance (FDF), with which they have an agreement concerning translation, and some 25 pages for other departments of the Federal Administration. It seems thus that there is no other group of texts which reaches the importance of Parliamentary Interventions in terms of translated pages per year. This also means that even though there may be much repetition in some of the other texts, if the tool is not useful when working with the Parliamentary Interventions, it would hardly be interesting for the German language section, simply because the text volume processed by the section is not great enough.

In view of this distribution of the texts, I have chosen to work only with Parliamentary Interventions here. As they constitute the largest group of translations done in the ZSD-D, most of the work with a potential TM system would be done with those texts too. However, there are other reasons as well. One of them is accessibility: Parliamentary Interventions are freely accessible in German, French and Italian, from Curia Vista, a database on the website of the Federal Assemblies.²⁹ It was therefore not necessary to ask different services for texts, which would have delayed or limited the project considerably. Also, in those services where a TM system already exists (see above), there are no TMs with Parliamentary Interventions. According to those using it, Multitrans is very helpful when it comes to translating legal texts or reports, but they have not used it for Parliamentary Interventions, either because they do not translate that many interventions, or because they judged the repetition in those texts to be below a useful level. It is therefore a new challenge to build a TM with those texts, and to see whether a TM system can actually be useful or not.

Parliamentary Interventions, as their name already says, are written by Swiss Members of Parliament in the Federal Assemblies. They can be submitted to the Parliamentary Services only during the quarterly sessions as well as the special sessions. Some of the interventions are the result of debates in commissions, those can also be submitted outside the regular sessions. Contentwise they consist of motions, initiatives and others requests to the executive power, the Federal Council and the respective offices, in order to obtain answers to urgent questions, to suggest subjects to be examined or changes in the legislation. There are different types of Parliamentary Interventions: Parliamentary and Cantonal Initiatives, Motions, Postulates, Interpellations, Questions and Questions for the Question hour which takes place the second and third Monday afternoon of every session. They all receive a number in the order of entry, composed of the year of entry (two digits), followed by a dot and three or four digits (e.g. 10.5621 is one of the last ones submitted in 2010). The type of intervention can be identified from this number. Parliamentary and

²⁹ Curia Vista, the database of parliamentary proceedings:

<http://www.parlament.ch/E/DOKUMENTATION/CURIA-VISTA/Pages/default.aspx> (12.05.2011).

Cantonal Initiatives, for example, count only three digits after the dot. The four-digit number of Questions for the Question hour start from 5000 (e.g. 10.5134), while the one of ordinary Questions starts from 1000 (e.g. 10.1162). All other interventions, i.e. Motions, Postulates and Interpellations receive a number in the order of entry starting from 3000 (e.g. 10.3698).

These texts can treat any subject which is of importance to Swiss members of Parliament and the Swiss people. They can concern policy in a wide range of domains such as agriculture, public health, nature conservation, migration, or any other, even though the subject would have to concern a federal level (as opposed to cantonal or communal level). Also, the texts can come in the form of questions, lists of questions, suggestions or accusations etc. They can be similar to legal texts, when a change of legislation is suggested, but they can also be simple prose argumentation. There is nothing, therefore, which suggests that the repetition in the texts is very high, except some standardised formulas. Also, they do not correspond to any sort of text normally suggested to be suitable for use with a TM. As mentioned above, in the Italian language section, no TM has been set up for Interventions because the repetition was thought to be below a useful level, and even in the German language section, the impression is that there might not be enough formal repetition in the texts. However, in terms of topic, there is at least some repetition, as there are several recurring themes which occupy Members of Parliament over a long period, or which for some reason are taken up repeatedly, for example because the solution is not considered satisfactory by some Members of Parliament. The next section will now analyse these texts in order to find out more about the extent of this formal repetition.

3.4 Content analysis of Parliamentary Interventions

In order to find out more about the repetitiveness of Parliamentary Interventions, I have carried out a small study with the tool Repetitiveness Checker, a tool developed by the Center for Sprogteknologi at the University of Copenhagen.³⁰ It uses a statistical method to find sequences of words that occur more than once in one or more texts. It can be used for many purposes, one of them being to assess "whether automatic translation by means of translation memory is a good option."³¹ This tool should give me some preliminary indication concerning the repetition in the texts and how extensive a potential Translation Memory would have to be.

30 CST's repetitiveness checker: http://cst.dk/online/rep_check/uk/ (04.02.2011).

31 Background information on CST's repetitiveness checker: http://cst.dk/online/rep_check/uk/ (04.02.2011).

Given the variety of contents the interventions can present, I have decided to select the texts with the help of a key word, limiting their scope to one of the recurring themes mentioned above. I have chosen to work with the keyword *LAMal* (Loi fédérale du 18 mars 1994 sur l'assurance-maladie), a theme which occupies Swiss Members of parliament very regularly even since before its coming into effect in January 1996. A full text search in French yielded 798 texts,³² enough to do extensive testing with Repetitiveness Checker. I have proceeded with 25 texts at a time, running the programme with the same settings, in order to find out if there are changes in repetitiveness, an increase, decrease or stagnation.

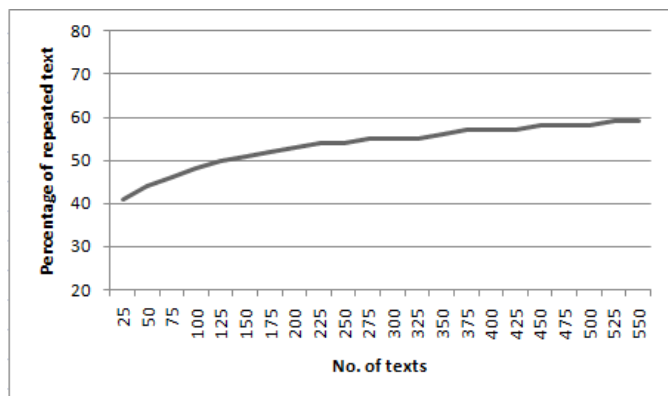


Fig. 6: Repeated Text, threshold 0%

A first series of tests, going up to 550 texts, revealed that there is, indeed, some repetition in the texts. The percentage of repeated text started at 41 percent for 25 texts, slowly edging up towards a final 59 percent for 550 texts (see fig. 6). This does seem like a good quantity of repetition, even though the slow

progression is not very promising. It suggests that a TM would have to be quite extensive in order to be useful for the translator. In this series of tests there was no limit imposed on the length of repeated word sequences, although the default minimal length is of two words. In terms of fuzzy matching this would correspond to a very low threshold and it is doubtful therefore whether a repetition of such low percentage is actually useful to the translator.

In order to dispose of more accurate information, I have therefore chosen to run a second series of tests, using the same number of texts, but this time imposing a fuzzy matching threshold of 75 percent. This threshold is not insignificant, of course. Indeed, studies have shown that fuzzy matches below this percentage are often not

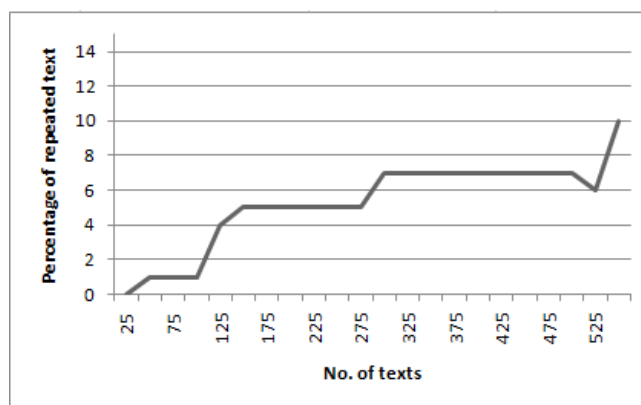


Fig. 7: Repeated Text, threshold 75%

useful any more.³³ The default threshold in SDL Trados Studio, for example, is at 70 percent, while in Wordfast it is at 75 percent. Imposing this threshold is therefore legitimate when assessing texts destined for the establishment of a TM. This second series of test has revealed

³² Reference date 1st September 2010.

³³ See e.g. Whyman and Somers 1999, p.1278; Bowker 2002, p.99-100.

a totally different picture (see fig. 5). Repetition was basically absent in the first 25 texts, and the 10 percent mark was only just reached with 550 texts. This confirms the impression mentioned by several people working in the FC, such as Mrs. Petrone or the head of the French language Section, Mrs. Bertagnollo. It seems that formal repetition is indeed not very high in Parliamentary Interventions. A potential TM would therefore have to be very extensive or offer other solutions in parallel in order to have an impact and for the tool to be useful for the German language section.

Concluding, we now know that the overall text volume is rather small and that the texts translated are very diverse. We also know that the repetition in the most important group of texts is not very high and that it would probably take a very extensive Translation Memory, and massive alignment of previous translations for the tool to be efficient and useful. This does not mean, however, that the usefulness of a translation memory can be ruled out categorically. As we have seen in the previous chapter, there are different ways a translation memory can be used, and if the alignment of the texts does not encounter too many problems, the time and effort required for the establishment of an extensive memory will not be imperative. However, for the following steps of this study, the information collected here and the results of these tests have to be taken into account.

4 The tests

To *evaluate* is to determine what something is worth to somebody.³⁴

On the basis of the information we have gathered throughout the previous chapters, we are now ready to look for answers to the questions underlying this study. Again, the most important question is whether a TM system would be useful to the ZSD-D and if an acquisition would be cost-effective, as well as what measures would have to be taken to introduce it. More generally speaking, this chapter also attempts to find out whether it would be useful for any translation service to work with a TM system, even if the repetition in the translated texts is limited. In order to find answers to these questions, a series of tests has been designed. First, section one will provide some theoretical information on the evaluation of translation tools and the methodology applied here. Section two then features the tests, for each of which the setup is presented, followed by the results and the impact of these results. Finally, section three will draw the conclusions from these tests and describe some other elements which have to be considered before a decision is taken as to whether or not a tool should be introduced in the ZSD-D.

4.1 *Evaluation and methodology*

Before describing the test arrangements themselves, this section provides the reader with an overview of what has been done until now in the field of translation tool evaluation, the methodologies that have been developed and the particular methodology that will be used in this study.

³⁴ EAGLES, 1996: 15.

4.1.1 Translation tool evaluation

When browsing through the history of translation tools and therefore also the evaluations that have been carried out, one finds many reports on machine translation systems, but not so many on CAT-tools. Indeed, as we have seen in chapter 2.2.1 (p. 12), the history of machine translation is very long in comparison with the history of Translation Memories and the same goes for product evaluation and reviews. In the field of machine translation, the first evaluations were probably carried out as soon as the first ideas were put forward. The importance of evaluation in the research and development process of software products has already been stressed by the ALPAC report in 1966, probably one of the most generally known evaluations of MT done so far (Quah, 2006: 129-130). For our purpose, most of these early evaluations are not very interesting because besides their obvious focus on MT, they were also conditioned by that perspective in that they generally evaluated the translation product of the tool, namely the target text, rather than the process and the tool itself (Quah, 2006: 136).

With the change of focus in the 1970s and 80s, evaluations were also carried out with other tools. However, these reports are still rather hard to find. The scarcity of texts on this topic is probably due to the fact, as King (1993: 261) and Quah (2006: 131) note, that these systems soon moved beyond the research status and that developing companies are generally not very willing to publish any evaluation reports. Indeed, with the commercialisation of a large number of tools since the 1990s, the developers are facing competition and therefore do not want to make their findings accessible to a larger public. What can be found, then, are mostly reviews which are carried out by potential customers or independent experts of the domain and published in translation journals, consumer magazines, newsletters or simply on the internet.³⁵ These reviews are interesting for users, especially with the multitude of tools available on the market today, but they do not generally present their research methodology and practices in great detail.

The idea of creating a standard methodology for such evaluations appeared at around the same time as the first Translation Memories hit the market in the early 90s. At that time, for every new evaluation "all the literature had to be searched to find suitable criteria" (Quah, 2006: 140), an enormous effort which evaluators naturally wished to avoid. The thought of a 'universal framework' that could be applied and adapted to specific circumstances was therefore obvious, but establishing it was rather difficult, not only because of the variety of

35 For examples of such reviews, see e.g. the journal *Multilingual Computing & Technology* <http://www.multilingual.com> or the German *MDÜ – Fachzeitschrift für Dolmetscher und Übersetzer* <http://www.bdue.de>, or Angelika Zerfass' site on <http://www.zaac.de> (31.03.2011).

tools available but also because of the variety of groups with an interest in testing (Quah, 2006: 129). Several researchers attempted to establish such a methodology at the beginning of the 1990s, one of which I will briefly present here. However, soon some more extensive projects were initiated to work on the subject, such as the ISO-Standards or the EAGLES working group, which had the most widespread impact later on.

In her 1993 article "Sur l'évaluation des systèmes de traduction assistée par ordinateur", Margaret King insists that CAT-tool evaluation cannot be taken lightly. On the contrary, she affirms that it is a complex interaction between needs, constraints, questions of product availability and the disruption of workflow and working environment, and even psychological considerations (King, 1993: 262). Because of this complexity of the issue, she wished to see more collaboration and exchange with respect to evaluation methodology and begins herself by identifying the actors involved in and concerned by CAT-tool evaluation. I will quickly present these here because the question of stakeholders is not to be underestimated in any kind of evaluation. According to the different phases in the development of a software product, the actors involved vary and with them their respective interests and priorities. In the research phase, these include the researchers, but also the research sponsors. In the development phase, there are again the developers and those who finance it. Finally, once a tool is ready and commercialised, there are those who decide whether or not to acquire a tool, as well as the actual end-users of the tool, namely the translators. This typology has not changed very much over the last two decades, as Quah (2006: 133) identifies more or less the same: researchers, developers, sponsors and end-users. Depending on at what moment of the development process a software product is evaluated, the different actors involved and their interests in the evaluation have to be considered.

In a second part of her article, however, King goes on to suggest a first possible 'skeleton' structure which could be used in a CAT-tool evaluation. This procedure includes four steps: (1) the requirements have to be identified and integrated into a list of criteria prioritised according to the importance of each criterion for the specific context; (2) an appropriate technique has to be chosen; (3) the technique is applied in order to collect and analyse the data; and (4) a judgement is formed on the basis of the data analysis (King, 1993: 266). As we shall see further on, this procedure already contains most of the key elements of a good evaluation procedure, even though there is yet room for more elaboration. Interestingly, though, King notes that steps 1 and 4 are dependent on the specific circumstances of the evaluation, while for steps 2 and 3 she suggests compiling a sort of 'catalogue' of research techniques from which the appropriate one can be chosen for each case. The basic ideas for a 'universal framework' are therefore already present in King's article.

Two years before Margaret King made these very practical suggestions on how to proceed for a CAT-tool evaluation, an ISO-Standard relevant to this issue was published which was to have a great influence on translation tool evaluation. ISO, the International Organisation for Standardization, is a developer of standards, "with the aim of ensuring that the development, manufacture and supply of products and services are efficient, safe and environment-friendly" (Quah, 2006: 140).³⁶ The ISO-Standard 9126 was published in 1991, and together with the ISO 14598, which was published much later in 1999, this standard is commonly thought to be the foundation for the general frameworks established later on. The ISO 9126 concerns Software Product Quality and therefore regulates the quality of natural language processing tools, though not their evaluation in particular (Quah, 2006: 140). It defines 6 key quality characteristics which describe a good software product: functionality, reliability, usability, efficiency, maintainability and portability. I will come back to these characteristics and describe them in more detail in section 4.2.1 (p. 36) below because they are relevant to this study. This basic catalogue of key characteristics still serves today as a basis for evaluation in that it provides a list of requirements which can be tested.

A few years later, in 1999, the ISO-Standard 14598 was published, which completes 9126 by covering Software Product Evaluation. Specifically, it lays down an evaluation process, which strongly resembles but at the same time refines the one suggested by King. In this five-stage process, stage one, Evaluation requirements, consists of identifying the criteria for evaluation. Stage two, Evaluation specifications, assigns measurement and metrics to each of the criteria, while in stage three, Evaluation design, the plans are drawn out and scheduled, and the testing methods are selected. Stage four, Evaluation execution, represents the collection of data by carrying out the elaborated plans, and in stage five, Evaluation conclusion, the findings are written up and presented (Quah, 2006: 141-142). Together with ISO 9126, this presents us with a rather complete methodology for software evaluation, with a list of criteria from which to select the ones pertinent for the specific context, and an easy to follow procedure to carry out the tests. However, there is still room for elaboration, for example through a list of methods, just as proposed by King. Indeed, the methodology described in the two ISO-Standards was further refined in several different projects, such as the Expert Advisory Group on Language Engineering Standards (EAGLES), International Standards for Language Engineering (ISLE), and others (Quah, 2006: 142).

Of the many projects initiated to work on the standard framework for translation tool evaluation, the results of the EAGLES Working Group are the only ones I will present here, because it is the most influential project, but also because I will use it for my own test arrangements. This research group was funded by the European Commission and launched in

³⁶ See the ISO website for more information: <http://www.iso.org> (19.04.2011).

1992, one year after the first ISO-Standard was published. It included several campaigns and published several reports, the final one in 1996 (EAGLES, 1996). This initiative built on the ISO-Standard 9126 for the creation of a standard but flexible evaluation framework, specifically in the evaluation of natural language processing tools. In their reports, we find again the idea of a catalogue of features and attributes, but also of techniques that can be combined to reflect the specific circumstances of a context. They elaborated three main components of evaluation: a set of attributes, specific requirements and their measurement as well as methods (EAGLE, 1996). In the first two, we find the quality characteristics from the ISO-Standard 9126 again, which are at the same time attributes of the system but also requirements according to the users and their context. These attributes should be measurable, adequate for the user requirements, and at the same time general enough to be usable in comparisons between different systems. They then have to be translated into appropriate values and metrics, in order to be tested scientifically, as King (1996: 74) notes:

"Defining relevant attributes is of no use if there is no way to measure the system's value with respect to those attributes. Measures must be both valid and reliable. They must measure what they are really supposed to measure and they must measure it consistently."

The last aspect concerning evaluation treated by the EAGLES report is the testing method, for which the EAGLES report defined test types, instruments and test materials, just as King imagined in 1993, a 'catalogue' from which the appropriate method can be chosen according to the evaluation context.

For our purpose here, the most important elements are the quality characteristics, in order to determine what to test, as well as the methods, and in particular the test types, to determine how to test them. They will therefore be described in more detail in the next section, which presents aspects of methodology relevant to this study.

4.1.2 Methodology

Clearly, then, for a good evaluation it is not quite enough to just ask the right questions. I will here describe in more detail several concepts which are of importance in order to understand the test arrangements below. These include the quality characteristics, as well as the test types and material used. First of all, however, it has to be understood that this is not a case of glass box testing³⁷, which implies that the evaluators have intimate knowledge of the programming code and algorithms and take this knowledge into account for the setup of the tests. This kind of testing is often done by researchers and developers, but for this study the

³⁷ Also called white box testing, clear box testing, code testing or structural testing.

method applied is black box testing.³⁸ This kind of testing looks at external elements, e.g. input and output or system requirements and focuses on overall performance without any knowledge of the system's internals. It is particularly suitable for user-oriented testing (Höge, 2002: 128), which is also pertinent to this study. I will be testing products which are already on the market, with respect to their suitability for a particular context. The important actors involved are therefore not the researchers or developers but the end-users as well as those who decide on whether a product is introduced or not.

I have listed the six quality characteristics before, but it makes sense to explain them in more detail here, as they will be used in one part of the subsequent test arrangements (EAGLES, 1996: 59):

- **Functionality** concerns the basic question whether the required functions are really present and is often tested as compliance or suitability.
- **Reliability** is the property of maintaining the level of performance in a specific situation. This is often tested through the maturity or fault tolerance of the system.
- **Usability** describes the effort needed to understand and use a system.
- **Efficiency** concerns the way a system uses resources, here processing time is most often tested.
- **Maintainability** describes the effort needed to make specific modifications.
- **Portability** is the compatibility of one system with another.

Obviously, other quality characteristics can be added to this list. For example, EAGLES later added **Customizability** as in "the ability to modify a product in order to satisfy a particular customer's needs" to the catalogue. (Rico, 2001: 5). Since these characteristics constitute a good natural language processing tool, they are at the same time quality requirements and as such the elements which need to be tested and compared. How these characteristics are used exactly in an evaluation will be described in more detail below.

When it comes to choosing a testing method, the first decision has to concern the test type. The EAGLES report suggests three test types: feature inspection, benchmark or systematic testing and the scenario test. The first of these three will only be used to a limited extent in our test arrangements. It describes in detail the technical features of a system, which is especially interesting for system comparisons (Quah, 2006: 144). In general, a checklist of features is used, and measured by boolean variables, e.g. the function is present or not present. This testing can be based on the manufacturer's data or done by the evaluator (EAGLES, 1996: 149).

³⁸ Also called behavioural testing, acceptance testing or functional testing.

The second type, systematic or benchmark testing will be used in one part of our test arrangements. It involves quantitative tests of the important functions with the help of specifically prepared test material. According to Monika Höge (2002: 142), it represents "measurement of system performance without being dependent on personal variables." This kind of test is normally carried out by evaluators and not by the end-user, even though the specific material, e.g. texts from the user context can be used. The 'human factor' in the evaluation is therefore minimised. This test type also allows for comparison, not only of the functions and features of different systems, as in feature inspection, but also of system performance. For systematic testing, the EAGLES-Group has put into place a procedure in seven steps in order to put into place an adequate test arrangement (EAGLES, 1999). This 'recipe' will be used as a basis for a part of the tests below. It is a step by step guide helping the researcher to ask the right questions at the right moment. The recipe begins with the most basic question: **Why is the evaluation being done?** This first step tries to find out the purpose of the evaluation and, consequently, what kind of system or what part of a system is being evaluated.

Once this basic information has been established, the second step includes **the elaboration of a task model**. Here, the relevant roles and agents are identified, in order to find out who will use the tool and what for. With the help of this information, the third step then consists in **defining the top level quality characteristics** from the list of key characteristics detailed above, to find out what features of the system should be evaluated. In step four, then, **detailed requirements for the system are produced**. This often means breaking down criteria into smaller and yet smaller units, until they become measurable and comparable. The fifth step includes **devising the metrics which are to be applied for the requirements** produced in step four. Indeed, the metrics applied should produce results as objective as possible. Also, scores should be applied for satisfactory or unsatisfactory performance, and most importantly, the different requirements should be weighed in order to count them according to their importance for the context. This is, of course, a very important step in the establishment of the tests. Steps six and seven then include the **preparation and design of the evaluation** as well as its **execution**. The test materials, for example texts, have to be prepared, as well as who will execute the evaluation, where and how. The execution not only includes the measurements, but also their rating and summarizing the results in an evaluation report. With this 'recipe', then, following the instructions, an adequate test arrangement can be put into place and executed very easily.

The third type of test proposed by EAGLES is the scenario test, which is different from the other tests in that it involves the end-user. Here, a system is put into its intended use by its intended type of user performing a standardised task, and the quality and efficiency of the

system are measured, but not only: the impressions of the users are also collected (EAGLES, 1996: 149). It is thus done by real translators with real data and generally aims at being as realistic as possible. The data collected can thus be qualitative as well as quantitative. This is particularly convenient to test a system's suitability for everyday routines. However, the fact that here the translators are involved also calls for particular caution. As King (1996: 76) notes: "Any human involvement raises the issue of the extent to which the human is being evaluated as much as the system's performance." This factor needs to be taken into account in any test where humans are involved, but particularly in the scenario test, where this involvement is much greater than in the other test types. For example, in scenario tests the time participants need for one task is often measured, for example to compare the time needed with the software tool to the time needed without it. In translation, however, the tasks are "likely to involve rather complex, individually varying, problem solving strategies" (Höge, 2002: 133). This makes comparison quite delicate and other factors such as computer literacy, motivation, day time, education or experience also need to be considered. Consequently, it is very important to have a representative number of users participate in the test (Höge, 2002: 134).

Besides the requirements and the test types, the testing material also needs careful preparation. For the tests, the evaluator can use different collections of texts. A test suite, for example, is a set of inputs which is carefully constructed to test very specific problems in a system. This type of material is most suited for researchers and developers (Quah, 2006: 136-137). I will here use a test corpus, i.e. "a collection of texts which attempts to represent naturally occurring linguistic data", which can be compiled to reflect the users' needs (Quah, 2006: 136). It is generally cheaper than a test suite because its construction is less complex.

4.2 The tests

Keeping these methodological aspects in mind, we will now actually set up the tests which are meant to find answers to our research questions. This section therefore treats the setup and execution of the tests for this study, as well as the collection and interpretation of the results. The recommendations drawn from these test results will be presented further on in section 4.3 (p. 73). First, the next section will present some aspects of the overall arrangement.

4.2.1 Overall arrangement

The tests will be carried out in three parts, each of which answers different questions but contributes to the overall result. The first part represents a benchmark test. It will look at how the texts of our specific context are treated by the tools, and more specifically by the alignment tools. The features of these tools will be systematically tested using the Parliamentary Interventions, in order to find out whether the programmes encounter any specific difficulties. This part therefore does not directly seek a definite answer as to which tool is better suited for our context, even if the results eventually contribute to this inquiry. Rather, it attempts to answer the question of how the tools need to be configured in order to work our specific texts most efficiently. The second part will try to answer the question of size for the TM, i.e. whether a small or an extensive one is more efficient. In chapter 3.4 (p. 29), we have considered the hypothesis that an extensive one will be more useful, but this question could not be resolved on with certainty from the tests with the Repetitiveness Checker tool. This second test will therefore compare two corpora for each tool, a small and an extensive one, and test the impact of each.

The results of these first two tests will be taken into account for the setup of the third part, which consists of a scenario test. In this part, a translation scenario is established which attempts to find out how useful and how user-friendly the tools are in a context which closely resembles the real one, i.e. the work in the ZSD-D. From the results of this part, recommendations as to which tool – if not manual translation without a TM tool – is most suitable for our context. The first two parts of the test arrangement thus specifically aim at creating the best possible conditions for the scenario test, with respect to the settings of the tools, but also the size of the TM database. For the overall result, only the scenario test will be relevant, but the conclusions drawn from the other parts will not be forgotten when it comes to formulating recommendations for future developments. Before describing the specific setup and the results of each part, we have yet to determine the main ingredients of the tests, i.e. the tools and the specific set of texts.

The Tools

The two systems examined here are SDL Trados and MultiTrans. Trados was one of the first commercialised translation memory systems, presented in 1992 at the trade fair CeBit in Hannover, Germany.³⁹ Today, the original German Trados company has joined forces with the British SDL, so the version used for these tests was SDL Trados Studio 2009 and the

³⁹ See SDL Trados: <http://www.trados.com> (27.02.2011).

integrated alignment tool WinAlign. This latest version of the tool represents an integrated translation workstation, including the translation memory, alignment, terminology and project management in one interface (Geldbach, 2009: 51). After having used MS Words as an editor for years, in the 2009 version the text editor with spell- and grammar-checker is integrated. The company Multicorpora is younger than Trados, having launched its flagship product MultiTrans in 2000.⁴⁰ It has two basic components, a translation memory "TextBase" and a terminology management module "TermBase." For the actual translation work, MS Office applications are used as editors (Wöllbrink, 2006: 31). MultiTrans is different from most other commercialised tools in that it is essentially corpus-based, meaning that the segments are stored within their context, and that the TextBase is simultaneously used as a translation memory and as a reference tool in the sense of a concordancer. The version used here is MultiTrans 4.4.

I have decided to restrict the number of programmes in order to have the time and resources to conduct more profound tests. The alternative would have been to test and compare a large number of different programmes, but to set up less elaborate tests. I believe this would not have been more profitable with respect to the research questions. On the contrary, the danger was to have information on many programmes, but not enough to decide if one of them is specifically adapted to my context. This choice will have some impact on the scenario test where, as we will see, the usefulness of the tool will be mixed up to some extent with user-friendliness, which is very personal and tool-specific. However, I still believe this choice to be more pertinent, as it also allows me to abide by the restrictions of an MA thesis with respect to size.

I have also chosen to work with those two programmes because they are based upon different architectures for their databases (see chapter 2.2.2, p. 14). Trados uses a database model while MultiTrans works with a reference model. Also, they offer different possibilities as to how the databases are used: Trados mainly relies on sentence-based retrieval or context-searches, while MultiTrans can be used for sentence-based retrieval, but also simply as a bi-text, i.e. a bilingual aligned corpus. Also, both offer terminology management. Among the tools using a database model, Trados was also chosen because it is one of the most widely known and used programmes (Lagoudaki, 2006: 23). MultiTrans, on the other hand, was chosen in particular because it is already used within the Federal Chancellery, and an acquisition and integration into the service would therefore be less time-consuming, as they could benefit from the installation and experience of the Italian and English language sections.

40 See Multicorpora: <http://www.multicorpora.com> (27.02.2011).

The Texts

The other vital ingredient for the tests are the texts themselves. The general content and repetitivity has already been examined in chapter 3.4 above. Here, I will therefore only mention how the set of texts used for the tests has been chosen and in what format they have been used. Out of the enormous quantity of texts available in Curia Vista, a certain number of texts had to be selected. I have taken the texts from this database because as of today, there is no other database where the different texts and translations produced for Parliamentary Interventions are stored. The originals are stored at the parliamentary services, the translations in the respective language services for French and German, and in the translation services of the concerned offices for Italian. It was therefore easier to take them from the internet rather than to contact all the different offices and services to receive the originals.

Personal experience has shown that the German translations, and therefore probably also all other translations, since they have to be sent to the parliamentary services in this way, are stored in MS Word files, without any special formatting. The texts were therefore drawn from the website by simply copy-pasting them into a Word document, saved under the number of the document plus a marker for the language (e.g. 03.4562_fr). It may of course be argued that this procedure does not correspond to what someone in the service would have to do in reality. If indeed a tool was introduced, an appropriate procedure for the collection of texts would have to be established, but as of now, there was no other possibility, except to give up any study with Parliamentary Interventions. I will come back to this point in section 4.3 (p. 73) below, but it has to be kept in mind that if a tool would eventually be introduced in the service, the collection of texts for alignment is a point which will need clarifying.

From the mass of Parliamentary Interventions available on Curia Vista, a subset had to be chosen for our tests. In order to reduce the scope of my text search I decided to use a keyword corresponding to a domain in which many texts have been produced, in the same way I proceeded with Repetitiveness Checker.⁴¹ I have decided to work again with the keyword *LAMal* (Loi fédérale du 18 mars 1994 sur l'assurance-maladie) for a search in French texts. This keyword would correspond to *KVG* in German (Krankenversicherungsgesetz), but I have decided to search in one language only, in order to avoid accidental doubles in my list of texts. The test corpus used here is therefore very similar to the one used with Repetitiveness Checker before, with the one difference of being bilingual. With Repetitiveness Checker, only French texts were tested, because as the source texts they would determine how much repetition could be detected by a TM System

Content	No.
Interventions (Title)	133
Interventions (Text)	367
Total Texts	500

Fig. 8: Statistics test corpus

⁴¹ See chapter Context, p.29.

when translating from French to German. Here, however, every French text has its German equivalent

The keyword was present in the title of 142 texts, a full text search even yields 818 texts,⁴² i.e. enough texts to create an extensive TM. Figure 8 shows the total amount of texts considered in this study, which includes all texts containing the key word in the title, as well as about half of the texts containing it in the text. The test corpus thus amounts to a total of 628 pages in French and 604 pages in German.⁴³ An overview of the text statistics can be found in Annex B (p. 98). The texts chosen to establish the databases date up to the 1st of September 2010, the Texts to test them against were chosen from the texts submitted after that date, i.e. dating back no longer than the autumn session 2010 which took place from the 7th to the 25th of September 2010. Having established the test material and the tools, we are now ready to carry out the tests themselves.

4.2.2 Alignment tests

As mentioned above, this first step of the test arrangements aims at an insight into how the two alignment components, WinAlign (Trados) and TextBase Alignment Agent (MultiTrans), cope with the specific texts, the Parliamentary Interventions. Of course, the ultimate goal of the overall test arrangement is to find out whether a TM system is usable and useful with the particular texts translated in the ZSD-D, and if yes, which one is more suitable (Trados or MultiTrans). In these preliminary tests, the aim is to find out more about the alignment component of the two tools, specifically technical information, as well as to find out how well they align the texts, in order to be aware of potential problems.

Test setup

For the setup of these preliminary tests, I will follow the 7-step recipe proposed by the EAGLES-Group (EAGLES, 1999). The components tested are the alignment components. Accordingly, the user here is not a translator working directly with the TM component but a person in the service responsible for the TM tool as well as the update of the database. This could also be an intern, but in any case this person would have at least some knowledge of the tool, as well as of the texts. Context factors, such as languages and text contents, but also the fact that a massive alignment may be necessary, will be taken into account in the following setup.

42 Search carried out on January 31st 2011.

43 The number of pages was calculated using the number of characters (with spaces), the administrative norm page counts 1800 characters per page (Communication I. Kamber, Annex A).

Top level quality characteristics. The alignment components need to be evaluated for the following features:

- A) **Compliance** with context criteria such as file formats and languages;
- B) **Efficiency**, i.e. time behaviour;
- C) **Reliability**, i.e. the quality of the alignment;
- D) **Maintainability**, i.e. the possibility to correct errors in the alignment as well as in the TM content;
- E) **Customizability**, i.e. the possibility to add abbreviations and abbreviation lists in case of repeated errors caused by this; and
- F) Some **accessory information** such as portability and price.

For each of these top level characteristics, specific tests and metrics have been devised and are detailed below. An overview of all characteristics and tests is available in Annex C, p. 99.

A) Compliance

Compliance with context factors is obviously an important criterion, because if the programme is not compatible with formal factors such as file format, then it is unusable, or at least the effort required for the introduction of the tool, specifically file conversion, is prohibitive. The tests here are therefore conditional. Compliance is needed in our context in particular with respect to two factors: file format and languages.

1 File format

The file formats used in the ZSD-D are Microsoft Office formats, specifically Word (.doc and .docx), as well as some Acrobat Reader .pdf. The working documents, however, are practically all in Word format, .pdf are generally converted before use. The metric applied is boolean, the test procedure is to check for functionalities in the user manuals as well as on the producers' websites.

2 Languages

Language compliance is also very important, but here we have to distinguish between the working languages, the possible interface languages, as well as the possibility to create multilingual TMs including all working languages of the service.

2.1 Working languages

The working languages of the ZSD-D are German (de), French (fr), Italian (it) and English (en). The majority of the texts translated in the service are either French or Italian. The tools therefore need to work with at least de, fr and it, the possibility of en would be an advantage. Again, the metric applied is boolean, the test procedure is to check for functionalities in the tool description.

2.2 Interface languages

All translators in the service are multilingual, or at least bilingual de and fr. However, their native language is German. It would therefore be appropriate that the interface of the tool is available in this language. This is not a condition, because a French or even English interface would be possible too, but they would probably feel more comfortable working in a de environment, especially in the beginning, when they have to get used to a new tool. Again, the metric applied is boolean, the test procedure is to check for functionalities in the tool description.

2.3 Trilingual databases

The texts in our context are written either in fr, it or en and are translated into German. In the Curia Vista database online, all the Parliamentary Interventions are available in three languages, de, fr and it, but not in en. It would therefore be helpful if a multilingual TM is possible with the tool, at least for the three main languages. This is not a condition, as the translators are never working with two source languages simultaneously, but it would simplify the setup and maintenance of the TM. Again, the metric applied is boolean, the test procedure is to check for functionalities in the tool description.

The tests for 'Compliance' can be summarized as follows:

Criterion			Test Procedure	Metrics
1. Compliance with file format		MS Word (.doc and .docx), Pdf	Check functionalities	yes/no (boolean)
2. Compliance with languages	2.1 Interface language	de	Check functionalities	yes/no (boolean)
	2.2 Working language	de/fr/it/en	Check functionalities	yes/no (boolean)
	2.3 Multilingual database	de/fr/it (en)	Check functionalities	yes/no (boolean)

Table 1: Summary Compliance

B) Efficiency

The efficiency of the tools can be measured by the speed with which they align the texts. For this test a certain number of texts will be aligned, in order to find out how long it takes the tools to carry out the task. Obviously, when aligning hundreds or even thousands of documents, speed is an important criterion. Therefore, the tools should align as fast as possible. The time necessary will be measured with a standard stopwatch.

1. Alignment of 5 texts

I have chosen five Texts from the corpus which present recurring patterns such as lists and bullets, abbreviations, titles etc. They were chosen randomly among texts presenting the search word in the title and with a minimum length of 400 words (Total of 3'141 words in French, 2'604 words in German). The de and fr versions of these texts are aligned. The five texts are:

Number		Author	Title de/fr	Words de	Words fr
1)	99.3309 - Motion	Berberat Didier	KVG. Schaffung einer Schweizer Risikogemeinschaft / LAMal. Création d'une communauté de risques au niveau suisse	502	563
2)	98.3487 - Motion	Saudan Françoise	KVG. Notwendigkeit einer effizienten Kontrolle / LAMal. Nécessité d'un contrôle efficace	367	442
3)	98.3377 - Interpellation	Borer Roland F.	KVG-Grundversicherung. Selektiver Rückzug eines Krankenversicherers / Assurance de base dans la LAMal. Retrait d'une caisse d'assurance-maladie de différents cantons	648	691
4)	09.3275 – Motion	Humbel Ruth	Einführung des Monismus im KVG / Instauration du système moniste dans la LAMal	373	543
5)	03.468 – Parlamentarische Initiative	Sommaruga Simonetta	KVG. Klare Bedingungen für die Grundversicherung / LAMal. Assurance obligatoire des soins. Clarifier la situation	714	902

Table 2: Texts for Alignment (5)

2. Alignment of 50 texts

As the number of texts necessary for these preliminary tests is quite small, I have chosen to carry out a second test, with 50 texts, also chosen arbitrarily among texts presenting the search word in the title, but without attention to their length (Total of 15'371 words in French, 12'782 words in German). This second test should serve as an indicator in case of

programmes that are slow in the beginning but speed up afterwards, so that the difference between aligning few and many texts is small. Again, the de and fr version of these texts are aligned.

The tests for 'Efficiency' can be summarized as follows:

Criterion		Test Procedure	Metrics
Time behaviour	Speed	1. Align fr/de texts (5 texts)	min:sec
		2. Align fr/de texts (50 texts)	min:sec

Table 3: Summary Efficiency

C) Reliability

Reliability, in other words the quality of the alignment, is a very important criterion, for any TM creation. For a small TM where alignment is verified, less errors mean less work, because fewer segments need to be corrected. If, however, the alignment is not verified, for example because the TM is simply too big, the quality of the alignment is even more important. The fewer alignment errors, the better the final quality of the TM. For this test, the five texts previously aligned for the efficiency test will be used again. With the alignment done I will then proceed to an error count. Out of the total number of segments I will count the number of fully exploitable ones, as opposed to incorrectly aligned ones. This will allow me to calculate the number of incorrect alignments as well as the percentage of correct alignments. The higher the percentage and the lower the number of incorrect alignments, the better, because this means less work and better TM quality.

The tests for 'Reliability' can be summarized as follows:

Criterion		Test Procedure	Metrics
Quality of alignment	1. No. of errors (incorrectly aligned segments)	Align fr/de texts, segment count	Incorrectly aligned segments / total number of segments
	2. No. of errors (percentage of fully exploitable segments)	Align fr/de texts, segment count	%

Table 4: Summary Reliability

D) Maintainability

Maintainability here means that it is easy to make changes, that the tool is easy to use in a concrete situation. The characteristic of maintainability can be divided into two main parts, correcting errors in the alignment and correcting errors in the TM content. For the errors in alignment, I will distinguish segmentation errors from alignment errors. A segmentation

error occurs, when the sentence boundaries were not respected in the segmentation, meaning that the sentence has been cut in the wrong place, as in the following example:⁴⁴

1. L'article 39 LAMal prévoit que la planification des besoins en soins hospitaliers établie par les cantons couvre les besoins de un ou de plusieurs cantons. La loi prévoit donc	1. Die geltende Regelung in Artikel 39 KVG sieht vor, dass die von den einzelnen Kantonen zu errichtende bedarfsgerechte Spitalplanung die stationäre Versorgung eines oder
déjà la possibilité d'opérer une planification hospitalière par-delà des frontières cantonales.	mehrerer Kantone umfassen muss. Das KVG sieht somit bereits die Möglichkeit von Spitalplanungen über die Kantonsgrenzen hinaus vor.

Table 5: Segmentation error

The term alignment error, however, concerns passages where the sentences were segmented correctly, i.e. according to their sentence boundaries. One correct segments was thus aligned with another correct segment, but not the corresponding one, as is the case in the following example:

1. L'article 39 LAMal prévoit que la planification des besoins en soins hospitaliers établie par les cantons couvre les besoins de un ou de plusieurs cantons.	Das KVG sieht somit bereits die Möglichkeit von Spitalplanungen über die Kantonsgrenzen hinaus vor.
La loi prévoit donc déjà la possibilité d'opérer une planification hospitalière par-delà des frontières cantonales.	Diese Möglichkeit bi- oder gar multilateraler Spitalplanungen wurde bis heute allerdings nur in sehr geringem Umfang wahrgenommen (Beispiel: gemeinsame Spitalplanung und gemeinsame Spitalliste der Kantone Basel-Stadt und Basel-Landschaft).

Table 6: Alignment error

A correct segmentation and alignment of this passage would look as follows:

1. L'article 39 LAMal prévoit que la planification des besoins en soins hospitaliers établie par les cantons couvre les besoins de un ou de plusieurs cantons.	1. Die geltende Regelung in Artikel 39 KVG sieht vor, dass die von den einzelnen Kantonen zu errichtende bedarfsgerechte Spitalplanung die stationäre Versorgung eines oder mehrerer Kantone umfassen muss.
La loi prévoit donc déjà la possibilité d'opérer une planification hospitalière par-delà des frontières cantonales.	Das KVG sieht somit bereits die Möglichkeit von Spitalplanungen über die Kantonsgrenzen hinaus vor.
Jusqu'à présent, très peu de cantons ont fait usage de cette possibilité en bilatéral ou en multilatéral (exemple: les deux cantons de Bâle, qui ont opéré leur planification ensemble et établi une liste des hôpitaux commune).	Diese Möglichkeit bi- oder gar multilateraler Spitalplanungen wurde bis heute allerdings nur in sehr geringem Umfang wahrgenommen (Beispiel: gemeinsame Spitalplanung und gemeinsame Spitalliste der Kantone Basel-Stadt und Basel-Landschaft).

Table 7: Correct segmentation and alignment

⁴⁴ Source: 03.3042, Postulate by Christine Wirz-von Planta

1 Correcting alignment errors

It is possible that some texts are aligned correctly and that there are no corrections to be made. However, when aligning automatically, it is more likely that there are at least some errors, in which case the correction should be possible and require as little effort as possible. For the tests under this heading, the five aligned texts from the previous test, i.e. the reliability test, are used and corrected. This test can be subdivided into three steps:

1.1 Possibility to correct alignment errors

For this preliminary question, the functionalities of the tools are checked and the metric applied is boolean.

1.2 Correcting alignment errors

Rather than correcting specific alignment errors, such as 1-1 or 2-1 alignments, I will correct all errors found in the five texts aligned for this test. I have chosen this procedure mainly because the incorrectly aligned passages are never completely identical. Also, I believe it to be more pertinent with respect to the final goal of these preliminary tests. As mentioned above, this test represents a test run for the scenario test, and should allow me not so much to compare the two programmes against each other but to detect specific problems and possibilities to avoid them in the scenario test. In this sense, correcting all alignment errors will give me an opportunity to see the causes for incorrect segmentations or alignments.

For these reasons, having chosen my sample of texts arbitrarily, the number of mouse-clicks necessary to correct all alignment errors, with five fully exploitable and correctly aligned texts as final result, gives me a better insight into the work required for the scenario test, because it will show me about how much work the correction of five texts potentially represents. I will therefore proceed to a correction of all errors, until all texts are fully exploitable. The metric applied here is again the number of mouse-clicks necessary to apply the corrections. Drag and drop actions count as 2 clicks.

1.3 Unsatisfactory correction

It is possible that not all errors can be corrected, due to some software-internal problem. Often, the cause of such a situation cannot be determined unequivocally. However, if there are many such alignment errors which cannot be corrected satisfactorily, this would constitute a serious problem for the alignment quality. I will therefore list the number of problematic passages which cannot be corrected entirely or in a satisfactory manner, if any should occur. The metrics applied is simply their number.

2 Correcting content errors

When it comes to content errors, typing errors for example, this can happen to anyone. Also, when checking the alignment, the attention paid to the text can reveal errors which should not be perpetrated through the TM. Of course, this does not include translation mistakes or mistakes in the content of the text but simple typing or spelling errors. Again, the test can be subdivided into two steps:

2.1 Possibility to correct content errors

For this preliminary question, the functionalities of the tools are checked and the metric applied is boolean.

2.2 Correction of a content error

For this test, the same typing error in one of the texts will be corrected in the tools. The metric applied is the number of mouse-clicks necessary to apply the correction, without counting the clicks or keyboard strikes needed for the correction itself.

The tests for 'Maintainability' can be summarized as follows:

Criterion		Test Procedure	Metrics
1. Correction of alignment errors	1.1 Possibility	Check functionalities	yes/no (boolean)
	1.2 Correction of alignment errors	Correct all alignment errors in 5 texts	no. of clicks
	1.3 Unsatisfactory correction	Count number	number
2. Correction of content errors	2.1 Possibility	Check functionalities	yes/no (boolean)
	2.2 Correction of content errors	Correct 1 content error	no. of clicks

Table 8: Summary Maintainability

E) Customizability

Once the tool is in place, it needs to be customized in order to achieve the best results. This is not only important for the following scenario test, where a great number of texts will be aligned, but particularly in a real situation where the quality of the TM and the alignment speed depend on an efficient adaptation to context requirements. Here, it is especially important to adapt the segmentation rules, for example to add abbreviations and abbreviation lists, as abbreviations often cause incorrect segmentation.

1 Addition of abbreviation

Sometimes it can be necessary to add just one abbreviation, especially if there are not many different abbreviations present in the texts. This part can be subdivided into two steps:

1.1 Possibility to add abbreviation

For this preliminary question, the functionalities of the tools are checked and the metric applied is boolean.

1.2 Addition of abbreviation

For this test, I will add an abbreviation to the segmentation rules of the tools. The abbreviations added are *al.* (for *alinéa*) for the French rules or *Abs.* (for *Absatz*) for the German rules. These are recurring abbreviations and their addition could potentially reduce the number of segmentation or alignment errors made by the systems. The metric applied here is the number of mouse-clicks necessary to apply the change. The keystrokes needed to write the abbreviation are not counted.

2 Addition of abbreviation list

When there are many specific abbreviations present in the texts, it might be worth while to make a separate list and add it to the segmentation rules. Again, this part can be subdivided into two steps:

2.1 Possibility to add abbreviation list

For this preliminary question, the functionalities of the tools are checked and the metric applied is boolean.

2.2 Addition of abbreviation list

For this test, I will constitute a short list of abbreviations found in the texts I will align. I will then add this list to the segmentation rules of the tools. The content of the list is secondary here, but it will of course include the abbreviations mentioned in the test 1.2 above. The metric applied is the number of mouse-clicks necessary to add the list, rather than the content. The clicks needed to browse in the system in order to find the file are not counted.

The tests for 'Customizability' can be summarized as follows:

Criterion			Test Procedure	Metrics
Possibility to change segmentation rules	1. Add abbreviation	1.1 Possibility	Check functionalities	yes/no (boolean)
		1.2 Adding abbreviation	Add abbreviation	no. of clicks necessary
	2. Add abbreviation list	2.1 Possibility	Check functionalities	yes/no (boolean)
		2.2 Adding abbreviation list	Add abbreviation list	no. of clicks necessary

Table 9: Summary Customizability

F) Accessory information

There is some information which is not important in itself to assess the tools or to better understand how they work the texts, but which is interesting with respect to a future acquisition of the tool. I have decided to include this information here for description only.

1 Portability

With respect to portability it is interesting to know whether the TM can be exported in a universally accepted format, i.e. Translation Memory eXchange (.tmx). There are only few occasions where external translators are contacted by the ZSD-D, but nevertheless this might be an interesting factor, keeping possibilities open. For this test the functionalities of the tools are checked and the metric applied is boolean.

2 Price

The amount of money necessary to purchase the system is another factor which cannot be decisive for this paper, since I have no authority in this matter, but which can give some indication as to whether it is realistic to consider an acquisition or not. For this test the tools' websites are checked and, if necessary, the developers are contacted for more information.

The tests here can be summarized as follows:

Criterion		Test Procedure	Metrics
1. Portability	.tmx export	Check functionalities	yes/no (boolean)
2. Price	for 1 licence	check on website/contact developers	CHF

Table 10: Summary Accessory Information

Results

The tests were carried out in January 2011, in the Salle informatique 6281 TAO, École de traduction et d'interprétation, Geneva. They were all carried out on the same computer, in order to avoid processor related differences. The presentation of results will follow the same structure as the above presentation and will be summarized in a table for each.

A) Compliance

Concerning the compliance criteria, both programmes satisfy the needs of our context. Both work with Microsoft Word documents (.doc as well as .docx). TextBase also opens Pdf files, which is not possible with WinAlign. However, since most of the working documents are Word documents, this should not cause any problems.

Also, both programmes comply with all the required interface and working languages. MultiTrans offers German as interface language from version 4.3. They also both allow the creation of multilingual databases, even though this version of the programme could be slightly more expensive. Concerning these formal factors, therefore, both programmes are equally suited for our context.

Criterion			Metrics	WinAlign	TextBase
1. Compliance with file format		MS Word (.doc and .docx), Pdf	yes/no (boolean)	yes	yes
2. Compliance with languages	2.1 Interface language	de	yes/no (boolean)	yes	yes
	2.2 Working language	de/fr/it/en	yes/no (boolean)	yes	yes
	2.3 Multilingual database	de/fr/it (en)	yes/no (boolean)	yes	yes

Table 11: Results Compliance

B) Efficiency

In this test the alignment speed was measured and both programmes are very close, for few texts as well as for many. The five main texts for these tests were aligned by both programmes in 6 seconds. The 50 texts were aligned by WinAlign in 25, by TextBase in 26 seconds. There may therefore be a slight difference in speed when more texts are aligned. However, with the tests done here this could not be confirmed, and even if this should be true, the difference would probably not be very great.

Criterion		Metrics	WinAlign	TextBase
Time behaviour	Speed	min:sec	00:06	00:06
		min:sec	00:25	00:26

Table 12: Results Efficiency

C) Reliability

For this test, the alignment errors were counted, in order to know the percentage of fully exploitable segments for each programme. WinAlign has segmented the five texts into a total of 139 segments, of which 127 were fully exploitable, with 12 incorrectly aligned ones. This corresponds to a total of 91.4 percent of fully exploitable segments. TextBase produced slightly fewer segments, namely 132, of which 117 were fully exploitable, and 15 incorrectly aligned. This corresponds to a total of 88.6 percent of fully exploitable segments. It therefore seems that WinAlign is slightly more appropriate for our type of text. However, the difference is not great enough to produce a clear preference at this stage.

Criterion		Metrics	WinAlign	TextBase
Quality of alignment	1. No. of errors (incorrectly aligned segments)	Incorrectly aligned segments / total number of segments	12/139	15/132
	2. No. of errors (percentage of fully exploitable segments)	%	91.4	88.6

Table 13: Results Reliability

D) Maintainability

With respect to maintainability, both programmes present the possibility to change alignment errors. However, this is not the case for the correction of content errors.

The correction of errors in the alignment was necessary with both programmes. Most problems were caused by incorrect segmentation and interestingly, both programmes had

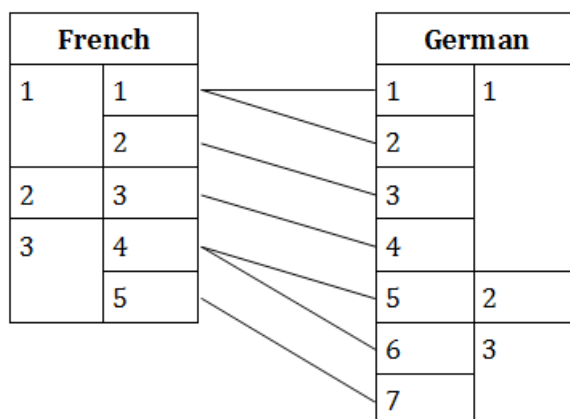


Fig. 9: WinAlign alignment error (99.3309)

difficulties with the same texts and passages. Two texts, 09.3377 and 98.3487, were aligned without problems, with 100% of fully exploitable segments. The most problematic text was 99.3309, where several passages were not segmented correctly. The other two texts, 03.469 and 09.3275, presented just one or two problematic passages. These incorrect

segmentations normally resulted in entire passages which were not well aligned. To illustrate, one example from WinAlign (fig. 9).

The columns on the far left and the far right represent the correct segmentation, respecting content and form of the text, while the inner columns represent the actual segmentation done by WinAlign. As you can see, there are not only unnecessary segmentations inside the sentences, but also wrong alignments of these segments. An incorrect segmentation of just one sentence can thus cause for incorrect alignment even several segments down in the text. The incorrect segmentation was often due to unknown abbreviations in the text,⁴⁵ as well as other punctuation marks, such as a colon. This resulted in a higher number of segments than actually required by the text. However, with WinAlign, even though sentences might be split into several segments, as in the above example, the correct segment boundaries were normally found and respected.

With TextBase, on the other hand, this was often not the case, as the following example shows (fig. 10). As you can see here, the correct segment boundaries were not respected, between S1 and S2 for the French text, and between S2 and S3 for the German Text. The result is therefore not necessarily a higher number of segments, but the same or a similar number with

French		German	
1	1	1	1
	2	2	
2		3	2
	3		3
3	4	4	

Fig. 10: TextBase alignment error (99.2209)

incorrect segment boundaries. In this case, a correction of the alignment is even more time-consuming because before realigning the segments, the correct segment boundaries have to be established. These difficulties were almost always caused by abbreviations or numbers followed by a full stop in the text.⁴⁶

Accordingly, correcting the alignment errors in TextBase was much more laborious, as the number of mouse-clicks necessary shows: 46 with WinAlign, 92 with TextBase, exactly twice as much. The difference is considerable, and it is only partly explained by the number of problematic passages for each programme (4 for WinAlign, 6 for TextBase). Where both programmes had difficulties with the same passage, TextBase consistently required more clicks than WinAlign. For example, in the case of Text 03.469, where the same passage was problematic for both programmes, WinAlign required only 8 clicks, while TextBase required 18.

45 Specifically *al.* for *alinéa* and *art.* for *article* in the French texts, as well as *Abs.* for *Absatz* and *Art.* for *Artikel* in the German texts.
 46 For example dates in German: 14. November.

For the third point examined concerning alignment errors, with WinAlign all problematic passages were finally well aligned after correction, whereas two situations in TextBase could not be resolved in a satisfactory manner. In one case, the only possible solution was to join two segments in the German text, since the two sentences could not be separated in the French text. The other situation concerned an incorrect segmentation involving a number followed by a full stop, where it was impossible to enforce the correct segment boundary. Here again, the final solution was not satisfactory. Both difficulties were located in the most problematic text, 99.3309. Overall, we can thus say that WinAlign not only seems to have less problems with the alignment of our specific texts but also that correction of alignment errors is easier and more efficient than with MultiTrans.

Tests on maintainability also included the possibility to correct content errors while aligning. This is possible and relatively easy, with only 4 clicks, with WinAlign, while TextBase does not offer this possibility. With TextBase, if the content of the texts is changed, they have to be realigned entirely. Here, WinAlign thus offers an option more than TextBase.

Criterion		Metrics	WinAlign	TextBase
1. Correction of alignment errors	1.1 Possibility	yes/no (boolean)	yes	yes
	1.2 Correction of alignment error	no. of clicks	46	92
	1.3 Unsatisfactory correction	number	0	2
1. Correction of content errors	2.1 Possibility	yes/no (boolean)	yes	no
	2.2 Correction of content error	no. of clicks	4	-

Table 14: Results Maintainability

E) Customizability

To test the possibility to customize, I have tried to adapt the segmentation rules. In both programmes it is possible to add an abbreviation list, and equally easily. When it comes to single abbreviations, customization is possible with MultiTrans, but it is quite laborious. It includes finding the preferences for the TextBase, opening the abbreviation list already activated, and finally adding the abbreviation. A total of 12 mouse-clicks were necessary to carry out this operation. This system is, however, quite convenient if one only wants to modify the default list associated with the TM, which is already quite complete. Adding one single abbreviation is not possible with WinAlign.

Even though this was not part of the tests carried out here, it deserves mentioning that with both programmes it is possible and fairly easy to modify or add segmentation rules.

This can be interesting for certain texts with specific sentence boundaries, or to exclude segmentations at colons for example. For our texts, this is not particularly important though.

Criterion			Metrics	WinAlign	TextBase
Possibility to change segmentation rules	1. Add abbreviation	1.1 Possibility	yes/no (boolean)	no	yes
		1.2 Adding abbreviation	no. of clicks necessary	-	12
	2. Add abbreviation list	2.1 Possibility	yes/no (boolean)	yes	yes
		2.2 Adding abbreviation list	no. of clicks necessary	5	5

Table 15: Results Customizability

F) Accessory Information

With respect to portability, both programmes satisfy our requirements in that they both allow exportation in the generally compatible .tmx format for translation memory databases.

For the price, I had to contact the two companies developing the programmes in order to find out the price of a server-based solution allowing for multilingual TMs. Unfortunately, and despite several attempts on my part, no answer could be obtained from either of the developers. Multicorpora does not publish any prices on their website, while for SDL Trados the price of the most sophisticated desktop version, SDL Trados Studio 2009 Professional is advertised at € 2'990 on their website. For server versions, no prices are published. This factor thus remains unknown.

Criterion		Metrics	WinAlign	TextBase
1. Portability	.tmx export	yes/no (boolean)	yes	yes
2. Price	for 1 licence	CHF	-	-

Table 16: Results Accessory Information

Conclusions and measures taken

Overall we have to say that both programmes are equally adapted or at least very close on most points. This is true for all compliance issues as well as portability. Also, even if WinAlign performed slightly better in alignment quality, the difference was small, as was the difference in the alignment speed. The main difference, then, is to be found in maintainability and customizability. For the latter, the difference lies mainly in the fact that with TextBase it is possible to add single abbreviations, while this is not the case for WinAlign. However, the only consequence of this test is that I will have to write and add a list of abbreviations for the TMs created for Trados, while for MultiTrans I can choose to use the default list and just add one or two abbreviations. However, the effort needed for this task is relatively similar with both

programmes. Overall, the difference here does slightly change my way of proceeding but it does not create a problem which could disqualify one or the other of the programmes.

With respect to maintainability, however, the difference in performance between the two programmes needs to be considered in detail. Indeed, the fact that the correction of segmentation and alignment errors is more time-consuming with TextBase can be an important factor for the scenario test. In terms of mouse-clicks, TextBase has required exactly double the effort necessary with WinAlign, a fact which definitely will have an impact on the time spent on the creation of any TM. As it is not possible to change the number of clicks necessary to carry out one correction, the important question here is whether it will be possible to improve the segmentation with certain measures, in order to maximise the alignment quality and to minimise the effort needed to correct remaining mistakes.

One of the possible measures is to add the problematic abbreviations to the systems, specifically *al.*, *art.*, *Abs.* and *Art.*, which have caused segmentation errors in the five texts aligned during these tests. As mentioned above, I will have to create an abbreviation list for WinAlign, while with TextBase I can add them directly to the pre-existing file. Subsequent tests with the same five texts have shown that this measure has eliminated all previous segmentation errors in WinAlign and some in TextBase, even though alignment errors still persist. The remaining segmentation errors in TextBase are more difficult to eliminate. As we have seen above, most segmentation errors in TextBase are caused by numbers followed by a full stop. The only possible measure was to add the numbers 0-9, followed by a period, to the abbreviation list, so that they would not be considered the end of a sentence. For the German texts, however, it was necessary to add numbers up to 31, because of the general format for dates (e.g. *14. November*).

With these two measures, i.e. the addition of abbreviations and of numbers, I have been able to slightly improve alignment quality. I have repeated the test for reliability carried out above with the same texts, after applying the measures. For WinAlign, the percentage of fully exploitable segments is up to 93.8 percent, with a total of 8 incorrect alignments out of 130. With TextBase, the percentage is still slightly lower, at 92.9 percent, but there are also only 9 incorrect alignments out of 127. It is therefore possible to improve the quality of alignment, and as a consequence, the effort needed for correction, even if the difference is small. To be noted also that the total number of segments for the five texts has decreased with both programmes, confirming the fact that unnecessary segmentations have been eliminated.

Criterion		Metrics	WinAlign	TextBase
Quality of alignment	1. No. of errors (incorrectly aligned segments)	Incorrectly aligned segments / total number of segments	8/139	9/132
	2. No. of errors (percentage of fully exploitable segments)	%	93.8	92.9

Table 17: Alignment quality (improved)

Overall, these preliminary tests aimed at a best possible preparation for the subsequent scenario test. We know now that there are no major problems encountered by the programmes when working with the required texts. There is, of course, room for improvement, but not too much. Also, these tests have allowed me to improve the alignment with a few measures. These measures, namely the abbreviation lists and the addition of numbers for TextBase will be retained for the setup of the scenario test

4.2.3 TM size

We now know more about how the alignment tools need to be configured in order to obtain the best results. However, before setting up the scenario test, where an actual translation situation will be recreated, there remains the question of size to be answered. In chapter 3.4 (p. 29) above, when testing the texts with the tool "Repetitiveness Checker", we stipulated that an extensive TM would probably be more useful than a small one. This hypothesis needs confirmation, and this will be attempted by this second part of the test arrangements.

Test setup

In order to compare the usefulness of a small and an extensive TM database, a total of four databases were created, two with each tool. For each tool, one relatively small TM was created, by aligning 15 texts and their translations, a so-called Mini-TM, as well as an extensive one, a Maxi-TM with 500 texts. In order to measure the impact of the different TMs, rather than translating a text with them, the available analysis features of the tool were used on one text.⁴⁷ Here, the goal is not yet to test the overall usefulness of the tools, but to find out which of the two, a small or an extensive TM, yields more results.

The texts for this test were selected from within the set of texts described above. For the Maxi-TM, all 500 texts were used, for the Mini-TM, a set of 15 texts were chosen from the set of texts featuring the key word in the title. Also, they were selected for their length, counting at least 500 words each, the longest counting 973 words. The alignment of the Mini-

⁴⁷ The test document was 10.3656, Interpellation by Liliane Maury-Pasquier, "Les primes d'assurance-maladie prennent leur envol et les réserves s'envolent" (754 words, 4857 characters with spaces).

TM was carefully checked and alignment errors were corrected, in order to obtain a high quality database. For the Maxi-TM, this was unfortunately not possible, due to practical reasons. Indeed, checking the alignment of 500 texts would have taken up too much time for the scope of this study. The fact that the alignment quality of the Maxi-TM will be slightly inferior to the one of the Mini-TM needs to be taken into account therefore. In the next two sections, the detailed setup and procedure will be described.

SDL Trados

With the SDL Trados alignment tool WinAlign, the two TMs were created, following the protocol established through the preliminary tests above with regard to abbreviations. For the Maxi-TM, the 500 texts were aligned without any checking of the segmentation and alignment quality. For the Mini-TM, however, the 15 texts were aligned and carefully checked. Any segmentation or alignment errors were corrected before use, ensuring 100 percent alignment quality. In SDL Trados Studio 2009, a project was set up for each TM, featuring the test document. The comparison was then done with two instruments. First of all, the text was analysed with the internal analysis tool, allowing for a comparison of the number of exact and fuzzy matches. Secondly, the text was opened in the editor, and a list of 10 terms was chosen for a context search.⁴⁸ These terms are:

primes d'assurance-maladie	Conseil fédéral
Office fédéral de la santé publique	transfert de réserves
procédure d'approbation des primes	redistribution des primes
cantonalisation	la maîtrise des coûts de la santé
base légale	augmentation des primes

Table 18: Search terms SDL Trados (Context search)

The usefulness of the two TMs was thus compared in terms of exact and fuzzy matches, but also for the impact of the context search. However, the quality of the proposed matches was not evaluated.

MultiTrans

For MultiTrans, too, the protocol established above was respected for abbreviations and numbers. The two TMs were thus created, and again, only for the Mini-TM the segmentation and alignment were checked and corrected where necessary. The Maxi-TM was aligned without checking. They were both connected in turns to the test document opened in MS Word. In order to compare the impact of each TM, the TextBase Agent was run, highlighting all

⁴⁸ In this function, the highlighted text or expression is looked up in the database and the segments containing it with their translations are proposed.

segments which can be found in the Database. The impact is compared through a word count of the highlighted segments. What needs to be taken into account, however, is that The TextBase Agent does not only highlight entire segments present in the database, but also terms of two or more words and other smaller chunks. Also, as with SDL Trados, the quality of the proposed material was not considered.

Results

For SDL Trados, a difference of impact could be felt, but mostly with respect to the context search. Indeed, the analysis of the test document was the same for both TMs. For a total of 33 segments, both TMs found 3 exact or almost exact matches (95 to 100%) and 3 repetitions. No correspondence was found for 27 segments. However, in the Mini-TM only 4 of the terms could be found, while the Maxi-TM featured 9 out of 10. With MultiTrans, the difference was even greater. With the Mini-TM, the TextBase Agent highlighted 81 of 754 words, i.e. 10.7%, while with the Maxi-TM it highlighted 250 words (33.2%).

Conclusions and impact

The above results show a clear difference between the impact of a small TM and that of an extensive one and the test thus seems to confirm the hypothesis stipulated in chapter 3.4 above. While the results were the same with respect to matches in SDL Trados, especially the search for terms and smaller chunks was much more helpful with the Maxi-TM. With MultiTrans, too, the difference between the impact of a Mini- and that of a Maxi-TM is clear. The usefulness of the systems does not so much repose on the retrieval of entire sentences therefore, but on the availability of smaller chunks and terminology. It is probable that the positive impact of the exact and fuzzy matches is reduced for the extensive TM by the fact that the alignment was not verified. For practical reasons, this was not possible here, but for future reference this fact needs to be kept in mind. The best option would be an extensive TM where the alignment has been checked and errors corrected. However, for our purpose here we will retain the general conclusion that an extensive TM, even if the segmentation and alignment had not been checked, is more useful than a small one, and this test result will be respected for the setup of the subsequent scenario test in that only Maxi-TMs will be created for it. This test result also gives a first indication as to what might be more important when working on texts with low repetitiveness. It does seem that the texts are too varied to benefit from many exact and fuzzy matches. On the other hand, it seems that the tool might still be useful for translation because of the terminology stored in the TM database, and therefore as a tool it might still be beneficial for a translation service such as the ZSD-D.

4.2.4 The scenario test

This third and main part of the tests consists of a scenario test and will attempt to find out more about how the two programmes and the TMs would work and how they could be appreciated in a concrete work situation. For this, the Maxi-TMs established for the above test of the TM size were used again, this time for translation with the tool, giving a more realistic idea as to how useful and how user-friendly the two programmes would really be in our context. What needs to be kept in mind, however, is that at least part of this test cannot be taken as statistically valid. As we have seen above, the number of participants in a scenario test is a crucial factor and here the size and purpose of this study unfortunately do not allow us to set up a group of test participants large enough to gain objective insights into the tools' usefulness. Also, when actually working with the programmes, the concepts of usefulness and user-friendliness are easily mixed up to create a general appreciation.

I therefore understand this part of the scenario test as more descriptive, a pilot project with few but well chosen test translators. It will not provide any definite answers, but rather ideas and suggestions about whether and how the tools could work in the specific environment of the ZSD-D, and thus information upon which a decision for further testing could be based. This quite responds to the questions of the ZSD-D, but when it comes to answering the more general question of whether a TM system would be useful in any translation service working with low repetitivity texts, a larger test arrangement would have to be set up with much more test participants. For this, of course the same methodology and procedure could be used again.

Test setup

A translation scenario was set up as close to the reality in the ZSD-D as possible. The texts were chosen from the texts actually translated in the service, and the participants translated four texts into German, with the following conditions: one was translated manually, without any TM system, one with the help of SDL Trados and two with MultiTrans, using the same programme in two different ways. For one translation, MultiTrans was used simply as a database, i.e. a bilingual aligned corpus. This TextBase Research could be used in a separate window or as automatic lookup. The only other MultiTrans function used was the TextBase Agent which highlights material existing in the TM. For the other translation, the MultiTrans Translation Agent was used, an interface similar to SDL Trados, where exact and fuzzy matches are suggested by the programme and can be inserted automatically. The fact that two texts are translated with MultiTrans accounts for the multiple way this programme can be

used, as well as for the clear difference in use between a bilingual aligned corpus and a translation memory with recall function.

With the help of a questionnaire information on usefulness and user-friendliness were gathered at the end, so that the impact of the different systems could be compared. Another instrument for comparison was the time necessary for the entire process, from the establishment of the TM to the translation. From these two inputs, a final ranking could be established, to see which translation process was the most efficient and appreciated. But let us first look again at the main ingredients: the participants and the texts. The TMs used in this test were created with the 500 texts which constitute the test corpus described in section 4.2.3 above. They were established respecting the conclusions drawn from the alignment tests, concerning abbreviations, numbers and the overall size of the database. The TMs are thus very close to what could be possible in a real situation, except for its limited size and thematic scope. If a TM of Parliamentary Interventions was created for the ZSD-D, it would of course be much more extensive, given the enormous amount of texts available, and not restricted to the theme of health policy, as is the case here.

The participants were chosen according to the following criteria: they had to have completed an internship at the ZSD-D within the last 2 years, thus ensuring that they know the specific context of this service, as well as the texts. Also, they had to be students or recent graduates from the ETI in Geneva, a factor which grants that the participants have at least some knowledge of the tools, since in the curriculum today, the introductory course into translation technology may not be compulsory but is generally difficult to avoid. Following these criteria, three participants could be found for the tests, myself and two others. This limited number obviously cannot be entirely satisfactory, but as the tests demanded considerable time and effort from the participants, it was difficult to find more than three. Still, and due to the selection criteria, I believe that the participants' opinion and impressions are relevant for our purpose and, if they cannot provide definite answers, at least they can provide a good first impression.

The texts for translation were chosen from the same database as the texts in the TMs, Curia Vista, and with the same key word. However, they were chosen from the more recent texts. The test corpus for the alignment tests and the TM size were established at the beginning of September 2010, including the texts submitted during the summer session which ended on the 18th of June 2010. The texts chosen here date back no longer than September 2010, thus ensuring that they are not yet included in the TMs established for the above tests. The four source texts in French are the following:

Number		Author	Title	Words	Characters
1)	10.3741 – Motion	Baettig Dominique	LAMal. Introduction du modèle de Singapour	315 (343)	2226 (2413)
2)	11.403 – Parliamentary Initiative	Meyer-Kaelin Thérèse	Plafonnement annuel de la contribution aux frais de séjour en cas d'hospitalisation	315	1996
3)	10.4036 – Interpellation	Maury-Pasquier Liliane	Des nouveaux-nés victimes de "désintégration"	317	2132
4)	10.3678 – Interpellation	Aubert Josiane	Pilotage et monitoring des coûts à charge de la LAMal	301 (358)	1910 (2279)

Table 19: Source texts for the scenario test

Texts one and four were slightly shortened to balance the length of the texts around 300 words, in both cases the last sentence was cut. Otherwise they were unchanged and presented in an MS Word document without any special formatting. In total they represent some 4.6 norm pages (1800 characters). The order of the four texts was chosen arbitrarily: since I was part of the test participants, I could not study the texts beforehand. For the translations, the participants received the text in electronic format, as well as a copy on paper.

Test procedure

A sequence of tasks was established for the test, which is represented in the following table:

Step	Manual	Trados	MultiTrans TextBase research	MultiTrans Translation Agent
1) TM Setup	-	Abbreviation lists	File list with ListBuilder	
	-	Memory setup	Abbreviations	
2) Alignment	-	Upload files, align	Upload files, align	
3) Translation	-	Project setup	Project setup	Project setup
	Translate text	Translate text	Translate text	Translate text
4) Evaluation	Questionnaire	Questionnaire	Questionnaire	Questionnaire

Table 20: Task sequence

Each step of the process was carried out according to a precise task list which can be found in Annex E (p. 102). Steps 1, 2 and 3.1 were carried out by myself, steps 3.2 as well as 4 by the test participants. Steps one and two were executed on the same computer in order to rule out processor-related differences. One important choice concerning the test setup was the order in which the different versions were carried out, because this can have an impact on the results. I have decided to begin with the manual translation, mainly because the participants do not use TM tools regularly for their work. This way, the participants were able to familiarise themselves with the general theme of the texts without yet having to worry about

using the tool properly. Second, I have chosen Trados before MultiTrans, mainly because of external reasons. After the first two translations, the participants are likely to demand a longer break, e.g. to eat lunch and refresh their concentration. Choosing Trados as the second version allowed me to have the two MultiTrans versions together, which reduces the loss of information on MultiTrans between the translations. It should not be forgotten that professional translators are generally expected to translate 4 to 5 pages per day, and that the 4.6 pages here thus represent roughly a day's work. Measures thus had to be taken to accommodate the test participants.

For the translation quality, participants were instructed to produce a version which they would submit to revision, not a final one. This instruction intended to lift pressure and ensure that they were able to use part of their time evaluating the tool during work. It was possible also because of the fact that the translation quality was not evaluated during this test. An analysis of the products would indeed have complemented this study in that it would have to some extent answered the question of whether the tools help the translators to translate better. However, the scope of this study unfortunately does not allow to explore all aspects of our subjects, and here the usefulness and user-friendliness of the tools received priority over the rather vague question of translation quality.

Time was measured for steps one to three, with the help of a standard stopwatch and added to obtain a total amount of time necessary for the setup and translation of each text. Information was collected through these time measurements as well as the questionnaire, allowing for two rankings: one according to the time necessary for the entire procedure, and one according to the questionnaire feedback. Different weights were attributed to the two ranks, 1 for the time rank, 2 for the feedback rank. The difference of weight reflects the relative importance of the information. According to the overall goals of this study, the usefulness in a concrete situation is more important than the time necessary for the TM establishment, since this would have to be done only once or only periodically. The test results can thus be summarised as follows:

Version	1)	2)	3)	Total time	Rank time	4)	Total points
Manual	-	-	min:sec	hh:min:sec	1-4	1-4	3-12
Trados	min:sec	min:sec	min:sec	hh:min:sec	1-4	1-4	3-12
MultiTrans 1	min:sec	min:sec	min:sec	hh:min:sec	1-4	1-4	3-12
MultiTrans 2	min:sec	min:sec	min:sec	hh:min:sec	1-4	1-4	3-12

Table 21: Result presentation

The total points will determine which translation was the most efficient overall and the most comfortable for the translators. However, the translators' feedback was not only collected through an overall ranking, but with the help of a questionnaire containing

qualitative questions.⁴⁹ The questions were devised with the ultimate goal of this study in mind, i.e. with the two main questions: *was this tool helpful?* and *was it easy to work with?* It was divided into a part on general information, which the participants filled in before starting the translations, and a part for each translation version, which they filled in directly after handing in the respective texts. A final part asked them to specify which translation was the most efficient and easy to them overall by distributing one to four points to the different versions. The answers to the qualitative questions will be discussed together with the ranking presented above, in order to establish a more complete image of the overall appreciation.

The translation part proceeded as follows: In a general introduction, the goal of the tests as well as the procedure was explained to the participants. General information on the participant and her previous experience with CAT-tools was collected. For the translations, the participants received the texts electronically, as a MS Word document, as well as on paper. They also received a short introduction to the tools, despite the fact that both of them were at least vaguely known to all participants. The goal of this introduction was to refresh their knowledge of the tools and to ensure that they use all the crucial functions correctly. For SDL Trados these functions included the integrated editor, since most participants only knew earlier versions of Trados working with MS Word, exact and fuzzy matches and how to insert them into the text, translation, as well as context search. For MultiTrans 1, the functions included the TextBase Agent, a tool which highlights terms and segments in the text which can be found in the database, and TextBase research, in a separate window and as automatic lookup. For MultiTrans 2, the functions included TextBase Agent and TextBase research, as in MultiTrans 1, but also the use of the Translation Agent.

Results

This section will present the results of the above test, namely the participants' answers from the questionnaire, as well as the measured time. We shall first look at the result of the time measurements and then at the participants answers to the questionnaire. After this general presentation of the results, the overall ranking will be summarised in table 24 below.

The following table shows the overall time necessary for the different steps according to the task sequence outlined in table 20 above, as well as in the task list for the scenario test.⁵⁰ The times for task 3.2, the translation, are the average times of all three participants.

⁴⁹ The entire questionnaire can be found in Annex F, p. 103 (in German).

⁵⁰ See Annex E.

Version	1)		2)	3)		Total time	Rank time
Manual	-	-	-	-	01:34:00	01:34:00	4
Trados	00:01:09	00:01:04	00:08:46	00:02:31	00:53:24	01:07:04	3
MultiTrans 1	00:01:46	00:02:05	00:06:26	00:01:32	00:53:38	01:05:27	2
MultiTrans 2	00:01:46	00:02:05	00:06:26	00:01:32	00:46:50	00:58:39	1

Table 22: Results time

Despite all the preparatory work necessary for the setup of the Translation Memories, the manual translation still clearly took more time than all the other versions. There are factors, of course, which could have had an influence on this result: As was mentioned in the questionnaire, two participants found the first text more difficult to translate than the others, especially because they first had to get used to the theme of health policy and its specific terminology. This could also explain why the translation time decreased quite steadily with the versions. Also, for versions two and three, SDL Trados and MultiTrans 1 the participants had to get used to a new system, while the difference between MultiTrans 1 and MultiTrans 2 did not imply as many new functions to learn. This could also explain the fact that the translation with MultiTrans 2 was the fastest. Unfortunately, for practical reasons it was not possible here to repeat the scenario test with different test participants, varying the order of the versions as well as the order of the texts. However, this should ideally be done in order to rule out this interference.

These results now need to be completed by the participants' answers to the questionnaire. For questions about the usefulness of a specific tool or function, the participants had to tick one of four possibilities: *not useful*, *not very useful*, *quite useful* and *very useful*.⁵¹ For questions about user-friendliness, the participants had to tick one of four possibilities: *very inconvenient* [to use], *quite inconvenient*, *quite convenient* and *very convenient*.⁵²

General information: The three participants, Dajana Abgottspon, Martina Novotny and Nina Walpen were all aged between 26 and 28 years, and all had been in contact with translation tools before to some degree, through courses at University or even at work. Dajana Abgottspon did the internship at the ZSD-D in June-July 2009, and worked there temporarily from August to December 2009. Nina Walpen did the internship in September-October 2009, and Martina Novotny in June-July 2010.

Manual translation: The translations were carried out in 01:07:17, 01:31:51 and 02:02:51 (average of 01:34:00). Besides MS Word, the following resources were used: Online

51 The terms in German are: *nicht nützlich*, *wenig nützlich*, *ziemlich nützlich* and *sehr nützlich*, see Annex F.

52 The terms in German are: *sehr unangenehm*, *ziemlich unangenehm*, *ziemlich angenehm* and *sehr angenehm*, see Annex F.

dictionaries (Leo, Wordreference, Reverso, Le Grand Robert & Collins, Pons, Wortschatz Uni Leipzig), which were considered very helpful by all; Terminology databases (Termdat), which was considered little to quite helpful; Google, which was quite to very helpful; and other Websites (admin.ch, parlament.ch, eru-lex.europa.eu, etc), which were also considered quite to very helpful. Also, one participant used the paper dictionary Duden Synonyms. The text in general was found to be quite difficult, and two participants found they needed time to get accustomed to the specific theme of the text and the medical terminology contained in it.

SDL Trados: The translations were carried out in 00:33:03, 01:01:57 and 01:05:13 (average of 00:53:24). The programme suggested between 3 and 5 exact matches, which were considered quite to very helpful. No fuzzy matches were proposed to any of the participants. The Context search was used and considered very helpful by all participants. As other resources, the same were used as during the manual translation, with the exception of the terminology databases. Overall, the tool was judged very helpful by all participants, and they declared that it was quite easy to very easy to use. Consequently, they all recommend the tool for use with this kind of text. This general appreciation was reflected in the comments, where especially the context search was noted as a helpful function, in particular for standard expressions, abbreviations and frequent terminology. The same goes for references to legal texts and earlier Parliamentary Interventions, which are quite frequent in these texts. For collocations, websites such as admin.ch and parlament.ch were more useful. However, as the Context search was generally used before undertaking a more extensive internet research, it was also helpful in finding parallel texts. Overall, the participants appreciated the fact that everything was present in one window, which reduced the number of clicks necessary to change between windows.

Negative aspects were also pointed out, the biggest of which being segmentation. Sometimes, a sentence was cut into three segments, each time after an abbreviation with a period. This was considered especially irritating because the participants felt that they had to respect the segmentation proposed by the programme, but the translation did not correspond exactly within these subsegments. It is possible to merge and split segments in the source text, but this was not part of the introduction to the tool the participants received. Another problem with respect to exact matches was pointed out: With most exact matches, the first proposition by the programme was incorrect, due to incorrect alignment. As mentioned above, the alignment in the case of the Maxi-TMs was not checked and corrected, as this would be very time-consuming. A good example would be the title *Texte déposé*, where the correct match is *Eingereichter Text*. However, the first few matches all resembled something like *Eingereichter TextKVG*, or even a totally unrelated sentence. In these cases, the

participants had to scroll down in the menu for the proposed matches in order to find the correct match.

One more point was mentioned, namely the importance of reading the entire text on paper before starting to translate. Even though they were working here on a rather short text, all participants felt that this was important for the coherence of the translation, and that working with the segments proposed by the programme made it difficult for them to keep a global vision of the text and its content. However, even if this aspect was pointed out by all participants, it was not felt very keenly, as the tool was still considered quite or very easy to work with. Overall, then, SDL Trados was felt to be most useful for the context search, while it proposed only few exact matches, and no fuzzy ones whatsoever. Also, in some cases the exact matches were not very helpful because of incorrect alignment. The context search was especially useful because it avoided unnecessary terminology searches on the internet, which is reflected by the fact that the participants did not use online terminology databases for this translation. Segmentation was mentioned as a reason to be cautious, but overall the translations were much faster than the manual ones and the tool is warmly recommended by all test participants.

MultiTrans 1: The translations were carried out in 00:46:18, 00:48:57 and 01:05:40 (average of 00:53:38). The Textbase Agent and the TextBase Research were both used by all participants and both considered not very or quite useful. As other resources, the same were used as during the manual translation, with the exception of the terminology databases. Overall, this tool was considered not very or quite useful and it was not very easy or quite easy to use. Consequently, two participants recommended this tool for use with this kind of texts, while one participant did not. This participant particularly felt that the programme did not find many of the terms and collocations she wanted to look up. Another point mentioned, even by the other participants, was a loss of time because of a frequent need to change between windows and the number of mouse-clicks which this implied.

On the other hand, the highlighting of chunks available in the database was considered helpful, even though it was not quite consistent. It seemed that some parts, such as titles, were not highlighted at all despite the fact that they most certainly are present in the database, while smaller chunks and terms were. Also, two participants pointed out that the search for terms and collocations, or even parallel texts, was much faster than on the internet, e.g. with the website admin.ch.⁵³ Finding the segments in their context was felt as an advantage, because seeing not only the sentence in which a term or expression occurred, but the entire document allowed for any ambiguity to be ruled out. Overall, then, translation here

⁵³ A search on admin.ch implies finding a term on a French page, memorising the location on this page, changing the language to German, and locating the corresponding German term on this page.

was again much faster than for the manual one, almost exactly as fast as with SDL Trados. With respect to usefulness of the database, the verdict is also the same as with SDL Trados, i.e. that it avoided terminology searches on the internet and because the segments were displayed in context. The tool is not recommended by all participants, however, particularly as it was clearly considered less easy to use because of the frequent changes between windows. The programme was thus felt to be useful for translation, but more time was necessary to get used to it, in order to use it more efficiently.

MultiTrans 2: The translations were carried out in 00:29:42, 00:45:49 and 01:05:00 (average of 00:46:50). The programme suggested about 3 exact matches, which were considered very useful by two participants and not very useful by the third. The programme also suggested fuzzy matches, the participants mentioned between 5 and 15, and they were considered quite useful by two participants, not very useful by the third. The TextBase Research was used again and considered quite useful. As other resources, the same were used as during the manual translation, with the exception of the terminology databases. Overall, the tool was judged quite useful by two participants, and not very useful by the third. Opinions also differed with respect to user-friendliness, one answer each for not easy, not very easy, and quite easy to use. Two participants still recommended the tool for use with this kind of text, while one did not.

As with the translation with MultiTrans as a corpus, the participants particularly appreciated the research for terminology and expressions which is easier than with internet. The link between MS Word and MultiTrans worked very well. Also, the fact that matches are displayed in their context was again considered positive. Two participants pointed out that an advantage in comparison with SDL Trados was the fact that the Translation Agent also proposed fuzzy matches or chunks of sentences automatically. The more problematic points, on the other hand, concerned user-friendliness rather than the usefulness of the programme. The participants mentioned that changing between different windows was not very convenient, especially because the windows of MS Word and the Translation Agent are not linked. In other words, one has to know exactly where to click in order to return to the right disposition of windows, and this obviously needs getting used to. The participants also had to get used to the navigation within the Translation Agent, for example to switch between segments. One participant made an entire segment disappear through an unintentional manipulation. Another point concerned the size of the field into which the translation has to be typed: it was considered far too small and therefore rather inconvenient, because the text is too small.

The tool was thus considered useful for research, the basic idea to have everything together, the TM, context search etc., was felt to be positive. It also proposed much more fuzzy

matches than SDL Trados, for example and the translation was thus the fastest of all versions. However, actually using the tool can be rather tiresome, because of the small window for the translation, or because of the many clicks necessary to change between windows. This is especially annoying with longer texts, and can decrease efficiency in the end. The tool was therefore recommended by only two of the three participants.

General appreciation: The participants had to attribute ranks to the different versions, which generally reflected the impressions described above:

Version	Participant 1	Participant 2	Participant 3	Total	Overall rank
Manual	4	2	4	10	4
SDL Trados	1	1	1	3	1
MultiTrans 1	3	3	3	9	3
MultiTrans 2	2	4	2	8	2

Table 23: General appreciation (ranking)

This ranking not only shows that translation with a tool is more appreciated than traditional manual translation, but also that there is a clear difference between the appreciation of SDL Trados and the two versions of MultiTrans. Interestingly, two participants agreed completely with respect to their appreciation, while participant two preferred to work without a tool to working with any version of MultiTrans. To the participants' appreciation we shall now add the first part of the testing, namely the timing of steps one to three:

Version	Rank time	Questionnaire	Total	Total rank
Manual	4	4	12	4
Trados	3	1	5	1
MultiTrans 1	2	3	8	3
MultiTrans 2	1	2	5	1

Table 24: Overall ranking

This overall ranking reflects both separate rankings in that the manual translation figures on the last rank. It also confirms the general impression that translating with a tool is perceived as more efficient than translating without. Finding SDL Trados on the first rank is not very surprising after having read the participants' comments in the questionnaires. However, thanks to its first rank in terms of time, MultiTrans 2 counts the same number of total points, which does not quite reflect the participants' impression. As we have observed above, this was the last translation done in the sequence, and therefore probably benefited from an increase of familiarity with the theme, probably even a certain weariness which might have sped up the translation. However, this should not lessen the clear result: Having a database of Parliamentary Interventions available when translating definitely is an advantage, and using

it as a TM seems even more appropriated. Also, we can say that SDL Trados and MultiTrans are equally suited for our purpose.

Conclusions and impact

This clear result, namely that working with a tool is more efficient than without, is rather surprising, given the fact that repetition in Parliamentary Interventions has been found to be very low. Some explanations can be found in the final comments in the questionnaire. Two participants pointed out that they particularly appreciated the compactness of the TM Systems, that the information is concentrated and ready to be used. This was appreciated especially for references to legal texts or earlier Parliamentary Interventions, but also for terminology and abbreviations such as the acronyms for the different departments and offices in the federal administration. Here, a gain of time was clearly felt in comparison with internet research, where efforts are often dissipated with endless clicking through different websites. It seems then that the acquisition of a TM System should not be done so much out of hope to automatically translate texts with the help of the database, but because it could facilitate and speed up the research without having to switch between windows and tabs for extensive internet searches.

One participant also mentioned that this reduction of research for terms in a sentence allowed her to concentrate on linguistic aspects of the texts. The language in Parliamentary Interventions is often very free, so that many passages are difficult to translate for linguistic rather than terminological or other reasons. This also means that the segmentation into sentences, at least when it was correctly done, did not so much disturb the participants here, as could have been expected from the studies presented in chapter 3.4 (p. 29) above. Also, seeing several translations of the same search term together was also felt to be interesting for terms and expressions which are notoriously difficult to translate. Good examples for this would be terms like *governance* or *management*, which are translated very differently according to the context, or structures such as gerunds or collocations. From the different segments, the right context and therefore the right translation can be chosen. Besides all these positive impacts of a TM System, it has to be kept in mind though that the use of such a tool would probably reduce the research for parallel text which are not necessarily present in the database. This means that on the one hand, the quality and content of the database is crucial, and on the other, that the translators have to show a certain discipline and continue to look for parallel texts when they have doubts about the translation proposed by the programme.

Concluding, then, we have to admit that this study clearly has its limitations, but that the result is nevertheless clear. Because of a small number of participants, the scenario test cannot be statistically valid, and for a more complete answer to our basic questions it would have been helpful to complement it with an analysis of the product, i.e. the translation produced by the participants. However, the scope of this study has set limits to what was feasible and this should not lessen the clear result, namely that the acquisition of a TM System could have many advantages in our specific context and can be recommended. Of course, the introduction should be framed with extensive teaching for the translators, to ensure their using it correctly and to make them aware of certain dangers when working with it. As for the specific tool, according to our tests either SDL Trados or MultiTrans are appropriate for use in the Federal Chancellery, but we must not forget that there are many others which might also be appropriate. Following this conclusion, the next section will treat other factors which need to be taken into account before deciding on whether or not an acquisition of a tool is possible.

4.3 Other factors

In answer to our basic question we can now say that yes, the above results have shown that a Translation Memory is a clear advantage, and having one available for the translators at the ZSD-D could assist them in their daily work and enhance their performance, if not in terms of speed, at least for the coherence of their texts. However, before deciding to acquire such a tool, a number of factors need to be considered. Indeed, for the test arrangements carried out above, we have left aside a few factors which will receive full attention here. This because the overall investment when introducing a tool in a service cannot be reduced to the actual price of the tool.

A first element to be considered are the translators who would actually be working with the tool. It is clear that some resistance or at least a certain scepticism has to be expected, especially when these tools are generally not very well known by the concerned, as is the case in the ZSD-D. This means that not only will it be necessary to convince the future users, but also to train them. Most developers offer initial training for services acquiring their tool, as well as continuing support,⁵⁴ but it will be necessary to have at least one so-called 'superuser', who could also assist colleagues in case of problems when the official support staff has gone home. With MultiTrans, the advantage would be that the tool, and with it

54 See e.g. Multicorpora services, <http://www.multicorpora.com/services/overview/> or Wordfast support http://wordfast.com/support_training.html (19.04.2011).

several superusers, are already present in the FC, the required training effort would therefore be smaller and expert information more accessible.

When it comes to Parliamentary Interventions, of course the person using the tool most intensively would be the intern who is present four times a year for six weeks. This means that the candidates applying for the internship would need at least basic knowledge of the tool, which is quite probable, according to translation curricula today, otherwise training would again be needed for every intern during the first days of their work. The superuser in the service would therefore not only have to instruct colleagues in case of problems, but also teach the interns if necessary. However, having the intern working with the tool, once s/he knows how to use it proficiently, can also have advantages: the colleagues revising would not need to pay attention to terminology as much, and since the intern would probably work faster, they would also have to translate less themselves during the session periods. Despite the potentially necessary training, then, the service could still benefit from the tool with respect to the intern, and for the intern this would probably be a valuable experience, given the fact that most translation services today are working with a tool.

There are technical issues, of course, which would have to be evaluated together with the IT and security staff of the Federal Chancellery. Besides the installation of the tool, regular updates will probably also be taken into account. The fact that MultiTrans is already present in two services of the ZSD shows that technically this is possible, even if installing and testing could take some time. In order to enlarge the resources for the ZSD-D, sharing databases with the ZSD-I could be considered, in case MultiTrans is acquired. A trilingual database (de/fr/it) for legal texts with the content of the Classified Compilation of Federal Legislation, for example, already exists and is updated regularly in the Italian language section. A collaboration here would have to be discussed and planned together with the Italian section and the technical staff, in order to make the databases on the servers available to everyone. It would also be possible to link the official terminology database of the Federal Administration, Termdat, to the programmes, as is already the case in the Italian and English sections working with MultiTrans. This would ensure even more control and quality with respect to terminology. These technological issues would probably take some time to be put into place, but they would not represent any definite obstacles once a decision is taken.

The most problematic issue, however, lies in the texts translated in the ZSD-D, and how they are received and stored. Indeed, if a tool were to be introduced, the workflow would have to be adapted in the first place. As it is, the texts are still received on paper or, if they are received electronically, scanned and saved as a PDF document. These texts are not usable in the tools, and would therefore have to be typed or passed through OCR software before use. In the parliamentary services, the texts are typed eventually, but normally they are first sent

to the translating services, in order to allow them to start working already. One solution would therefore be to wait for the typed documents, so that they can be directly received in MS Word format. Another solution would be to treat the PDF documents with an OCR software,⁵⁵ which is quite fast, but which would of course increase the work load of the person receiving and distributing the interventions when they come in. The best course of action would probably be to find a solution together with the Parliamentary Services in order to receive them directly in .doc or .docx format.

If the workflow is adapted in order to accommodate the tool, there still remains the question of the database which has to be established first. Filling it up along the way clearly would not make much sense because in this case the impact of the tool would probably not be felt for a long time. Here, the problem concerns the way the translations were stored until recently. A new Content Management System, GEVER, was introduced in autumn 2010, where texts and their translations are stored thematically, but this does not include texts translated before that date.⁵⁶ Rather, before the introduction of GEVER, the originals were only stored at the parliamentary services, and the translations in the different services which translated them: the ZSD-D for German translations and the ZSD-F for French ones, but the Italian ones are translated in the services of the respective Departments or Offices concerned by the intervention. For the vast majority of Parliamentary Interventions, then, the easiest source remains the database of the Parliamentary Services, Curia Vista, but drawing texts from there by copy-pasting them into MS Word is such an effort as cannot be imagined even for an intern. The best solution again seems to be a collaboration with the parliamentary services, in order to receive the complete texts from there, but even then, the texts would have to be prepared individually for the alignment. This solution would also require some effort on the part of the parliamentary services, which could generate some resistance.

According to the type of tool acquired, the alignment of the texts would require a different investment. With a tool like MultiTrans, once the texts are prepared, the alignment can be scheduled for the night and can thus be carried out without too much time and effort. Also, since the segments are not stored independently but in their context, the alignment would not necessarily have to be checked and corrected. For SDL Trados, the alignment would probably take more time, but if the alignment quality is not checked for every text, then the effort necessary would not be beyond reasonable. However, as we have seen through our tests above, just like any programme using the database model, SDL Trados relies on the quality of the alignment, so even here in order to attain the best possible quality for the database, the alignment should be checked and corrected, which implies a considerable increase in the

55 E.g. OmniPage by Nuance, <http://www.nuance.de/imaging/products/omnipage.asp> (19.04.2011).

56 See the GEVER programme, <http://www.gever.admin.ch> (23.03.2011).

effort. This factor could not be applied in the above scenario test, but it manifested itself clearly during translation, as the test participants noticed.⁵⁷ Overall, then, it is the text preparation and the alignment which represent the greatest investment in terms of time, a task for which an intern with some experience in bilingual text alignment could be recruited in order to alleviate the financial weight. There is little hope that the situation with respect to the texts will improve within the next months, despite the new content management system, GEVER. This because the system is restricted to the FC, at least for the moment, and therefore the ZSD-D would still not have access to all Italian translations, since most of them are translated outside the FC. In terms of effort and investment, the collection and preparation of texts, together with alignment, is therefore probably the most important factor to consider.

Of course, it should not be forgotten that the service translates other texts than Parliamentary Interventions as well, for which collection would be easier and which might even present more repetition than the interventions. Letters and e-mails, for example, often contain standard expressions which could be inserted automatically with the tool. The same goes for job advertisements and other standardised documents. Another good example would be the Objectives of the Federal Council, which are published every year with few changes. Here, all the unchanged parts could directly be inserted with the tool. The tool could therefore also be useful with other texts than with Parliamentary Interventions and assist the translators of the ZSD-D. However, an acquisition would probably only pay off if a database with Parliamentary Interventions can be established, and therefore the above mentioned difficulties need to be addressed first. Once a tool is purchased and introduced, and a database installed, adding these other texts would not represent too much additional effort.

Once the TM is established and the translators are working with the tool, it would be dangerous, however, to just go on without monitoring the development of the database. One or more superusers would have to be responsible for this maintenance. According to the tool acquired, maintenance can take different forms. With a tool working with the reference model, for example with MultiTrans, maintenance generally means realigning texts. For example, if the legal texts in the Classified Compilation change, the texts are simply aligned again in their most recent version, and the database is up to date. And since alignment can be scheduled for the night, the translators' work would not be interrupted and even the superusers would not spend too much time on it. For sentence-based tools like SDL Trados, maintenance is a bit more complex. Normally, when a text is translated, the translated segments are directly stored in the database. If the translation is done by an intern, chances are high that the first draft will be revised and corrected later on. One quality assurance tool

⁵⁷ The participants noted that even for exact matches, the propositions were not always correct because of segmentation or alignment errors, see scenario test results (from p.66).

would be to assign different roles to different translators. The superuser would be the administrator, while the intern and other translators would be simple users. Translated segments coming from the intern or other users would not be saved in the TM until revised, or until the administrator was able to check them. This procedure would increase the quality control of the TM, however, when texts such as legal texts are changed over time, these would have to be realigned and the wrong translations eliminated from the database. This would also imply a considerable additional effort to be made by the superuser in charge of maintenance. Besides establishing the TM and training the translators, maintenance thus has to be added to the list of elements which constitute the total investment.

One last factor is of course the financial investment. It is unfortunate that the developers did not react to my queries, but even so it is clear that good programmes will not come free of charge. Some of them are quite expensive, SDL Trados Studio 2009 Professional, for example, is advertised at 2990 €, even without server solution and floating licences. Other solutions are cheaper, Wordfast and Heartsome, for example, cost around 600 € for the most complete desktop versions, Wordbee Freelance costs 290 € per year. For more elaborate versions with floating licences, the price will be higher, of course, especially if multilingual databases have to be possible. Multicorpora unfortunately does not communicate any prices on their website. Normally, the more costly the solution, the more solid the customer support, but there are exceptions of course. Generally speaking, then, there are many different tools with as many different catalogues and prices, but a certain investment for the tool itself has to be expected, even without counting the investment in terms of working time for technical issues, installation, training, or text preparation and alignment.

The question of whether the financial investment overall will pay off cannot be answered definitely here. On the one hand, some key information is missing here, such as the price of the tools for a specific solution. On the other hand, the gain in productivity, even if easily imaginable, cannot be quantified easily. This is especially the case since translation represents only 20 percent of the overall work load of the service. Consequently, even if a positive impact is measurable in terms of translation speed and quality, the impact on the ZSD-D as a whole will probably be small, even if present. On the other side of the balance, the list of investments in terms of time and money has by now become rather long, including purchase of the tool, training, text preparation and alignment, maintenance, as well as technical issues such as installation and updates.

In conclusion it is thus possible to say that a tool would certainly be useful in the context of the ZSD-D, that it would very probably have a positive impact on the work of the translators in the service, in terms of speed and coherence. In our test, SDL Trados and MultiTrans both equally outmatched the manual translation and the general conclusion is

therefore that any tool is better than none. Between the two tools, MultiTrans not only has the advantage of being already present in the Central Language Services, which would reduce the necessary investment for training and installation. It also has the advantage that alignment and maintenance are less time-consuming than with SDL Trados. However, there are a number of other tools on the market which could be considered and not all of which are very expensive. Otherwise, what needs to be considered is that the situation concerning the texts, for the workflow as for the database, represents a serious obstacle which is difficult to overlook. Despite all the advantages of such a tool and all the positive feedback from the test participants it is therefore by no means certain that such a tool will eventually be introduced in the ZSD-D.

5 Conclusion

On arrive donc au moment du bilan basé sur une seule considération : l'utilisation d'un système informatisé rend-il le processus de traduction plus efficace?⁵⁸

In this paper we have tried to find out whether a Translation Memory System would be useful and possible to introduce in a translation service where the repetitiveness in the translated texts is not very great and in particular, whether this is the case for the ZSD-D. For this, we have explored the general field of Translation Memory technology as well as the specific context of the ZSD-D. Based on this information, we have also devised a set of test arrangements, specifically designed in order to find answers to the above questions. The test arrangement included a benchmark test, where two alignment tools were tested with respect to the texts of the ZSD-D, i.e. the parliamentary interventions, and a second test which attempted to confirm the hypothesis that, given the small amount of repetition in the texts, an extensive TM is more useful than a small one. The third part of the test arrangements, the scenario test, was set up to test how useful and user-friendly the tools would be in a situation closely resembling the context of our study, namely the ZSD-D.

The test results have indeed been very revealing. Through the benchmark test we found out that the alignment tools do not encounter any major problem when processing the parliamentary interventions, even though some alignment errors do occur. The second test, concerning the size of the TM indeed confirmed that a small TM, even if it is checked and corrected, has less impact than an extensive one. The best solution would be to have an extensive one which has been checked and where alignment errors have been corrected, a solution which for practical reasons was impossible to realise for this study. The scenario test finally showed that the tools are indeed useful: Even though the low repetition in the texts was confirmed by the small amount of exact and sometimes even fuzzy matches proposed by the systems in the scenario test, they were not necessarily the feature most appreciated and judged most useful by the participants. In fact, the context or terminology search functions

58 (King, 1993: 265)

were used more frequently and judged more useful. It seems then that Elina Lagoudaki's opinion is confirmed in that the tools nowadays are less dependent on the text retrieval through exact and fuzzy matches, but can offer other solutions as well (Lagoudaki, 2006: 16). Especially a synergy between TM technology and efficient terminology management seems to be important and useful. The conclusion, based on the results of the test arrangement, is therefore that a TM system would indeed be useful for the context of the ZSD-D and that the introduction of a tool can clearly be recommended.

On a general level, the scope of this study does not allow us to definitely answer the question whether a TM system is useful for a context where texts show little repetition and further investigation into this topic is therefore necessary. The number of participants for the scenario test here, for example, was too low for it to be statistically valid, even though the basic methodology of the scenario test could of course be used again in further pursuit of the question. It would indeed be interesting to carry out a scenario test on a larger scale, also varying the sequences of the texts and tools, in order to have more data and to draw more objective conclusions. With more participants, the quantitative data from the questionnaires could also receive more weight. Finally, it would be interesting to complete such a larger scenario test with an analysis of the translations done by the test participants. The scope of this study has not allowed for this to be included here, but an analysis of retrieved segments and terminology in the translations, for example, could have completed the data in order to form a more complete picture. In any case, independent research on the usefulness of TM systems, in particular in contexts where texts are not very repetitive, should be continued, because many companies which carry out such studies for their internal services are not willing to publish them. The question of whether such a tool is useful to a specific context, or not, is one of the most basic questions in this field and further research into it may also enable the developers to adapt the tools more specifically to the reality of translation business.

On the part of the ZSD-D, further investigation will also be necessary. Whether the acquisition of a tool will be cost-effective cannot be definitely answered by this study because other, circumstantial reasons have to be considered. For the ZSD-D, the availability of the texts and the intensive training necessary for the users are reasons which could render the necessary investment prohibitive. The main reason, however, is the fact that translation is not the main task this service has to accomplish, but represents only around one fifth of their overall work load. For clear answers to this question, further investigation inside the ZSD-D and the Federal Chancellery are thus necessary. If the service were to pursue the idea of introducing a tool, then they should start a series of tool testing in their service, for TM as well as terminology. For this study, I have limited the number of tools to only two, but there are many others, often cheaper ones. Such a test series could also give a more detailed idea on

how such a tool would be received by the translators in the service. This is especially important because, as we have seen in the scenario test, usefulness and user-friendliness are mixed up as soon as the actual users are working with the tool. Potential resistance could therefore probably be reduced by involving the translators in the selection process and tool testing because, as King (1993: 265) noted, no system can be efficient if translators refuse to use it.

Instead of a series of tool testing, the direction of the ZSD-D could also skip or shorten the tool testing stage and directly decide to introduce MultiTrans, simply because it is already present in the Federal Chancellery. The investment to introduce it would be considerably smaller, on the one hand because knowledge and installations are already in place and accessible, and on the other hand because the effort for the setup of useful databases as well as maintenance is smaller than with a tool like SDL Trados. Sharing resources in this way, would not avoid all investment costs, but would at least reduce them considerably. In both cases, however, besides budgetary discussions, early meetings with IT and security staff of the Federal Chancellery would also be beneficial for the project, in order to solve technical difficulties early and avoid slowing down the introduction process later on.

In the end, one fact should not be forgotten though, namely that the aim of such Translation Memory Systems is to facilitate the everyday work of translation professionals, to make it more efficient, as the above quote states. This can be in terms of productivity and speed as in terms of general comfort. Depending on the work context, such help is indeed necessary, in the private as in the public sector. Expectations concerning the work of translation professionals seem to be ever increasing. However, it is the context, the texts and the people involved that finally determine whether or not there is an increase in efficiency, and this should be kept in mind throughout any investigation in this field.

6 Bibliography

Austermühl, Frank. 2001. *Electronic tools for translators*. Manchester: St. Jerome.

Beurdeley, Manuelle. 2010. *Comparaison de deux outils d'alignement dans le contexte du CICR: Alignfactory et WinAlign*. Mémoire de master présenté à l'École de traduction et d'interprétation de Genève. ETI, Geneva. <http://archive-ouverte.unige.ch/downloader/vital/pdf/tmp/gqt9c9m1o7us8u08ln1monkmv4/out.pdf>, (12.05.2011).

Bowker, Lynne. 2002. *Computer-Aided Translation Technology: A Practical Introduction*. University of Ottawa Press.

Carter-Sigglow, Janet. 2004. "Who's afraid of CAT? Redefining the Boundaries of Translation." In: *Translating and the Computer 26: Proceedings of the Twenty-sixth International Conference on Translating and the Computer, 18-19 November 2004, London*. Aslib/IMI, London.

Chama, Ziad. 2009. "25 Jahre gute Memories. Die Entwicklungen des Unternehmens SDL Trados Technologies." In: *MDÜ – Fachzeitschrift für Dolmetscher und Übersetzer*, 02/2009, p.48-50.

Dillon, Sarah & Janet Frazer. 2006. "Translators and TM: An investigation of translators' perceptions of translation memory adoption." In: *Mach Translat*, 20, p.67-79.

EAGLES Expert Advisory Group on Language Engineering Standards. 1996. *Evaluation of Natural Language Processing Systems – Final Report*. Issco Institute, Geneva. <http://www.issco.unige.ch/en/research/projects/ewg96/ewg96.html>, (07.04.2011).

EAGLES Expert Advisory Group on Language Engineering Standards. 1999. *The EAGLES 7-step recipe*. <http://issco-www.unige.ch/projects/eagles/ewg99/7steps.html>, (08.09.2010).

Estrella, Paula, Andrei Popescu-Belis & Nancy Underwood. 2005. "Finding the System that Suits you Best: Towards the Normalization of MT Evaluation." In: *Translating and the Computer 27: Proceedings of the Twenty-seventh International Conference on Translating and the Computer, 24-25 November 2005, London*. Aslib/IMI, London.

Fenneteau, Hervé. 2007. *Enquête : entretien et questionnaire*. Paris: Dunod.

Freigang, Karl-Heinz. 2009. *Developments in Translation Memory*. Presentation on the occasion of the International Conference on Technologies for Translation, Forlì, 9-10 October 2009. http://tetra.sslmit.unibo.it/convegni/tetra_1/allegati/freigang.pdf, (12.05.2011).

- Geldbach, Stefanie. 2009. "Integration und Innovation. Neuerungen bei SDL Trados Studio." In: *MDÜ – Fachzeitschrift für Dolmetscher und Übersetzer*, 02/2009, p.51-53.
- Gross, Alex. 1992. "Limitations of computers as translation tools." In: Newton, John. *Computers in Translation: A practical appraisal*. London and New York: Routledge.
- Höge, Monika. 2002. *Towards a Framework for the Evaluation of Translators' Aids Systems*. University of Helsinki, Faculty of Arts, Department of Translation Studies.
<http://ethesis.helsinki.fi/julkaisut/hum/yleis/vk/hoge/towardsa.pdf>, (07.04.2011).
- Hutchins, John. 1995. "Machine Translation: A Brief History." In: Koerner, E.F.K. & R.E. Asher (eds.). *Concise history of the language sciences: from the Sumerians to the cognitivists*. Oxford: Pergamon Press. <http://www.hutchinsweb.me.uk/ConcHistoryLangSci-1995.pdf>, (12.05.2011).
- Hutchins, John. 1998. "The origins of the translator's workstation." In: *Machine Translation*, 13/4, p.287-307.
- Hutchins, John. 2009. *Compendium of Translation Software: directory of commercial machine translation systems and computer-aided translation support tools*. (15. Ausgabe, Januar 2009).
<http://www.hutchinsweb.me.uk/Compendium-15.pdf>, (08.09.2010).
- Hutchins, John & Harold Somers. 1992. *An Introduction to Machine Translation*. London: Academic Press.
- Jelinek, Richard. 2004. "Modern MT Systems and the Myth of Human Translation: Real World Status Quo." In: *Translating and the Computer 26: Proceedings of the Twenty-sixth International Conference on Translating and the Computer, 18-19 November 2004, London*. Aslib/IMI, London.
- Kästner, Astrid, Bernard Rouvière & Janine Thiele. 2007. "Standardmässig schnell, sicher, effizient. Einsatz der TM-Technologie beim Corporate Language Management von Daimler." In: *MDÜ – Fachzeitschrift für Dolmetscher und Übersetzer*, 04/2007, p.26-29.
- Kay, Martin. 1980. "The Proper Place of Men and Machines in Language Translation." CSL-80-11, Xerox Palo Alto Research Center. <http://mt-archive.info/Kay-1980.pdf>, (17.01.2010).
- Kay, Martin & Martin Röscheisen. 1993. "Text-Translation Alignment." In: *Computational linguistics*, 19:1, p.121-142.
- King, Margaret. 1993. "Sur l'évaluation des systèmes de traduction assistée par ordinateur." In: Pierrette Bouillon (ed.). *La traductique*. Montréal: AUPELF/URELF, p.261-269.
- King, Margaret. 1996. "Evaluating Natural Language Processing Systems." In: *Communications of the ACM*. 39:1, p.73-79.
- Lagoudaki, Elina. 2006. *Translation Memory systems: Enlightening users' perspective*. Translation Memory Survey 2006, Imperial College London.
<http://www3.imperial.ac.uk/portal/pls/portallive/docs/1/7307707.PDF>, (21.04.2011).
- L'homme, Marie-Claude. 2001. *Initiation à la traductique*. Montréal: Lingatech.
- Levitt, Gary. 2003. "Internet-based Sharing of Translation Memory." In: *Multilingual Computing & Technology*. 14/5, p.38-41.

Macdonald, Kirsty. 2001. "Improving Automatic Alignment for Translation Memory Creation." In: *Translating and the Computer 23: Proceedings of the Twenty-third International Conference on Translating and the Computer, 29-30 November 2001, London*. Aslib/IMI, London.

Massion, François. 2007. "Welcher Anbieter hat die besten Karten? TMS aus der Sicht eines Übersetzungsdienstleisters." In: *MDÜ – Fachzeitschrift für Dolmetscher und Übersetzer*, 04/2007, p.30-32.

Melby, Alan K. 1982. "Multi-Level Translation Aids in a Distributed System." In: *COLING 1982. Proceedings of the Ninth International Conference on Computational Linguistics, Prague, July 5-10, 1982*. Amsterdam: North Holland, p.215-220.

Melby, Alan K. 1992. "The Translator Workstation." In: Newton, John (ed.). *Computers in Translation*. London, Routledge Press.

Mikheev, Andrei. 2004. "Text Segmentation." In: Mitkov, Ruslan (ed.). 2004. *The Oxford Handbook of computational linguistics*. Oxford: OUP.

Mossop, Brian. 2006. "Has Computerization Changed Translation?" In: *Meta*, 51/4, p.787-305.

Quah, Chiew Kin. 2006. *Translation and Technology*. Basingstoke: Palgrave.

Reinke, Uwe. 2004. *Translation Memories. Systeme – Konzepte – Linguistische Optimierung*. Frankfurt am Main: Peter Lang.

Rico, Celia. 2001. "Reproducible models for CAT tools evaluation: A user-oriented perspective." In: *Translating and the Computer 23: Proceedings of the Twenty-third International Conference on Translating and the Computer, 29-30 November 2001, London*. Aslib/IMI, London.

Schnell, Bettina & Marcos Aranda. 2007. "Engeres Blickfeld, andere Reformulierungen. Translation-Memory-Systeme und der Übersetzungsprozess." In: *MDÜ – Fachzeitschrift für Dolmetscher und Übersetzer*, 04/2007, p.34-38.

Seewald-Heeg, Uta. 2007. "Vielfalt auf dem Markt. Wachsende Anforderungen – neue Entwicklungen." In: *MDÜ – Fachzeitschrift für Dolmetscher und Übersetzer*, 04/2007, p.12-25.

Somers, Harold (ed.). 2003. *Computers and translation: a translator's guide*. Amsterdam, Philadelphia: J. Benjamins.

Véronis, Jean & Philippe Langlais. 1997. "ARCADE: Evaluation de systèmes d'alignement de textes multilingues." *ACADE Project: Final report of the first campaign (1996-1997)*. <http://aune.lpl.univ-aix.fr/projects/arcade/report1-fr/index.html>, (21.04.2011).

Whyman, Edward K. & Harold Somers. 1999. "Evaluation Metrics for a Translation Memory System." In: *Software – Practice and Experience*. 29:14, p.1265-1284.

Wöllbrink, Birgit. 2006. "Multitrans 4 – Unkomplizierter Allesverwerter." In: *Technische Kommunikation*, TEKOM, 4/2006.

Zerfass, Angeilka. 2002. *Evaluating Translation Memory Systems*. <http://mt-archive.info/LREC-2002-Zerfass.pdf> (04.11.2009)

Zerfass, Angelika. 2002. „Comparing Basic Features of TM Tools.“ In: *Multilingual Computing and Technology, Language Technology Supplement*, No.51, vol.13, issue 7, p.11-14.

Zerfass, Angelika. 2005. *Translation Memory – eine Einführung*.
<http://www.transline.de/infocenter/wissenswertes/translation-memory-eine-einfuehrung.html> (12.06.2011).

Online resources

TIM/ISSCO: <http://www.issco.unige.ch>

EAGLES-Group: <http://www.issco.unige.ch/en/research/projects/eagles/index.html>

CST's Repetitiveness Checker: http://cst.dk/online/rep_check/uk

Project LISA, TMX: <http://www.lisa.org/Translation-Memory-e.34.0.html>

International Organisation for Standardization ISO: <http://www.iso.org>

Bundeskanzlei: <http://www.bk.admin.ch>

Programm GEVER Bund: <http://www.bk.admin.ch/themen/04609>

Wikipedia, Computer-assisted translation: http://en.wikipedia.org/wiki/Computer-assisted_translation

MdÜ: <http://www.bdue.de>

Multilingual: <http://www.multilingual.com>

Angelika Zerfass (Information on Translation Memories): <http://www.zaac.de>

Nuance, Omnipage: <http://www.nuance.de/imaging/products/omnipage.asp>

Commercialised tools

(This list is not exhaustive)

Across: <http://www.across.net>

Atril Déjà Vu: <http://atril.com>

Fusion: <http://www.jivefusiontech.com>

Heartsome Technologies: <http://www.heartsome.net>

MemoQ: <http://kilgray.com>

MetaTaxis: <http://www.metataxis.com>

Multitrans: <http://www.multicorpora.com>

OmegaT: <http://omegat.org>

SDL Trados: <http://www.trados.com>

Similis: <http://similis.org>

Star Transit: <http://www.star-transit.com>

Swordfish Translation Editor: <http://www.maxprograms.com/products/swordfish.html>

Wordbee: <http://www.wordbee.com>

Wordfast: <http://www.wordfast.net>

7 Annexes

Annex A – The interviews

ZSD-D, Isabel Kamber, 17.06.2010

Interne Struktur:	
Eigene Funktion innerhalb der Sektion	Stellvertretende Leiterin, Übersetzerin
Organisation der Sektion (verschiedene Abteilungen, Kompetenzbereiche)	Alle übersetzen, einige lieber, einige weniger gern Koordination: wer Zeit hat, übersetzt
Wie gross ist die Sektion? Wie viele Übersetzer?	Personen: 11, 9 Redaktoren, 1 Sekretariat, 1 LeGes (übersetzen auch); Vollzeitstellen: 8
Übersetzung: interne und externe Übersetzer	Externe sehr selten, vielleicht 1 Mal pro Jahr (Die Sektion ist sehr flexibel, übernimmt die Texte selbst, Ausnahme: sehr dringende und lange Texte)
Textvolumen, ev. Anteil der Übersetzung an Gesamtarbeit	Textvolumen: ca. 800 Seiten/Jahr 2009 (2 Listen, Vorstösse, 464 Seiten 2009, und andere, 337 Seiten 2009) Anteil ca. 20%, schwankend
Spezielles	Normseite Bund: 1800 Zeichen/Seite

Workflow und Texte	
Woher kommen die Texte, und in welchem Format?	Auftraggeber: BK, Bundespräsident/in, z.T. Finanzdepartement Format: meistens Word, manchmal Power Point oder Excel; Vorstösse noch immer auf Papier (sollte im Laufe des Projekts e-parlament ändern, Warten darauf, dass sie abgetippt sind ist keine Alternative, zu kostspielig)

Inhalt der Texte (ev. verschiedene Textsorten, Untergruppen)	Vorstösse (Anteil: ca. 58%) BK: alles Mögliche, Berichte, Gesetzesentwürfe, e-mails, Briefe, Medienmitteilungen, alles, was nicht strikt intern ist Bundespräsident/in: Reden, Mitteilungen (nur wenn es in der Funktion des/r Präsidenten/in ist) Webinhalt: ch.ch (Webseite, für die die BK zuständig ist)
Revision	Intern, für ALLE Übersetzungen, Vieraugenprinzip, Gegenlesen bzw. laut Vorlesen Praktikum: Revisionsprozess ist viel intensiver, gründliches Gegenlesen, Recherchen, Korrektur
Haben Sie den Eindruck, dass es in den Texten viel Repetition gibt?	Nein (Grund, warum bisher immer gedacht wurde, dass ein Tool nicht nötig bzw. nützlich ist) Am ehesten die Vorstösse (allerdings eher Repetition des Themas, Vokabular, nicht unbedingt ganze Sätze) Viel Repetition in Geschäftsberichten, Jahresberichten, Legislaturplanung (Wird häufig Material vom Vorjahr verwendet) Reden sicher nicht, Medienmitteilungen wenig
Arbeitsethik innerhalb der Sektion, gibt es einen grossen Zeitdruck?	Zeitdruck immer keine Weisungen, Qualität ist meist sehr wichtig (ausser Ausnahmen, z.B. grundsätzliche Aussage eines Briefes) Keine Einteilung der Übersetzungen in Bezug auf Qualität
Weitere Hilfsmittel	Termdat, Internet, Publikationen (auch eigene, Leitfaden), elektronische Wörterbücher, Bibliothek, Fragen an Muttersprachler und Autoren
Spezielles	Die Situation mit dem Praktikum ist eine etwas spezielle Situation, anderer Benutzer (weniger Erfahrung), intensiver Revisionsprozess

Persönliche Erfahrungen	
Welche Erfahrungen haben Sie bisher mit Tools gemacht, was haben Sie darüber gehört?	Keine Erfahrung, Präsentation Abgottspon/Walpen letztes Jahr, Gerüchte, Frage: braucht man das? Auch nie überlegt, dass das Tool auch für die Redaktion hilfreich sein kann
Welche Erfahrungen bzw. Vorwissen ist bei Ihren Kollegen/innen vorhanden?	Vermutlich sehr wenig Barbara Grüter hat beim Kanton Freiburg (ein bisschen) mit Transit gearbeitet, Rebekka Bratschi hat vor ca. 10 Jahren auch schon mit einem solchen Tool gearbeitet.
Haben Sie auch schon Negatives über diese Art von Übersetzungstool gehört?	Grosser Aufwand für die Einspeisung von Material und Maintenance Kompatibilitätsprobleme (jeder schwört auf sein eigenes Programm, Rivalität)
Was versprechen Sie sich von der Nutzung eines Übersetzungstools?	Erhoffen: Hilfsmittel, Inspirationsquelle beim Übersetzen, vergrösserte Einheitlichkeit der Texte, Profitieren von der Arbeit der anderen Sektionen durch Zusammenlegen der Datenbanken
Vorstellungen über eine eventuelle Nutzung, obligatorisch für alle?	Ist wahrscheinlich nur sinnvoll, wenn alle es benutzen Aber: Widerstand bei Einigen ist möglich (Kontext: GEVER-Einführung im September, es besteht momentan eine gewisse Sättigung in Bezug auf neue Tools bzw. Neues lernen)

Spezielles:	<p>GEVER: Einführung im September (2010), Ablage nicht mehr pro Sektion, sondern pro Dossier, AT und die verschiedenen Übersetzungen sind im selben Dossier abgelegt, es ist nicht mehr möglich, Dokumente zu löschen, alte Versionen bleiben neben der Neuen bestehen</p> <p>Momentan ist das Grundkonzept bekannt, aber gesehen haben sie noch nichts davon, wie es konkret benutzt wird, wissen sie erst im Sept. (Ausbildungen, Floorworkers)</p> <p>Momentan besteht kein Austausch über Tools oder Hilfsmittel zwischen den verschiedenen Sektionen, auch Texte können nicht konsultiert werden (Deutsche Sektion hat keinen Zugriff zu Texten der Italienischen oder Französischen Sektion)</p>
-------------	--

ZSD-F, Fabienne Bertagnollo, 22.06.2010

Interne Struktur:	
Eigene Funktion innerhalb der Sektion	Leiterin der Sektion, übersetzt auch
Organisation der Sektion (verschiedene Abteilungen, verschiedene Kompetenzbereiche)	Alle müssen alles machen können Praktisch alle sind ausgebildete Übersetzer (z.B. ETI, oder haben zumindest schon immer als Übersetzer gearbeitet)
Wie gross ist die Sektion? Wie viele Übersetzer?	12 Personen Vollzeitstellen: 10,5
Übersetzung: interne und externe Übersetzer	Keine externen, Sektion kann auch lange, dringende Texte ziemlich gut auffangen
Textvolumen, ev. Anteil der Übersetzung an Gesamtarbeit	Anteil ca. 30%, sehr intensiv während und kurz nach der Session, sonst weniger Hauptaufgaben: Revision und Korektion Textvolumen: keine Antwort

Workflow und Texte	
Woher kommen die Texte, und in welchem Format?	<p>Verschiedene Auftraggeber: BK, Bundesratspräsident (Alles, was von der BK produziert wird)</p> <p>Parlament: Vorstösse</p> <p>Format: Word, Excel, Pdf, Vorstösse immer noch auf Papier/Fax</p> <p>e-parlament: Momentan in Testphase oder Vor-Testphase, sollte bis Ende Jahr umgesetzt werden</p>
Inhalt der Texte (ev. verschiedene Textsorten, Untergruppen)	<p>Briefe (oft vertraulich), Botschaften und Stellungnahmen des BRs, Berichte, Legislaturprogramm, auch Webinhalt</p> <p>einige Texte können ziemlich technisch sein, da die Sektion Web der BK auch für gewisse Dinge für die ganze Bundesverwaltung zuständig ist</p>

Revision	Übersetzungen: intern, mündliches Gegenlesen, nur die übersetzende Person macht Änderungen Revision: alles elektronisch, Word
Haben Sie den Eindruck, dass es in den Texten viel Repetition gibt?	Wenig (Grund, warum sie kein Tool haben) Am ehesten noch in den Vorstössen, vor allem, wenn diese mehrere Male eingegeben werden, oder zum selben Thema, aber dann können diese schnell wieder gefunden werden.
Arbeitsethik innerhalb der Sektion, gibt es einen grossen Zeitdruck?	Wenig Zeitdruck, ist eher eine Frage, wie die Fristen verhandelt werden, oder die Länge des Texts aufs Jahr hinaus rechnet man mit ca. 5 Seiten/Tag
Weitere Hilfsmittel	Termdat, Publikationen der Sektion, SR, Internet, elektronische Wörterbücher

Persönliche Erfahrungen	
Welche Erfahrungen haben Sie bisher mit Tools gemacht, was haben Sie darüber gehört?	Ausprobieren: Trados und MultiTrans, musste aber nie damit arbeiten Trados: sehr negativer Eindruck MultiTrans: Oberfläche ist nicht schön, ist aber etwas flexibler
Welche Erfahrungen bzw. Vorwissen ist bei Ihren Kollegen/innen vorhanden?	Ein Kollege hat vorher damit gearbeitet, war sehr zufrieden, es handelte sich aber um eine andere Arbeitsethik Sonst: Praktisch niemand hat damit gearbeitet, obwohl einige im Laufe der Ausbildung eine Einführung hatten
Haben Sie auch schon Negatives über diese Art von Übersetzungstool gehört?	Sehr aufwändige und zeitraubende Einrichtung Sieht vor allem auch die Probleme in der Revisionsarbeit, Fehler werden immer wieder gemacht, das Gehirn wird "abgeschaltet" Grundhaltung: kann eine grosse Hilfe sein, Zeitersparnis, dies hängt jedoch stark von der Situation ab! Für die eigene Situation sicher nicht!
Haben Sie schon über eine Einführung nachgedacht, bzw. schon entsprechende Schritte eingeleitet?	Die Frage kam auf, als die italienische Sektion MultiTrans angeschafft hat, es wurde untersucht (nicht nur diskutiert, aber offenbar keine breit angelegte Studie), und abgelehnt Gründe: nicht hilfreich (bringt keine Zeitersparnis) und ist nicht wünschenswert für die Erhaltung des Gehirns bzw. der Intelligenz
Spezielles:	Insgesamt sehr negative Einstellung gegenüber der Tools

ZSD-I, Angela Petrone, 08.04.2010 and Simone Cardelli,

17.05.2010

Mit Kommentaren (*kursiv*) von Jean-Luc Egger, Leiter des italienischen Sprachdiensts

Interne Struktur:	Angela Petrone	Simone Cardelli
Eigene Funktion innerhalb der Sektion	Übersetzerin, de/fr>it	Übersetzer, de/fr>it

Organisation der Sektion	Keine klare Aufgabeteilung, alle machen Alles, einige etwas mehr Revision	
Wie gross ist die Sektion? Wie viele Übersetzer?	Ca. 15 in Bern, ca. 15 in Bellinzona, plus einige für die Terminologie <i>EGGER: 32 Personen, ca. 27 Vollzeitstellen</i>	
Übersetzung: interne und externe Übersetzer	<i>EGGER: Netzwerk von Externen, die angefragt werden bei Überlastung oder Zeitdruck, sehr komplexe oder normative Texte werden nur im Ausnahmefall nach aussen gegeben</i>	
Textvolumen, ev. Anteil der Übersetzung an Gesamtarbeit	Pro Person: vielleicht 4 Seiten pro Tag, aber nicht klar geregelt, man ist freier, speziell für spezifischere Texte, niemand kontrolliert. <i>EGGER: Jahresdurchschnitt der Sektion bei insgesamt ca. 19000 – 21000 Seiten/Jahr, Übersetzung macht ca. 60-70% der Arbeitszeit des Übersetzers aus, ansonsten Revision, Terminologie, Arbeitsgruppen und Meetings, Dokumentation, Ausbildung</i>	
Spezielles:	Hat vorher mit Trados gearbeitet	

Workflow und Texte	IT (Petrone)	IT (Cardelli)
Woher kommen die Texte, und in welchem Format?	Normalerweise .doc (Word Dokumente) nur wenig Pdf, Vorstösse auf Papier (sollte sich jedoch noch in den nächsten Monaten ändern).	
Inhalt der Texte (ev. verschiedene Textsorten, Untergruppen)	Es wird für die BK und fürs Parlament übersetzt: Gesetzestexte, Verordnungen, Botschaften, Berichte, Pressemitteilungen, Briefe, Vorstösse Webtexte, ch.ch,	
Revisionsprozess	<i>EGGER: Revisionsprozess intern und gegenseitig, wichtige, technische oder komplexe Texte manchmal extern durch Spezialisten des Fachgebiets. Hilfsmittel: MultiTrans</i>	
Haben Sie den Eindruck, dass es in den Textes viel Repetition gibt?	Kommt auf die Textsorte an. Viel Wiederholung bei Botschaften, oder bei Gesetzen, wenn es Entwürfe oder Änderungen gibt, aber insgesamt nicht wie in der technischen Übersetzung. Vorstösse nicht (immer neue Argumente)	
Für wen werden die Texte übersetzt? (Zielpublikum, Auftraggeber)	Keine Richtlinien für Zielpublikum, der Übersetzer muss diese Arbeit meist selber machen (Recherche, Überlegung) Bei Botschaften klar, auch bei Berichten (sind für Fachleute).	
Arbeitsethik innerhalb der Sektion, gibt es grossen Zeitdruck?	Zeitdruck ja, hängt meist von externen Faktoren ab, wie Bundesratssitzungen, Sessionen etc. Keine Kontrolle des Volumens, keine statistische Erhebungen (theoretisch möglich, wird aber nicht gemacht)	
Weitere Hilfsmittel	Termdat, Google, neu: etools.ch (Metamotor, sucht in 14 Suchmaschinen), ansonsten keine spezifischen Instrumente	

Spezielles:	<p>Terminologie: hier kein direkter Austausch, aber Verknüpfung des Programms mit Termdat ist vorhanden (funktioniert aber auch sehr gut ohne Verknüpfung).</p> <p>Gesetzestexte, z.B. Verordnung: arbeitet viel mit MultiTrans, für die Kontrolle, macht Analyse, was ist schon vorgekommen, Konsistenz der Terminologie (ist am wirksamsten, wenn die Anpassungen klein sind)</p> <p>--> Alignment für Gesetzestexte ist sehr einfach und schnell, weil die Texte sehr strukturiert sind.</p>
-------------	--

Tool	IT (Petrone)	IT (Cardelli)
Waren Sie bei der Einführung dabei?	Ja, aktiv, Tests 2006	Ja
Wenn ja, wie ist das abgelaufen?	<p>Viel Ablehnung. Initiative ging von Stephen Frost aus (Leiter der englischen Sektion), der MultiTrans für seine Sektion anschaffte (Kauf von 5 Lizenzen). So hat die it. Sektion testen können (durch 2 Übersetzer), und dann zuerst 2 Lizenzen gekauft, später mehr. Multicorpora sind gekommen, um das Tool vorzustellen, zu erklären</p> <p>Arbeiten noch heute mit der englischen Sektion zusammen, v.a. für technische Aspekte</p>	
Haben alle Übersetzer in Ihrer Sektion Zugang zum Tool?	Insgesamt 27 Lizenzen, es können also nie alle gleichzeitig damit arbeiten, aber im Normalfall genügt dies (Teilzeitarbeit).	
Wer benutzt es?	Drei Superuser, plus etwa 70% der anderen Übersetzer vielleicht 10 benutzen es regelmässig (von ca. 30)	
Gibt es Ablehnung? Aus welchen Gründen?	<p>Angst vor Streichung von Arbeitsplätzen, auch keine Lust auf Veränderungen</p> <p>Ja, Ablehnung von Leuten, die keine Erfahrungen bzw. Ausbildung mit Tools haben</p>	
Wer kümmert sich heute um das Tool? Wer macht den „Unterhalt“?	<p>3 Superuser: Angela Petrone, Simone Cardelli und Mara Ninni</p> <p>AP macht heute die Koordinierung, ansonsten nach Devise „Wenn wir Zeit haben machen wir es“ (Texte einspeisen), in Zukunft soll es etwas systematischer werden, auch Aufgabenteilung: Ninni kümmert sich um die Terminologie, Cardelli für Updates der TM (Gesetzestexte)</p> <p>Update der TM für Gesetzestexte: alle drei Monate (CD von SR), 14'000 Texte in nur 30 Minuten. Texte hochladen ist kein grosser Mehraufwand, Aufwand gibt es vielleicht eher noch für die Informatiker.</p>	
Haben Sie eine Datenbank für alle? Wie greifen Sie darauf zurück? (Server-Lösung)	<p>Server-Lösung, mehrere Memories, und Terminologie-Datenbanken</p> <p>Es gibt mehrere Datenbanken: eine für Gesetzestexte, eine für Stellen-ausschreibungen, und eine generelle Datenbank für den Rest. In Zukunft wird es nur noch zwei geben, eine für Gesetzestexte und eine generelle.</p> <p>Vorstösse sind momentan nicht in die Datenbank integriert</p> <p>mehrsprachige Memories (alle Texte in derselben Datenbank, aber das Programm zeigt nur die Sprachen, die für die Übersetzung relevant sind)</p>	

Instruktionen für die Übersetzer	Guides von MultiCorpora, kurze Anleitung auf italienisch von AP, zwei kleine Kurse durch AP in Bern und Bellinzona (MultiCorpora macht normalerweise 3 Tage Ausbildung gleich nach der Installierung, hier war es jedoch nur 1 Tag, dann ein Jahr später noch ein Tag)
----------------------------------	---

Persönliche Nutzung	IT (Petrone)	IT (Cardelli)
Wie oft brauchen Sie das Tool?	Täglich (alle Superuser) die anderen eher unregelmässig, kommt auf den bearbeiteten Text an!	Praktisch täglich
Welche Teile des Tools brauchen Sie am häufigsten (Terminologie, TM)	Alle ausser dem Translation Agent (nie) Die meisten nutzen MultiTrans einfach als Korpus, Kontextsuche (kleineres, spezifischeres Google)	Am häufigsten für Kontrolle Hat meistens beide Textbase gleichzeitig geladen Translation Agent eigentlich nicht, ist zu kompliziert, bringt zu viele Restriktionen mit sich (man muss das Segment sofort übersetzen), nur, wenn es viel interne Repetition hat oder nur kleine Anpassungen im Text zu machen sind.
Für welche Texte brauchen Sie es?	Botschaften (sind lange, werden auch auf verschiedene Übersetzer aufgeteilt) nie für Vorstösse	Für Reden lohnt es sich nicht, weil die Terminologie nicht sehr schwierig ist, für Vorstösse auch nicht, weil die zu verschieden sind
Was schätzen Sie am meisten an diesem Tool? Generelle Vorteile	Heute viel mehr Konsistenz, vor Allem bei Änderungen in Gesetzestexten, auch in der Terminologie, die Konzepte sind gleich übersetzt Zeitersparnis nur wenn es viel Repetition hat (z.B. Botschaften)	Nicht unbedingt Produktivitätssteigerung, kommt auf den Text an, aber für die Kohärenz der Terminologie, schnelleres Arbeiten wegen schnelleren Suchen, automatisches Alignment
Welches sind die Nachteile einer Arbeit mit dem Tool?	Man hat eher Tendenz, das Material, das in der Datenbank ist, ohne Hinterfragen zu nutzen, gerade unter Zeitdruck	Keine, (hat lange überlegt) Das Tool ist nur so gut wie die TextBase eben ist – QA, Kontrolle ist sehr wichtig (nur revidierte Texte hochladen), weil ja nur die „eigene Arbeit“ kompakt angezeigt wird.
Übersetzen Sie anders seit Sie mit dem Tool arbeiten?	Ich arbeite besser, vor allem mit der Kontextsuche, auch schneller	Ev. etwas schneller, ist aber schwierig messbar, ist sicher eine wichtige Hilfe, überhaupt für Kontrolle bzw. Revision
Würden Sie die Nutzung eines Tools für die ganze Bundeskanzlei empfehlen? Wenn ja, warum?	Ja, möchte dass die ganze Bundesverwaltung dasselbe Tool hat, ist aber nicht möglich (gewisse Departemente arbeiten mit STAR, oder Trados)	Ja, gerade für Kohärenz, aber: Ist auch eine Frage des Budgets.

Spezielles:	Hat lieber MultiTrans als Trados: Problem mit der Segmentierung in Trados, ist wie ein Puzzle zusammenzusetzen, nicht mehr Übersetzen. AP ist deshalb gerne auf MultiTrans umgestiegen. Auch: schnelleres Alignieren weil keine Kontrolle von Hand nötig ist.	Erfahrungen mit anderen Tools: Keine, Trados nur Hörensagen (viel Zeit für Alignment) Benutzt Analysetool regelmässig, um zu sehen, ob es sich lohnt, den Translation Agent zu benutzen, ist aber selten der Fall.
	Zusammenlegen der Datenbanken: Weiss nicht, Texte müssen gut kontrolliert werden Für deutsche Sektion nicht sehr sinnvoll, weil die nicht viel übersetzen (ausser Vorstösse, und die sind sehr verschieden). Für die französische Sektion könnte es sinnvoll sein, aber Kontrolle ist wichtig Für SR wäre ja keine neue Datenbank notwendig, für generelle Texte könnte vielleicht auch die italienische Sektion profitieren?	

ZSD-T/E, Stephen Frost, 12.05.2010

Interne Struktur:	
Eigene Funktion innerhalb der Sektion	Leiter (ich mache alles, aber vor allem übersetzen)
Organisation der Sektion (verschiedene Abteilungen und Kompetenzbereiche)	angegliedert an Sektion Terminologie (Leiter: Franco Fomasi), v.a. aus politischen Überlegungen
Wie gross ist die Sektion? Wie viele Übersetzer?	5 Personen, ca. 3,20 Vollzeitstellen, alle Übersetzen, einer macht auch englische Terminologie (50%), ein Mitarbeiter wird vom BBT (Bundesamt für Berufsbildung und Technologie) bezahlt
Übersetzung: interne und externe Übersetzer	Kann vorkommen, aber meist intern, da kein Budget für externe Übersetzung
Textvolumen, ev. Anteil der Übersetzung a Gesamtarbeit	Ca. 2300 Normseiten/Jahr, ohne Rechtstexte

Workflow und Texte	
Woher kommen die Texte, und in welchem Format?	Word, Pdf, nichts auf Papier (sollte mit der Einführung von GEVER sowieso nicht mehr vorkommen)
Inhalt der Texte (ev. verschiedene Textsorten, Untergruppen)	Sehr unterschiedlich, etwas Rechtstexte, Webinhalt, Reden, Botschaften
Haben Sie den Eindruck, dass es in den Texten viel Repetition gibt?	Nein, nur Bruchteile, Segmente, Terminologie, die Textinhalte sind zu verschieden

Für wen werden die Texte übersetzt? (Zielpublikum, Auftraggeber)	Übersetzung auf Wunsch von Ämtern, Magistratspersonen, Departementsvorsteher Alles, was innerhalb der BK erwünscht ist, ch.ch, oder admin.ch (ist aber nicht so viel)
Arbeitsethik innerhalb der Sektion, gibt es einen grossen Zeitdruck?	Je nach Text (Fristen werden oft gesetzt) Ziel ist der BK-Schnitt von ca. 4 Seiten/Tag (wird aber nicht streng kontrolliert, ist als Mass vor allem für Planung und Organisation wichtig)
Weitere Hilfsmittel (Termdat)	Wurden Bundesweit oder BK-weit von Fomasi eingeführt, OED, Duden, Larousse (zweisprachige), etwas Leo
Spezielles:	Rechtstexte erst seit 2007: Es gab eine grosse Aufräumaktion für die Gesetzestexte auf der Webseite, Organisation in der SR. In der englischen Sektion übersetzen sie Rechtstexte normalerweise nicht systematisch, sondern nur auf Wunsch der Ämter, mit guter Begründung (5-6 Kriterien) --> keine „Gratisübersetzung“ für andere Ämter! Übersetzung: ZGB – haben Rechte gekauft, update vorgenommen, fast fertig SteGB und OR sollen auch noch übersetzt werden

Tool	
Waren Sie bei der Einführung dabei?	Ja, hat sogar alles losgetreten. SF war 00-02 im EFD, wo Trados vorhanden war, hat dies jedoch nicht wirklich benutzt. Als er 02 zur englischen Sektion kam, war kein Tool vorhanden.
Wenn ja, wie ist das abgelaufen?	Die anderen Sektionen wussten von diesen Tests.: Test von Trados, Star, MultiTrans und MetaRead Warum MultiTrans: MetaRead war noch nicht so ausgereift / Keine Zeit für Alignment / wenig Repetition in Texten / Flexibilität in der Anwendung / Preis (Trados damals viel zu teuer) Ein wichtiges Kriterium war, wie schnell und einfach man starten kann, ohne noch wochenlang Texte vorbereiten und einspeisen zu müssen Zuerst 5-6 Lizenzen, dann 17, heute 27 (für englische und italienische Sektion) --> floating Lizenzen, für Teilzeitarbeitende
Haben alle Übersetzer in der Sektion Zugang zum Tool?	Ja
Wer benutzt es?	Alle müssen, aber keine Vorgaben, wie genau es benutzt werden muss
Gibt es Ablehnung? Aus welchen Gründen?	Keine Ablehnung hier, auch von älteren Mitarbeitern (dank Flexibilität von MT), einige hatten schon vorher Erfahrung mit Tools
Wer kümmert sich heute um das Tool? Wer macht den „Unterhalt“?	SF, Einspeisung der Texte etwa alle 2 Wochen, scheduling, wird über Nacht gemacht.

Haben Sie eine Datenbank für alle? Wie greifen Sie darauf zurück? (Server-Lösung)	<p>Server-Lösung (haben in den letzten Monaten auf neue Server gewechselt)</p> <p>1 Datenbank für alle Texte (findet, dass es keinen grossen Unterschied macht, ob man eine für alles hat oder 4 verschiedene, die man dann alle offen hat)</p> <p>Idee wäre, eine separate zu machen für Rechtstexte (wegen Organisation und Verwaltung, weil diese ja oft angepasst werden) und eine für die BBT-Texte</p> <p>Die eine Datenbank hat sich auch bisher bewährt, weil ihre Textinhalte so verschieden und breit sind, z.B. bei Reden, wäre schwieriger, eine angemessene Unterteilung zu finden. Das muss sich aber noch einpendeln, sie sind noch am ausprobieren</p>
Instruktionen für die Übersetzer	<p>Ausbildung ursprünglich durch MultiTrans, dann wieder durch MT beim Kauf neuer Lizenzen. Diese Ausbildung ist seit der ersten Zeit auch viel strukturierter geworden</p> <p>Bei neuem Personal: durch SF, mit relativ wenig Aufwand (MT ist relativ einfach für den Einstieg!</p>

Persönliche Nutzung	
Wie oft brauchen Sie das Tool?	Täglich, hat Konsole, kann sehen, wer online ist Für SF ist es kaum vorstellbar, als Übersetzer heute noch ohne Tool zu arbeiten.
Welche Teile des Tools brauchen Sie am häufigsten (Terminologie, TM)	Translation Agent am wenigsten, weil wenig Repetition TextBase am häufigsten, auch einfach als Referenztool Terminologie für „Raritäten“, oder besondere Termini, auch für bevorzugte Übersetzungen bestimmter Termini
Für welche Texte brauchen Sie es?	Probiert, es für alle zu benutzen, am nützlichsten ist es für Webtexte, Broschüren
Was schätzen Sie am meisten an diesem Tool? Generelle Vorteile	Flexibilität in der Anwendung, unkompliziert, einfache Verwaltung
Welches sind die Nachteile einer Arbeit mit dem Tool?	Generelle Softwareprobleme, das ist ein permanenter Verbesserungsprozess
Übersetzen Sie anders seit Sie mit dem Tool arbeiten?	Einfacher und effizienter, die Vorschläge kommen schneller
Würden Sie die Nutzung eines Tools für die ganze Bundeskanzlei empfehlen? Wenn ja, warum?	Für die Übersetzer ja, sieht aber einen klaren Unterschied zur deutschen Sektion, die eine andere Aufgabe hat. Aber auch die könnten es als Referenztool benutzen

<p>Spezielles:</p>	<p>Erfahrungen mit anderen Tools: Testphase der Einführung Trados: Die Segmentierung ist nicht grundsätzlich schlecht, wenn viel Repetition in den Texten vorhanden ist, es hätte da aber auch mit mehr Resistenz von Seiten der Mitarbeiter gerechnet! Für Rechtstexte kein Grund zum Zusammenlegen, weil die gesamte Terminologie schon in Termdat ist Verlinkung mit Termdat existiert schon: wenn Terminologiesuche, auch in Termdat! Wunsch nach mehr Austausch mit Termdat ist da – wenn die Übersetzer einen Terminus nicht in Termdat finden, und diesen übersetzen, sollten sie diese Übersetzung an die Terminologiesektion weitergeben, die dann eine richtige Fiche daraus machen.</p>
--------------------	--

Annex B – Statistics ZSD-D

Statistik Vorstösse 2009

	FS 09	SonderS1 09	SS 09	SonderS2 09	HS 09	WS 09	Total 2009
Total Anschläge Franz.	199802	43298	166694	2060	116121	234108	
Total Anschläge Ital.	19711	14908	11130		7786	30544	
Total Seiten Franz.	111	24.05	92.60	1.14	64.51	130.06	
Total Seiten Ital.	10.95	8.28	0.05		4.32	16.96	
Total Seiten 2009	121.95	32.33	92.65	1.14	68.83	147.02	463.92

ÜBRIGE ÜBERSETZUNGEN

Total Seiten 2009	337.1
davon Französische	283.0
davon Italienische	42.9
davon Englische	11.3

Total Seiten 2008	337.1
Total BK (inkl. EDÖB)	235.8
Total Bundespräsident	52.7
Total EFD	5.3
Total Parlamentsdienste	19.0
Total übrige Verwaltung	24.2
Total Französische	288.9
Total Italienische	47.8
Total Englische	0.4

Source: E-mail from I. Kamber, 21.06.2010.

Annex C – Overview tests

Characteristic	Criteria			Test procedure	Metrics
A) Compliance	1. Compliance with file format		MS Word (.doc and .docx), (pdf)	check functionalities	yes/no (boolean)
	2. Compliance with languages	2.1 Interface language	de/fr	check functionalities	yes/no (boolean)
		2.2 Working language	de/fr/it/en	check functionalities	yes/no (boolean)
		2.3 Multilingual database	de/fr/it (en)	check functionalities	yes/no (boolean)

B) Efficiency	Time behaviour, rapidity	Speed	align fr/de texts (5)	min:sec
			align fr/de texts (50)	min:sec

C) Reliability	Quality of alignment	1. No. of errors (incorrectly aligned segments)	align FR/DE text	Incorrectly aligned segments / total number of segments
		2. No. of errors (percentage of fully exploitable segments)	align FR/DE text	%

D) Maintainability	1. Correction of alignment errors	1.1 Possibility	Check functionalities	yes/no (boolean)
		1.2 Correction of alignment errors	Correct all alignment errors in 5 texts	no. of clicks necessary
		1.3 Unsatisfactory correction	Count number	number
	2. Correction of content errors	2.1 Possibility	Check functionalities	yes/no (boolean)
		2.2 Correction of content errors	Correct 1 content error	no. of clicks necessary

E) Customizability	Possibility to change segmentation rules	1. Add abbreviation	1.1 Possibility	check functionalities	yes/no (boolean)
			1.2 Adding abbreviation	Add abbreviation	no. of clicks necessary
		2. Add abbreviation list	2.1 Possibility	check functionalities	yes/no (boolean)
			2.2 Adding abbreviation list	add abbreviation list	no. of clicks necessary

F) Accessory Information	1. Portability	.tmx export	check functionalities	yes/no (boolean)
	2. Price	for 1 Licence	check on homepage /contact developers	CHF

Annex D – The texts

Overview

Content French	No. of texts	Words	Average Characters	Average Characters (w. spaces)	Average
Interventions (title)	112	33880	303	188579	1684
Initiatives (title)	21	6877	327	38544	1835
Interventions (text)	317	107366	339	601445	1903
Initiatives (text)	50	24429	489	135273	2705
Total	500	172552	364.5	963841	2031.75
Pages		575			628

No. of pages is determined according to norm page of 300 words / 1800 characters

Content German	No. of texts	Words	Average Characters	Average Characters (w. spaces)	Average
Interventions (title)	112	28125	215	189372	1691
Initiatives (title)	21	5706	271	38756	1846
Interventions (text)	317	86470	273	592256	1868
Initiatives (text)	50	19404	388	132272	2645
Total	500	139705	286.75	952656	2012.5
Pages		465			604

A complete list of texts is available from the author.

Contact: ni.w@gmx.ch

Annex E – Scenario test task list

This task list describes in detail the different steps in the scenario test:

Stage	Manual	Trados	Multitrans 1	Multitrans 2
1	1	- - open Windows NotePad - write abbreviation list (art., l'art., al.) - save file as <AbbreviationsFr.abr> - write abbreviation list (Art., Abs.) - save file as <AbbreviationsDe.abr>	- open Multitrans ListBuilder - choose languages - choose folders - check file list - close, save list	
	2	- - open SDL Trados Studio 2009, open WinAlign - create new project - choose languages - source segmentation: add abbreviation list - target segmentation: add abbreviation list - choose file type - select .tmx export	- open Multitrans 4 - open preferences (tab TextBase, abbreviations) - open file <french.abv> - add abbreviations (art., l'art., al., 0.-9.) - save file - open file <GERMAN.abv> - add abbreviations (Art., Abs., 0.-31.) - save file - close preferences	
2	-	- add files, align file names - close project settings - align project - export project as .tmx, save	- open TextBase Alignment Agent - modify list of languages - open file list - name new TextBase - check abbreviation lists - unselect creation of TermBase - align texts - save TextBase	
3	1	- - open SDL Trados Studio 2009 - create new project - choose languages - load document - load TM - open project file for translation	- open MS Word - connect Multitrans	- open MS Word - connect Multitrans - open document - open Translation Agent
	2	- translate - short introduction of the tool and functionalities - translate	- short introduction of the tool and functionalities - translate	- short introduction of the tool and functionalities - translate
4	- answer questionnaire	- answer questionnaire	- answer questionnaire	- answer questionnaire

Annex F – Scenario test questionnaire

Szenario-Test: Fragebogen

Übersetzerin: _____ Alter: _____

Erfahrung mit Informatik

SDL Trados: _____

Multitrans: _____

1. Manuelle Übersetzung

Welche Hilfsmittel haben Sie benutzt:

☐ Online Wörterbücher – falls ja: Welche? _____

Wie nützlich waren sie?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

☐ Terminologiedatenbank – falls ja: Welche? _____

Wie nützlich waren sie?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

☐ Google – falls ja: wie nützlich waren es?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

☐ Webseiten – falls ja: Welche? _____

Wie nützlich waren sie?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

☐ Weitere Hilfsmittel?- falls ja: Welche? _____

Weitere Kommentare? _____

2. Übersetzung mit SDL Trados Studio 2009

Wurden Ihnen exakte Übereinstimmungen (100%) vorgeschlagen?

☐ Nein ☐ Ja (Anzahl: ____)

Waren diese nützlich?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

Wurden Ihnen ähnliche Segmente (fuzzy match) vorgeschlagen?

☐ Nein ☐ Ja (Anzahl: ____)

Waren diese nützlich?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

Haben Sie den Context Search benutzt? ☐ Nein ☐ Ja

Falls ja, war er nützlich?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

Welche anderen Hilfsmittel haben Sie neben SDL Trados noch benutzt?

☐ Online Wörterbücher ☐ Terminologiedatenbank ☐ Google ☐ Websites

Insgesamt, war dieses Tool hilfreich bei der Übersetzung?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

War es angenehm, mit dem Tool zu arbeiten?

1 – sehr unangenehm 2 – ziemlich unangenehm 3 – ziemlich angenehm 4 – sehr angenehm

Würden Sie dieses Tool für diese Art Text empfehlen? ☐ Nein ☐ Ja

Warum / warum nicht? _____

Weitere Kommentare? _____

3. Übersetzung mit Multitrans als Corpus

Haben Sie den Textbase Agent benutzt? ☐ Nein ☐ Ja

Falls ja, war er nützlich?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

Haben Sie die TextBase für Recherchen benutzt? ☐ Nein ☐ Ja

War sie nützlich?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

Welche anderen Hilfsmittel haben Sie neben Multitrans noch benutzt?

☐ Online Wörterbücher ☐ Terminologiedatenbank ☐ Google ☐ Websites

Insgesamt, war dieses Tool hilfreich bei der Übersetzung?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

War es angenehm, mit dem Tool zu arbeiten?

1 – sehr unangenehm 2 – ziemlich unangenehm 3 – ziemlich angenehm 4 – sehr angenehm

Würden Sie dieses Tool für diese Art Text empfehlen?

☐ Nein

☐ Ja

Warum / warum nicht? _____

Weitere Kommentare? _____

4. Translation Multitrans Translation Agent

Haben Sie den Textbase Agent benutzt?

☐ Nein

☐ Ja

Falls ja, war er nützlich?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

Wurden Ihnen exakte Übereinstimmungen (100%) vorgeschlagen?

☐ Nein

☐ Ja

(Anzahl: ____)

Waren diese nützlich?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

Wurden Ihnen ähnliche Segmente (fuzzy match) vorgeschlagen?

☐ Nein

☐ Ja

(Anzahl: ____)

Waren diese nützlich?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

Haben Sie die TextBase für weitere Recherchen benutzt? ☐ Nein ☐ Ja

War sie nützlich?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

Welche anderen Hilfsmittel haben Sie neben Multitrans noch benutzt?

☐ Online Wörterbücher ☐ Terminologiedatenbank ☐ Google ☐ Websites

Insgesamt, war dieses Tool hilfreich bei der Übersetzung?

1 – nicht nützlich 2 – wenig nützlich 3 – ziemlich nützlich 4 – sehr nützlich

War es angenehm, mit dem Tool zu arbeiten?

1 – sehr unangenehm 2 – ziemlich unangenehm 3 – ziemlich angenehm 4 – sehr angenehm

Würden Sie dieses Tool für diese Art Text empfehlen? ☐ Nein ☐ Ja

Warum / warum nicht? _____

Weitere Kommentare? _____

Generelle Einschätzung

Bitte ordnen Sie die vier Übersetzungen, die Sie gerade gemacht haben, entsprechend Ihrer generellen Einschätzung (Ränge 1-4). Rang 1 entspricht dem Tool, das in der Übersetzung für Sie am nützlichsten und am angenehmsten war, d.h. das Tool, das Sie für diesen Kontext am ehesten empfehlen würden.

Manuell Rank:____

SDL Trados Rank:____

Multitrans 1 Rank:____

Multitrans 2 Rank:____

Abschliessende Kommentare?_____
