Measuring Consistency in Translation Memories: A Mixed-Methods Case Study

Joss Moorkens

B. A.

Thesis submitted for the degree of Doctor of Philosophy

School of Applied Language and Intercultural Studies

Dublin City University

May 2012

Supervisors:

Dr. Dorothy Kenny, Dr. Sharon O'Brien,

and Mr. Reinhard Schäler

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____

ID No.: 54144086_____

Date: _____

Acknowledgements

This research is supported by the Science Foundation Ireland (Grant 07/CE/I1142) as part of the Centre for Next Generation Localisation (www.cngl.ie) at Dublin City University.

This thesis would not have been possible without the generous assistance of my supervisors, Dr. Dorothy Kenny, Dr. Sharon O'Brien, and Mr. Reinhard Schäler. They have provided support, encouragement, positive criticism, and wisdom far beyond the call of duty and I am incredibly grateful and proud to be associated with them. I would also like to thank staff and colleagues at the Centre for Next Generation Localisation, the School of Applied Language and Intercultural Studies in DCU, the Centre for Translation and Textual Studies in DCU, and the Localisation Research Centre in the University of Limerick for their assistance over the past (almost) four years. I am enormously grateful to my examiners Dr. Lynne Bowker and Dr. Ryoko Sasamoto for their advice and assistance.

My sincere thanks are also due to the companies that allowed me to visit their offices and see their localisation operations (Symantec, Vistatec, Simultrans, and SDL), the companies that provided TM data, and all of the interviewees for their time and interest. I am grateful for the support and friendship of the other DCU humanities PhD students over the past few years, particularly Stephen for stats help and coffees, and Manuela, Midori, and Yanli for my initiation to graduate scholarship! I am indebted to my family and friends: my mother Joy and my sister Simone and her family for their support and encouragement, my son Evan for brightening every day, and especially my wife Giita, to whom this thesis is dedicated, for unwavering love, patience, and good humour. Finally, I would like to thank Science Foundation Ireland and CNGL for providing me with the wonderful opportunity to carry out this research.

Table of Contents

Abstract

Introduced in the early 1990s, translation memory (TM) tools have since become widely used as an aid to human translation based on commonly-held assumptions that they save time, reduce cost, and maximise consistency. The purpose of this research is twofold: it aims to develop a method for measuring consistency in TMs; and it aims to use this method to interrogate selected TMs from the localisation industry in order to find out whether the use of TM tools does, in fact, promote consistency in translation.

The research uses an explanatory, sequential mixed-methods approach. Following a pilot study, the first phase of the research involved a quantitative study of two English-to-German and two English-to-Japanese TMs. Inconsistencies found in these TMs were categorised and counted. The research found inconsistencies of letter case, spacing, and punctuation in source texts, and inconsistent terminology, formatting, and punctuation in target texts despite the restrictive nature of TM tools.

In a follow-on qualitative phase, thirteen interviews were conducted with translators and others from the localisation industry with experience of TMs. Interviewees believed inconsistency to be a problem in translations completed using TM tools and confirmed that the findings from the quantitative phase corresponded with their experiences. Furthermore, they expressed their frustration with recent developments in TM tool functionality that, they say, do not address their needs and concerns. The thesis collates interviewees' procedures for minimising inconsistency in TMs and suggests changes to the functionality of TM tools that may improve consistency and prove beneficial to translation professionals.

Publications and Presentations from this Research Project

Moorkens, J. 2013 (forthcoming). The Role of Metadata in Translation Memories. *IN:* Pelatt, V. (Ed.), *TEMPT: Text, Extra-Text, Meta-Text, and Para-Text in Translation*. Newcastle-upon-Tyne, UK: Cambridge Scholars.

Moorkens, J. 2012 (forthcoming). A Mixed-Methods Study of Consistency in Translation Memories. *Localisation Focus*, 11(1).

Moorkens, J. 2012. A Case Study of Inconsistency in Translation Memories. 4th International Association for Translation and Intercultural Studies Conference, 25th July. Queen's University, Belfast, UK.

Moorkens, J. 2011. Translation Memories guarantee consistency: Truth or Fiction? *IN: Proceedings of Translating and the Computer 33 (digital)*. London, UK.

Moorkens, J. 2011. The Role of Metadata in Translation Memories. TEMPT: Translating and Interpreting Conference. 9th September. University of Newcastle, Newcastle-upon-Tyne, UK.

Moorkens, J. 2009. A Case Study of Consistency in Computer Aided Translation. International Post-graduate Conference in Translation and Interpreting 5, 21st and 22nd November. Heriott-Watt University, Edinburgh, UK.

Moorkens, J. 2009. Total Recall? A Case Study of Consistency in Translation Memory. *IN: Proceedings of LRC XIV*. 24th and 25th September. Limerick, Ireland. p101-110.

List of Figures

List of Tables

List of Abbreviations

CAT – Computer-Aided Translation

CL – Controlled Language

CMS – Content Management System

CSB – Character String Within a Bitext

LSP – Language Service Provider

MT - Machine Translation

NP – Noun Phrase

QA – Quality Assurance

SL – Source Language

SMT – Statistical Machine Translation

ST – Source Text

TL – Target Language

TM – Translation Memory

TT – Target Text

TWB – Translator's Workbench

UI – User Interface

1 Introduction

Progressive globalisation and the growth of the Internet mean that demand for translation is higher than ever before (Bowker and Barlow 2008, p2). Language service providers (LSPs) are under pressure to provide high quantities of localised content with a fast turnaround and at low cost. This has led to increased use of technology throughout the translation industry and the development of computer-aided translation (CAT) tools such as multilingual terminology databases, multilingual concordance tools, word processors, and translation memory tools.

A translation memory (henceforth TM) is a repository of previously translated text that has been divided into segments. Each segment is usually a sentence, a heading, or a list element. Segments in the source language are aligned with those in the target language so that they can be recycled within a TM tool. A TM tool manages the translation process, providing a user interface (UI) for the translator to see both source and target texts and automatically creating a TM during translation by saving a segment of source and target text together as a translation unit (TU). In the case of reappearance of a previously translated segment the TM software will propose the previous translation to the translator. Depending on the parameters set by the translator, the TM system will also suggest partial or 'fuzzy' matches, based on a percentage of similarity between a new ST segment and a source language (SL) segment (or source-language segments) already in memory.

The core assumptions behind the use of TM tools are that they reduce the cost of translation, save time, and remove inconsistency (Webb 1998, p43; Austermühl 2001, p140; Olohan 2011, p343). They allow the opportunity to leverage legacy materials and have been shown to increase the productivity of translators, thus saving time and cost (Webb 1998, p20). Costs are further reduced as translators are often paid based on TM match analyses, with full payment offered for translation from scratch, partial payment offered for editing fuzzy matches, and a small (or sometimes no) fee paid for reviewing 100% matches. In theory TM tools should produce consistent translations as previously

translated work is recycled. This research aims to discover whether this is true in practice.

While there has been some research on the introduction of errors (García 2008) and error propagation in TMs (Bowker 2005; Ribas López 2007), there appears to be little research on consistency in TMs[1]. We propose to find a method of interrogating TMs for consistency, then to apply that method to measure and categorise inconsistencies in sample TMs in a case study. In order to maximise ecological validity, contemporary TMs from the localisation industry will be used. A disadvantage of case studies is the inherent threat to external validity. This is addressed by using a mixed-methods research design.

This research will use a sequential explanatory mixed methods design, beginning with a quantitative study of TMs in the first phase, and following with a series of qualitative interviews in the second phase. The interviews are intended to explain and add richness to the quantitative results, demonstrating whether the findings from the first phase are applicable to TMs generally in the experience of translators and other TM users from the localisation industry. The intended outcomes of this research are to find a method of measuring inconsistency in TMs, to show what types of inconsistencies (if any) commonly occur in TMs, and to suggest methods of minimising inconsistency in translation using TM tools. The specific research questions addressed in this study are outlined in section 1.1.

[1] When words or segments differ when we could expect them to be formally identical, we consider them inconsistent. The concept of inconsistency is operationalized in section 4.4.1.

1.1 Research Questions

The first question to be answered in this research is:

RQ1: *Are TMs consistent?*

One of the assumptions behind the use of TM in the localisation industry is that legacy materials are recycled consistently with minimum editing, saving time and cost, and leaving orderly and consistent TMs for future re-use.

RQ2: *How can consistency be identified and measured in TM data?*

If the assumption of consistency in TMs is to be proved or disproved, a method for measuring consistency in TMs that could be applied beyond the present study is required.

RQ3: *In the opinion of translators and those working in the localisation industry, is inconsistency a problem?*

RQ4*: In the opinion of those same people, what causes inconsistency?*

RQ5*: What procedures could be implemented to minimise inconsistency?*

A final intended outcome is a series of recommendations as to how to maximise the leverage and associated savings offered by TM. We return to these questions in more detail in Chapter 4.

1.2 Thesis Structure

In the current introductory chapter, some background to the use of TM has been provided. A more detailed discussion of the development of TM tools and a critical analysis of TM research follows in Chapter 2. Chapter 3 presents an initial attempt to measure inconsistency in a TM, detailing the methodology used in a pilot study and the lessons learned thereafter.

Chapter 4 presents the methodology chosen to measure TMs for consistency in the main study, considering threats to validity and reliability, and offering a rationale for choices made at each step. The first phase of the mixed methods study produces the quantitative results presented in Chapter 5, which reveal the extent of consistency and inconsistency in four TMs from the localisation industry.

Chapter 6 draws on the aforementioned results in a series of qualitative interviews with various TM users from the localisation industry. It seeks to find recurring themes in these interviews and offers explanations for the quantitative findings presented in Chapter 5. It also suggests ways of minimising inconsistency in TMs and improving the performance of TM tools.

The final chapter of the thesis contains a summary of the findings from this research, discusses some strengths and weaknesses, and suggests how this thesis may be added to in future research.

2 Literature Review

2.1 Introduction

Translation technology forms a comparatively minor part of the translation studies canon. This is partly because translation successfully took place for centuries before the relatively recent phenomena of computers and computer-assisted human translation. Moreover, translation theory is also most often concerned with literary translation, whereas computers are of most use in areas where texts are repetitive and functional, for example in localisation.

TM tools have, from a vague initial concept in the 1960s to humble beginnings in the late 1980s and early 1990s, risen to become an indispensable part of the localisation industry. Essentially, the job done by TM tools in recycling material from the TM has changed little since their initial release, although many tools now provide a translation environment with inbuilt terminology tools and may suggest automatic machine translations. Introductory guides introduced translators to the concept of leveraging previous translations, citing lower costs, increased productivity, increased quality, and increased consistency as benefits of the tools, often without a great deal of critical analysis.

This chapter reviews the available literature related to TM, beginning with an exploration of the origins of TM software, providing background for the current study and details of how translation with TM was envisaged prior to the availability of commercial TM tools. Section 2.2 contains a review of introductory guides to TM written for translators and students of translation technology. These guides are used to prepare translators for work with TM tools and may influence how tools are perceived by users. Section 2.3 investigates how TM is used in practice. Section 2.4 is primarily concerned with translators' attitudes to TM, while technical evaluations of TM are featured in section 2.5. Sections 2.6 and 2.7, on TM quality and the future development of the software respectively, begin a discussion that will be continued by interview participants in Chapter 6.

2.2 Historical Overview

This section presents the sequence of development of TM tools, beginning with the origin of the concept of TM with the ALPAC report in 1966 and following the progress of ideas through several research papers (section 2.2.1). Section 2.2.2 details the release of the first commercial TM tools from the early 1990s and section 2.2.3 shows how the tools have evolved subsequently and how the TM tools market has developed. It also discusses the introduction of standard file formats.

2.2.1 Conceptualisation of TM

Although commercial TM tools first appeared in the early 1990s, the idea of TM predates these tools by many years. A 34-page report from November 1966 by the Automatic Language Processing Advisory Committee (ALPAC), commissioned indirectly by the US government and titled “Languages and Machines: Computers in Translation and Linguistics”, cited the following “crucial problems” in translation: “quality, speed, and cost” (p16). ALPAC dismissed machine translation (MT) as a solution to these problems, writing “we do not have useful machine translation [and] there is no immediate or predictable prospect of useful machine translation" (*ibid*. p32-33). Rather, the paper suggested “speeding up the human translation process” and “evaluation of the relative speed and cost of various sorts of machine-aided translation” (*ibid*. p34).

One cited example of machine-aided translation as practised in the Terminological Bureau of the European Coal and Steel Community (CECA) involved matching and retrieval from a database:

> The computer goes through a search routine and prints out the sentence or sentences that most clearly match (in lexical items) the sentences in question. The translator then retrieves the desired items printed out with

> their context and in the order in which they occur in the source. (CECA 1966, p27)

The report stresses that the translation is not done by computer and adds that "since the data produced by each query are added to the database, the more the system is in use, the greater is the probability of finding sentences that have the desired term in the proper context" (1966, p27). This concept of building a database from which to draw matches is key in today's TM systems.

The idea of searching for matches in a term database was expanded upon in 1971 by Friedrich Krollmann and Erhard Lippman, but the first research paper that suggested "translation by text-retrieval" in the style of current TM tools was by Peter Arthern from the Council of European Communities at the 1978 *Translation and the Computer* conference (Hutchins 1998, p5). Arthern posits "the possibility of recording standardized chunks of text to be used repeatedly in various combinations" and later proposes a system called Terrier that can store information so that "any given portion of text in any of the languages involved can be located immediately". The method of comparison between new and stored text would be "probably sentence by sentence" (Arthern 1979, p82, p95).
In his 1980 paper, Martin Kay went further, envisaging a 'Translator's Amanuensis' with a well-designed user interface containing two windows – source text in the upper window, and target text below. Prior to translation, the Amanuensis would suggest "statistically significant words and phrases" that had appeared in previous stored translations, and could suggest dictionary entries, or previously translated portions of selected text interactively (Kay 1997, p14). As opposed to MT, Kay stresses that the "kind of translation device I am proposing will always be under the tight control of a human translator. It is there to help increase his productivity and not to supplant him" (1997, p20). Alan Melby expanded upon this in 1982, suggesting a three-level workstation, where level one provided terminological suggestions, level two included translator-controlled dictionary lookup, and level three translation aids integrated the "translator workstation with full-blown MT" (1982, p218).

2.2.2 Introduction of Commercial TM Tools

The first commercial instantiation of a translator's workstation was the ALPS (Automated Language Processing Systems) Translation Support System (TSS) that ran on the IBM AT Personal Computer in the mid 1980s, which was capable of multilingual word processing, dictionary and term lookup (Hutchins 1998, p12). Repetitions could be added to a repetitions file during translation so that they did not have to be re-translated. With increasing standardisation of hardware platforms to the IBM PC or Apple Macintosh model, more translation tools were made commercially available. Brian Harris's paper from 1988 is credited with suggesting the use of aligned segments in a bi-text and those bi-texts, along with improved methods of text alignment, meant that TM became a reality (*ibid*. p13). The first commercial TM tools were released in the early 1990s; Trados (Translation and Documentation Software) had been a service provider since 1984, and released Multiterm in 1990, followed by the DOS-based Translators Workbench II, featuring Translation Memory, in mid-1992. This first Trados TM tool supported German, English, Spanish, Italian, and French; words not contained in the source segment were highlighted in the case of fuzzy matches, and 100% matches were automatically propagated (Brace 1992, p5-6).

IBM's Translation Manager/2 (with support for 19 languages), STAR Transit, and XL8 were also released with TM functionality in 1992 (Brace 1992, p7). Déjà Vu followed in 1993 as the first TM tool that was Microsoft Windows-compatible and which worked with Microsoft Word, and then, having received a great deal of investment, Eurolang Optimiser followed in 1994 (García 2005). The 1994 release of a Windows-compatible Trados Translator's Workbench (TWB) with the addition of the T-Align tool for alignment, and a marketing push that included opening an office in Brussels, put Trados TWB in an advantageous position in the translation marketplace (García 2005). At this stage TWB supported text, WordPerfect, and Word files, but in 1997 International Translation and Publishing (ITP) in Dublin, in collaboration with Trados, developed S-Tagger, so that Framemaker files could be converted to and from a TWB-compatible format.

This was a major development, as 60% of software industry documentation was typeset in Framemaker (O'Brien and Barany 1996). The addition of Tageditor in May 1999 allowed TWB users to translate HTML and other tagged formats (Nakada 1999).

By the end of the 1990s, Déjà Vu had "emerged as Trados' major competitor in the freelance sector" (García 2005). A Windows 95-compatible version of Déjà Vu was released in 1996. It integrated functions such as terminology lookup, TM, and alignment into a single package, which differentiated it from the Trados suite of tools with a proprietary interface rather than working as a macro with Word. This dichotomy between TM tools that work as a macro for Word, and give a view of the formatted text (What You See Is What You Get or WYSIWYG) and tools that display only text within a proprietary text processing environment continued until the late 2000s, with translators tending to have a preference for one or the other (Lagoudaki 2006, p30).

2.2.3 Evolution of Commercial TM Tools

As tools diversified, interoperability between TM tools and translation vendors became an issue for LSPs. LISA (Localization Industry Standards Association), a group containing members from various companies involved in localisation and translation that was active from 1990 to 2011, formed a special interest group called OSCAR (Open Standards for Container/Content Allowing Re-use) in 1997 to define standards and improve interoperability. TMX (Translation Memory eXchange), an XML-based standard mark-up language for TM, was suggested as a tool-independent format by OSCAR in 1998 (LISA 2004), although it would be 2003 before a TM tool – SDLX – was certified by OSCAR for fully supporting the TMX format (Waßmer 2003), by which time a further exchange format that may be used for TM, XLIFF (XML Localization Interchange File Format), had been standardised by OASIS (Organization for the Advancement of Structured Information Standards, a consortium that develops web standards). OSCAR also

maintained the TBX (Term Base eXchange) format for terminological resources and SRX (Segmentation Rules eXchange) until March 2011[2]. There are still concerns about TMX format support at the time of writing, with reports of "very poor implementations from certain tool providers" (Guillemin and Trillaud 2012, p41).

In 2005 García wrote that the major TM tools were Trados, the market-leader with many major contracts (and thus many detractors, as seen in section 2.4) and, as shown by Lagoudaki (2006 and 2008), the largest market share[3], followed by Déjà Vu, SDLX, Star Transit, and (to a lesser extent) Wordfast. Since the late 1990s, Idiom Technologies' Worldserver Translation Management System (TMS), including a TM tool, had been gaining market share, by 2007 incorporating terminology management and machine translations of ST segments (Global Watchtower 2007).

More recently, due to the improvement in performance of statistical machine translation (SMT), MT has become a common feature in TM tools. MT output is suggested to translators in the absence of 100% or fuzzy matches, or in a separate pane in the user interface, with further integration of TM and MT widely predicted (Pym 2011, p1; Zetzsche 2012, p51; section 6.7). Both SDL and Star (developers of SDLX and Transit) are LSPs, so the news that SDL had purchased Trados for 60 million US dollars in June of 2005 caused some concern within other LSPs (DePalma 2005, p7). SDL's purchase of Idiom Technologies for 21.7 million US dollars in 2008, along with plans to consolidate their advantageous market position by merging the Trados TWB, SDLX, (TWB and SDLX were merged with the launch of SDL Trados Studio in 2009) and Idiom's

[2] Since 2011 the LISA OSCAR standards have been maintained by an industry specification group called Localisation Industry Standards (LIS), part of the European Telecommunications Standards Institute (Guillemin and Trillaud 2012, p38).

[3] SDL claimed to have sold 185,000 licenses for SDL Trados TM tools by October 2011, according to a presentation at the SDL Trados Roadshow at the Clarence Hotel, Dublin, on October 13th 2011.

Worldserver products, caused further trepidation among LSPs (Global Watchtower 2008). There have been some concerns that the manoeuvring for market share on the part of the software developers has not led to improvements in TM tools. These concerns are addressed in more detail in section 2.4, based on existing research, and section 6.4, based on the primary sources consulted in the current research.

2.3 Introductory Guides

Although TM tools first emerged in the early 1990s, it took several years before independent introductory guides to TM were published. These guides prepare translators for working with TM and in some cases perpetuate the axioms (savings of time, cost, and maximisation of consistency) associated with TM. In this section we review each of these introductory guides in order of date of publication. Electronic Tools for Translation *(2001) by Frank Austermühl and* Computer-Aided Translation Technology *(2002) by Lynne Bowker are intended as aids for students of translation technology and provide an introduction to CAT tools. Edited and part-written by Harold Somers,* Computers and Translation: A Translator's Guide *(2003), is, as the title suggests, intended for professional linguists and translators. Chiew Kin Quah gives an overview and discusses current trends, particularly in MT, in* Translation and Technology *(2006).*

2.3.1 Electronic Tools for Translation

Austermühl (2001) begins with an explanation as to why technology is increasingly vital for the translation industry. He believes that "the growing demand for high-quality translations of technical texts is no longer manageable without the use of computer-based methods" (2001, p8), citing increasing digitization and networked workflows amongst his reasons. Although the principles of CAT tools are clearly explained, a great deal of the book assumes a rather basic level of computer usage. Many current students of translation are relatively experienced Internet users, but this may not have been the case when this book was written. Moreover, due to technological progress, sections on modem and even ISDN use seem quite antiquated as ADSL, cable or wireless broadband is increasingly the norm. In contrast, the chapter on TM and screenshots of Translator's Workbench 2 "for Windows 95/NT" (2001, p144) look identical to the TWB user interface up to SDL Trados 2007, unintentionally

highlighting how little the TWB interface had changed in the intervening years, despite the pace of progress in the IT domain. While Austermühl does write that some texts are more suited to TM use than others and introduces the factor of initial software cost, his discussion of TM is largely uncritical and stresses only the benefits of TM. He writes that "the use of TM can result in enormous savings, both for the client and the translator", citing "increase in income", "elimination of repetitive tasks", and consistency amongst further benefits (*ibid*. 2001, p140). While the book may provide a good introduction for training purposes, subsequent works have superseded a great deal of the content.

2.3.2 Computer-Aided Translation Technology

Bowker (2002) is similarly intended for translator training – particularly with regard to CAT tools, and also begins with an explanation of the necessity of technology in translation, and of technology in the translation classroom. Bowker chooses not to "attempt to explain the computational algorithms or technical implementation strategies behind CAT tools" (2002, p5), believing the general concepts are more relevant to translators than technical detail. Her discussion of the merits and drawbacks of TM software in the fourth chapter provides a good introduction to the hot topics of rates of pay, which may be altered based on percentage matches, and quality issues within the TM. These quality issues may arise not just from poor translations, but formerly correct translations "may become inaccurate over time" (2002, p116) she writes, introducing the difficulty of TM maintenance as source materials and terminology are changed and updated, and adding that reduced consistency in translations may result (see also section 6.4). Bowker's introduction to the issues of pay rates, file formats, and TM ownership gives the reader a concise synopsis of the arguments, but, as the subtitle *A Practical Introduction* suggests, the book represents only the 'tip of the iceberg' with regard to these and other topics. The section on integration of MT and TM tools somewhat predates the level of integration that is present in the current translation market, as TM tools and LSPs increasingly use MT to

maximise leverage. However, Bowker does explain that she foresees a "continued movement away from stand-alone systems" (2002, p138) into more involved translation workflows, which proved to be the case, particularly in enterprise translation.

2.3.3 Computers and Translation: A Translator's Guide

Somers (2003) is intended to explain how computerisation has impacted and will continue to impact on the target audience of translators and language professionals, while reassuring them that, rather than being a threat, the computer "can become an essential tool which will make your job easier and more satisfying" (2003, p1). The opening chapters concentrate mostly on TM: Somers introduces the Translator's Workbench, along with some historical context, and discusses TM systems in chapter 3, ending with a brief comparison between TM and EBMT (Example-Based Machine Translation). Much of the remainder of the book is concerned with machine translation, explaining why translation is difficult for computers, describing commercial and online MT, and evaluation metrics. At the time of publishing this may not have been relevant for many translators, although even at that juncture Bennett and Gerber, in chapter 11, suggest the use of MT "as a production tool for professional translators in a commercial setting" (2003, p188).

Nyberg, Mitamura and Huijsen's introduction of controlled language (CL) in chapter 14 differentiates between human and machine-oriented CL. Although some difficulties for translators are discussed (such as repetition and uniformity), the benefits cited from controlling ST in terms of terminological accuracy and reduction in ambiguity concur with interviewees' opinions in the current research as presented in section 6.3.

Despite targeting translation professionals, Somers's book is perhaps better directed at students of translation. It gives an overview of translation technology,

yet has an emphasis on MT that may not be appropriate for many translation professionals. Somers gives an explanation of TM systems, but stops short of analysis of or comparison between individual CAT tools, which, although helpful for professional translators looking to learn practical information about TM, would also ultimately date the book and limit its use as software evolves. Nonetheless, this book provides a clear introduction to many translation technologies, from where the reader could move on to texts that focus more strongly on their chosen field of interest.

2.3.4 Translation and Technology

Quah (2006) attempts to give a technical overview of current trends in translation. Although the book is largely concerned with MT, the section on CAT tools, standards and file formats should be beneficial for the target audience of postgraduate students and professionals. There is no comparative discussion of CAT tools, and a discussion of evaluation metrics deals mostly with MT output. In his review of the book, Benis (2008) criticises the CAT section particularly, citing the omission of MT integration in the overview of TM tools, and overemphasis on TM to the exclusion of other valuable aspects of CAT programs. He writes that "many software tools in common use are either referred to only in passing or not at all", and believes that the book offers a "sketchy map of the terrain" (Benis 2008, p263). The chapter on MT is occasionally confusing and, as reviewed by Kenny (2007), contains "several instances of unfortunate exemplification, contradictions and less-than-careful wording" (2007, p295).

2.4 Translation Memory and Localisation

The introductory texts referred to in section 2.3 are practical rather than theoretical, and are intended for translator training purposes or as a reference for translation professionals. In this section we consider attempts to address this dearth of localisation theory and discuss works that take a critical view of the advent of TM tools.

Anthony Pym attempts to formulate a theory of localisation in his book *The Moving Text* (2004). Sections of the book are reworked and updated from the author's earlier book *Translation and Text Transfer* (1992). In his first chapter, Pym posits that, rather than a rewriting or version of a source text, a localised target text is part of the process of distribution of the source text. When efforts to widen distribution to new markets cease, distribution will "diminish, [...] and the text will eventually be without function" (2004, p6). Target texts are created not for nations, but for locales, defined by their need for localised materials and their resistance to non-localised texts.

Part of the process of localisation for more than one language is 'internationalisation', which Pym describes as a production process of removing culturally embedded material prior to and in order to assist translation into the target language or languages, although the process needs also to take account of language, input, and writing conventions that may vary between languages (see section 6.5). Pym appears to see the process of internationalisation positively, since it allows for translation into more and less popular languages, which should "put paid to ideologies of English-the-killer-language working hand-in-hand with technology" (2004, p40). Pym ignores the other side of this process, namely that localisation allows companies greater distribution, increasing the ubiquity of their product. As a result, a decision not to localise a popular product can have far-reaching effects. For example, if Microsoft decides that a localised version of Windows is not viable for a language, they have the power to change the

language of computing within this linguistic locale (Hilmarsson-Dunn 2001, p306). Since few companies localise into lesser-used languages, speakers of these languages have little choice other than to use products in languages other than their own.

As part of his discussion of the cost of localisation, Pym writes that "the transaction costs should be less than the projected benefit" (2004, p141). He points out that language learning is of better value than translation in the longer term, since translation costs rarely decrease. He believes that translation memory software, while helping to maintain terminological consistency, involves too much management to bring about great productivity gains. Elsewhere, Pym has opined that translation memories "lead to a progressive degradation of translation quality" (2008, p43). He argues that segmentation, as occurs within TM software, results in non-contextual translation, despite the translator's access to context within most TM tools. This concurs with Dragsted (2005), who found that professional translators' processing was complicated by the imposition of TM segmentation, and that segmentation tends to make translators review each segment separately rather than in context. Others have noted the effects of segmentation, such as Mogensen (2000, p28) who writes that target text segments tend to be written in a minimal style to maximise recyclability, and Heyn who says that this minimal and technical style, avoiding pronouns, leads to "peep-hole translations" (1998, p135). Pym, however, also feels that translators should make translation decisions based on a set of ethical values, only translating if their work will improve "cross-cultural co-operation" (2004, p177).

In his view, this tendency towards fragmentation in the localisation industry results in work that is "not just boring but also dehumanizing" (2004, p198). In an article in the Chinese Translator's Journal, he addresses these concerns specifically, stating that translators are "virtually obliged to accept the renditions that come from the client (the glossaries and translation memories); they will not challenge deceptive "exact matches"; they are not motivated to improve the

quality of the memories" (2008, p44). Pym's position misrepresents the translator's role somewhat, as in many cases translators are obliged to accept exact matches, and, rather than being deceived by these matches, are not paid or permitted to change them (see section 6.5.1 on client expectations of 100% matches). This is also often the case for 100% matches in context (In Context Exact or ICE match in Idiom Worldserver; PerfectMatch in SDL Trados TWB) which may be locked so that the translator cannot edit them (see section 6.4). Bowker also notes that, when presented with a TM match proposal, it may be "difficult for the translator to think of a different way of expressing that notion" (2007, p182). Nonetheless, this does not change Pym's negative perception of TMs as a vehicle for error propagation, a view that recurs in other literature (Bowker 2005; Ribas López 2007), but is mostly missing from more general works on computers and translation due to the axiomatic view of consistency in TM, which we intend to challenge in this thesis.

2.5 TM Tools in Practice

This section presents research into use of CAT tools by translators. Section 2.5.1 reviews quantitative and qualitative research on users' attitudes to TM tools. Opinions on tool development are discussed in section 2.5.2, drawing on surveys and studies of views expressed on discussion boards. Finally, two studies on translators' use of CAT tools other than TM tools are reviewed in section 2.5.3.

2.5.1 Translators' Attitudes to TM Tools

Several works have explored the perception of translation technology within the translation industry, most often by surveying freelance translators, but also by analysing user comments on mailing lists and online fora. Dillon and Fraser (2006) investigated translators' perceptions of TM in an attempt to understand why all freelance translators had not adopted TM despite pressure from the industry. The authors' hypotheses were that inexperienced translators have a positive perception of TM; that translators who have adopted TM feel positively about it; and that how translators value TM is unrelated to their IT proficiency.

The authors received 59 usable responses, mostly from freelance translators, to an email questionnaire that examined ease of use, compatibility, 'trial-ability' and several other factors related to TMs. In their analysis of the responses received, the authors write that newly qualified translators have a positive perception of TM users, possibly due to a respect for those more established in the industry. Translators with little experience felt that TM use was positive for a translator's profile, whereas those with over 10 years' experience felt the opposite. The authors believe that the low uptake (at the time of publishing) from recently qualified translators can be attributed to a lack of TM training, despite the fact that contemporary translation courses are likely to have included a translation technology component (Bowker and Barlow 2008, p8). The

second hypothesis - that TM users feel positive about TM - was proved largely true for this study. IT proficiency was shown to have a large importance for TM adoption, possibly because "those with strong IT skills view TM as less of a threat to their craft" (*ibid*. p76). While the article added to the small amount of quantitative research available at the time, it is difficult to make broader generalisations about the industry based on 59 respondents. However, Lagoudaki's survey from the same year proved a better barometer on the state of the industry with regard to TM.

Lagoudaki's survey (2006) sought to present opinions of and suggestions for improvements to TM from a large sample of translators and project managers. She begins her 2006 paper with an analysis of the TM systems on the market at the time of writing, differentiating between storage techniques, matching algorithms, and whether or not the program works as a macro with Microsoft Word. She then explains how the survey was designed, piloted and carried out (via the Internet), before analysing the data created by 874 translation industry respondents, 699 of whom completed the survey in full. 90% of respondents were translators, 73% worked on a freelance basis, and the majority were university-educated with 'good' self-assessed competence using computers.

Unsurprisingly for a web-based survey, the percentage of respondents who used TM was high (82%) and found to be highest for those who translate technical or repetitive texts, and those with IT skills. Despite SDL Trados TWB being the most widely used TM tool and the one that most non-users had heard of, Déjà Vu and Wordfast were rated higher by users. Most of those surveyed, possibly because many used TWB (which at the time worked with Word), stated a preference for a Word-macro TM system that would "show the text you are working on as it will finally be seen by the end user" (Benis 2003), rather than the contextless single string view in a tool such as Idiom Worldserver. All respondents had Internet access (as might be assumed for an online survey) and most had a DSL connection.

In the aftermath of this survey there was a small amount of argument online as to where notice of the survey had been posted and whether all translators' groups had been alerted (ATA Language Technology 2007), but the survey details were posted on many Internet translation boards and portals, and the number of respondents is impressive. Despite the average user ratings, the level of TWB use reported shows the benefit of marketing and market ubiquity. The responses from non-users of TM show barriers, such as the learning curve, that TM system developers need to be aware of when analysing their software's ease of use, although Bass (2006) suggests that "most who adopt CAT tools can be back up to baseline productivity within five days of starting to use the software" (p76). García (2008) also writes that TM is "an obtrusive tool that translators need to learn and adjust to before they can achieve the gains in productivity and consistency it promises" and believes there should be more focus on what he calls the "specifity of translating with TM" in work on the topic (García 2008, p55).

García (2006) conducted a qualitative review of TM usefulness based on comments by contributors to a popular web-based translation mailing list. Using a corpus gleaned from contributions by 134 users, García writes of benefits of TM adoption – productivity increases, especially for repetitive texts – and drawbacks – frustratingly high learning curves and productivity increases that are offset by discounts based on the match percentage. Correspondents complained of different behaviour by TM systems depending on the source file or computing environment, a problem exacerbated in the case of TWB by reports of poor customer service.

Once TMs were successfully deployed, comments were generally very positive about productivity increases, but translators were frustrated by clients' requests for a "Trados discount" (fuzzy match discount) with "very little grounds for negotiation" (*ibid*. p102). Even without repetition in the texts, users found TM beneficial for terminological consistency, although others found the distraction of suggested translations irksome. This mixed bag of favourable and

unfavourable opinions is consistent with previous quantitative research, leaving no fixed conclusion for the paper other than one translator's description of TM as "another case where the latest and greatest technology turned out to be different than imagined" (*ibid*. p104).

García writes about translators' tendency to be undervalued and how "translation was never likely to be the master of its own destiny" (2007, p56) as reasons why TM is, in many cases, moving the focus away from the translator. He presents a scenario where translators are expected to be skilled, efficient, have high computer literacy and problem solving skills, yet then are increasingly straitjacketed by client demands and restrictive TM environments. García recounts that, while initially TM put translators in an advantageous position, the power has shifted to (especially larger) clients as localisation has given way to internationalisation. Translators also increasingly see TM as vital for their work. García explains how enterprise translation companies are retaining their TMs by pre-translating prior to sending the work to freelancers, rather than letting the translators work interactively. These companies can thus circumvent arguments about the ownership of TMs, or what Smith, in an article on the topic of ownership, calls the "translation family jewels" (2008, p23).

Elsewhere, Smith argues for the translator as owner of a TM based on EU laws on database ownership and the lack of originality inherent in most source texts intended for localisation; however he notes that the common practice is for "corporate LSPs to assign the TM and the rights to it directly to the customer, as added value in the translation service provided" (Smith 2009, p5). Gow does not believe that assignation of TM copyright would necessarily be helpful, but rather suggests that translation clients should "learn to profit from their potentially valuable linguistic assets" and that translators should be allowed to reuse TMs internally as long as non-disclosure obligations are met (2007, p188).

García discusses the appearance of the first web-based TM applications, with the TM itself retained at the server side, and the implications of this for freelance

translators. He believes that, while initially this may bring a more level playing field, where less technologically adept translators have equal opportunities for work, this migration to web-based tools may ultimately lead to a widespread de-skilling of the translator's job, leaving them "anonymous and easily interchangeable" (García 2007, p66).

A more recent study by McBride (2009) looked at translators' attitudes to TM based on 937 posts to online fora at TranslatorsCafe.com and ProZ.com. Users on these fora may post under assumed names and provision of personal information is not necessary, thus input may be candid and opinionated. 219 posts contained little information and were removed from the study, leaving over 700 posts of which many (30.8%) related to purchasing a TM system. Thereafter, most popular topics were source and target formats (23.4%), client brand and format requirements (19.2%), files created by the TM system (13.4%), and the price of TM tools (12.1%).

Contributors' comments on file formats and compatibility "generally arose from client, agency, or outsourcer requirements" that files would be delivered in a specified format, usually an SDL Trados TWB or other proprietary format (McBride 2009, p122). McBride says that SDL Trados TWB is the industry standard and notes that "a large number of translators are not necessarily content with this situation" (2009, p135). This dominant position combined with what the contributors perceive as a high pricing policy leads to TWB becoming the focus of opprobrium throughout the corpus. One contributor opined that SDL Trados "hopes above all to sell to giant corporations, who will put pressure on translation agencies to buy, who will likewise pressure translators to buy" (*ibid*. p125). Another wrote more generally: "a lot of purely commercial software is designed so that you are locked into buying upgrades and so on, which seems maddeningly exploitative" (*ibid*. p115). McBride believes that users need to understand prior to purchase that TM systems are not a once-off investment; "they must be updated like any other software application, piece of technology, or other translation resources" (2009, p151).

TWB was variously described as being “very complicated”, “a big pain in the neck”, “difficult to learn and very frustrating to use”, and “impossible to master without a guide [manual]” (*ibid*. p116). Contributors complained about time lost while learning to use TM tools, and felt the benefit of TM was mitigated by fuzzy match discounts. One wrote “There is obviously downward pressure on rates due to CAT tools. It's ironic that we're expected to shell out $1,000 or more, [and] then offer discounts to boot” (*ibid*. p136).

Based on the study, McBride makes some recommendations that she feels will lessen the schism between some developers and translators. She notes that TM tools’ promotional material emphasises technical aspects of the software rather than benefits. When benefits are mentioned, one is increased income; however no mention is made of payment practices or fuzzy match discounts. She feels that benefits, learnability, support, and ease of use should be more explicit in TM tool developers’ promotional material (2009, p144). She also notes that many TM tools are moving to a user interface with a proprietary text editor rather than incorporating a third-party word processor, and suggests that developers should maintain contact with their customers to “determine attitudes towards the current state of integration and realize that users who are content with their current integrated systems may not easily migrate to a proprietary text editor” (*ibid*. p148).

Finally, she argues for further interoperability between tools and between vendors, with an emphasis on the importance of open standards, to “combat vendor lock-in and compatibility issues” (*ibid*. p154). However, her argument that this would “increase customer satisfaction and would facilitate the process of switching from one TM system to another” (*ibid*. p154) may give pause to those who benefit from existing vendor lock-in.

Many of the findings from studies such as Lagoudaki (2006) and McBride (2009) recurred in a survey from 2012 focussing on translator status. Marshman found that 79% of 229 translators used TM for their work, although 9% (22

respondents) described themselves as “former users of TM tools”. 79% of respondents said that TM made them feel more in control of translation quality as it “facilitates consistency” and gives “fast, easy access to a large amount of data” (Marshman 2012). However, 54% of respondents said that they feel they have less control of remuneration due to TM, primarily due to the associated discounts expected by clients.

2.5.2 TM Tool Development

Lagoudaki believes that users’ needs are not always catered for in TM tool development as, in TM system beta tests, translators are only “invited to provide feedback on an almost finished product with limited possibilities for changes” (2006, p1). She has also written that “systems usability and end-users’ demands seem to have been of only subordinate interest” in TM system development (Lagoudaki 2008, p17), a view reiterated by interviewees in this study (see section 6.4.2).

In Lagoudaki’s survey, respondents’ views on the TM systems that they currently use are of particular interest, as are comments regarding features that users would like to see implemented in future (see also section 6.6). Given the stated willingness of 69% of respondents to get “involved in the development process of a tool that will meet their needs in a better way” (2006 p32), there may in the future be less disparity between user requirements and software functionality. Lagoudaki’s survey also highlights a possible flaw in the software development cycle, when consumers’ views are not always taken into account.

This topic recurs with more focus in Lagoudaki (2008), where she writes that TM research is mostly motivated by “technical improvement of the TM system and not how the TM system can best meet the needs of its users” (p17). She believes that, because system development is “driven by profit potential”, development has served the needs of language service providers, whereas translation

professionals are in a position of relative weakness, and thus have "been forced to use the technology that was designed to serve industry goals and not their needs" (Lagoudaki 2008, p58). As an example of this, she cites the Requirements Engineering cycle for a commercial TM system, where the focus is on previous research rather than the end user.

Based on her research, the message from the users of TM is occasionally unclear or conflicting. However, she interprets that an overall message is clear: users want simplicity. This does not necessarily mean fewer features; rather they want a streamlined process with compatibility between languages and scripts. They want ease of access, meaning "affordability of the system, not only in terms of purchase cost, but also in terms of upgrade, support and training costs" (Lagoudaki 2008, p203). Based on her respondents' answers, she foresees the future of TM involving increased interaction with web-based corpora and other online resources such as dictionaries and glossaries.

Lagoudaki's example of user consultation coming towards the end of a product development cycle contrasts with the modus operandi of User-Centred Design (UCD) as advocated by Saffer (2007). UCD practitioners view users as co-creators and involve users at every stage of a project (Saffer 2007, p32). Saffer writes that, in the past, software engineers have "churned out software that made sense in terms of the way computers work, but not in terms of the way that people work" (2007, p31).

It could be argued that TM systems should use activity-centred design, in other words, they should be tools for a specific activity as opposed to for specific users. Although some amount of complexity is inherent to every process, and there is a point beyond which you cannot simplify a process any further (*ibid*. p54), many TM systems would benefit from increased 'affordance' of commands. In other words, results of commands or button presses should be implicit and intuitive, in keeping with perceptions of the user. Furthermore, interaction designers believe that errors should be constrained in design, forcing users to correct erroneous

actions. Saffer also points out that “human beings have an amazing tendency to become accustomed to the terrible, inconvenient, and awkward,” until this is changed by “something we may not have even known we needed” (2007, p21).

Olohan (2011) refers to this process of adaptation to a TM tool using a model borrowed from the field of sociology called the ‘dance of agency’. She analyses (via posts on an online forum) a translator’s narrative of his struggle with the installation of SDL Trados Studio 2009, to which he has ascribed human attributes. The software does not behave as expected or provide feedback for the user’s actions, and instead he describes it “staring you blankly in the face” (*ibid*. p348). As other forum members become involved in the discussion, the difficulty for the software’s developer becomes clear. A more technologically adept user, who was one of “thousands of testers” of the software prior to release, believes that the problems encountered by the original poster (OP) were of his own creation (*ibid*. p350). The OP replies that the developers “thoroughly lack insight into the mind of the regular user” (*ibid*. p351). Some respondents state their desire for a simplified system, others for more options; some respondents blame the developers for problems encountered, while others blame the translators involved in the testing programme. The challenge for software developers is to accommodate users with different levels of technical ability who each want different things from their TM tool. In this case, some users express satisfaction with their experience of the TM tool and the OP eventually decides to stop using the product, as “some people are not compatible with some programs” (*ibid*. p350).

2.5.3 CAT Tools Used by Translators

Fulford and Zafra (2005) report on freelance translators’ usage of online tools and web resources based on the results of a survey carried out in 2003/2004. They received responses from 390 freelance translators, with experience ranging from 1 to 51 years, 97% of whom had Internet access. At that time, two thirds

accessed the Internet via dialup – a situation that is likely to have changed in the meantime. Most respondents used email, search engines and online glossaries. 51% used newspaper archives, 30% online journals, 6% used Usenet groups, and 3% online MT. The respondents could be put into three groups: those who used the Internet as a communication tool, those who used it also for some glossary and online search, and those who made wide use of the Internet, including for collaborative work, marketing and terminology retrieval. The third group was more likely to be younger and hold a university qualification.

A more recent study by Désilets et al. (2009), for which all subjects had university training in translation and held full-time positions as translators, found that translators use an average of ten tools to solve "translation problems". The research group used Contextual Inquiry, a technique based on observing subjects at work and asking questions to gain further insights, to understand work practices of eight translators. For translation problems caused by, for example, unprescribed terminology, highly polysemic vocabulary, or colloquialisms, the translators used various public and private terminology resources, dictionaries and lexicons, TMs, previous translations, and style guides.

They used more public (78%) than private material, more bilingual than unilingual resources, and in 35% of cases, searched for further solutions despite having already found one. The research group found "no universal tools", and that "new tools do not necessarily replace older ones" (Désilets et al. 2009, p5). Although their flow was interrupted by the "real inconvenience" of navigating between tools and there were often further delays when "the first resource consulted… turned out to be a bad choice" (*ibid*. p4), translators appeared unperturbed by sifting through noisy data, and "seemed very adept at scanning a list of potential solutions" (*ibid*. p7). The translator's ability to sift through noisy data has also been noted by Lauriston, who writes that translators "are usually able to separate the wheat from the chaff and even turn the chaff into palatable solutions to a particular communications problem" (1997, p180).

2.6 Technical Evaluation of TM Tools

The studies surveyed so far have dealt with translator attitudes to TM, and with the impact of technological developments on users of TM tools. Another strand of research takes a more technical approach to evaluating the retrieval performance of TM tools, on the one hand, and the human effort involved in editing matches from memory, on the other. Studies belonging to this latter strand are discussed in this section, beginning with TM match retrieval in section 2.6.1. Thereafter, section 2.6.2 contains details of studies comparing the effort required to edit TM and MT target text proposals and some related work on the integration of TM and MT.

2.6.1 TM Match Retrieval

Whyman and Somers (1999) attempt to evaluate the retrieval of 'useful' fuzzy matches in a TM system and create an evaluation metric for usefulness of TM matches. Matches that are considered to be relevant to the translator are labelled a hit, irrelevant matches a miss. The chosen methodology is based on measurements of precision and recall in the field of Information Retrieval; the performance of a TM system is measured based on the fuzzy match parameters as set by the user, and the number of keystrokes required to produce the desired translation. This methodology was tested on a French-English corpus of which eight pages had already been translated. In this case study 43 matches were found, of which 33 were considered 'hits'. However, a problem with this methodology is the subjective decision as to whether a match is 'useful' or 'relevant' or not. However, the authors state that this is "an initial attempt at formalising some quantitative evaluation techniques for TM" and hope that many more will follow (Whyman and Somers 1999, p1283).

Gow's MA thesis in 2003 also investigated metrics for evaluating TM systems, dividing programs into those that store "sentence-based" segments of text, and those that search for a "character-string-within-a-bitext" (CSB) (2003, p iv). She chose SDL Trados TWB and Multitrans as representatives of the two approaches. The work begins with a review of TM literature, during which Gow notes that "TM evaluation has so far been less concerned with determining the single best tool and more concerned with helping translators and project managers evaluate the best tool for their own needs" (*ibid*. p12). She also notes that improved consistency in translation "can reduce the chances of litigation," (*ibid*. p14) an insight perhaps from the author's legal studies and work for the Canadian government (*ibid*. p4). She does not, however, measure whether TM does actually improve consistency.

Gow notes some drawbacks of TM, such as the steep learning curve, difficulties of database management (and thus error propagation), and pay issues, in addition to the problem of disputes over TM ownership. She quotes Kay's intention for TM, that having uninteresting work done automatically would make the work "more rewarding, more exciting, more human," (Kay 1997, p3) without pointing out that this has not necessarily transpired, although her discussion of problems of file format, style and revision tacitly imply that TMs do not offer the translator as smooth a ride as was initially envisaged.

Gow makes these problems more explicit when discussing each TM approach. Of the sentence-based approach, she lists problems of TM quality for translators who do not revise as they work, stylistic inconsistency (Bédard's famous "sentence salad" (2000)), and a prevalence of unhelpful fuzzy match suggestions or 'noise' (*ibid*. p28). Her discussion of the CSB approach highlights the advantage of providing suggested matches in context, potentially providing guidance in new domains. She also stresses that because sentential alignment is not required in the CSB approach, faulty alignments do not result in the insertion of incorrect segments into a translated text. However, this approach has disadvantages in areas of networked or shared TMs and may cause increased

noise. She also notes that translations are not added to the database as the translator works, mitigating the benefits of TM for repetitive texts.

For her pilot study, Gow created a scoring system to measure the 'usefulness' of a TM, based on use of TM suggestions and time wasted by noise, and refined this methodology for application on a larger test corpus. She found that TWB received fewer time penalties than Multitrans, although it required more database management. Multitrans created noise through reading left-to-right without segmentation, but in other cases suggested matches that TWB missed due to rigid segmentation. In cases of exact matches, TWB's automatic insertion did save time. Gow mentions limitations of her methodology in conclusion, but asserts that her thesis that the two methodologies are empirically comparable has been proved. She mentions that her method may be applied differently by different translators and adds that a TM retrieval function is only one component of a suite that is usually capable of more functions. Evaluation of the 'usefulness' of memory-based systems was also suggested by Hodász (2006), but like Gow, Hodász found that usefulness is difficult to quantify.

The limitations of TM retrieval are addressed by Reinke (2004), who provides results of his own empirical research. He believes that there has been insufficient research into TM tools and creation of TM tools, and writes:

> bislang nur wenige der am Markt verfügbaren Übersetzungswerkzeuge aus nationalen oder internationalen Forschungsprojekten hervorgegangen sind (Reinke 2004, p9).
>
> *So far, very few translation tools on the market have emerged from national or international research projects*

His discussion of methods of data retrieval leads into an analysis of similarity, and of the difference between human and computational notions of similarity between texts. Fuzzy or even exact matches, may prove deceptive, as:

> identische Satzinhalte [können] identische oder unterschiedliche Ausdrucksformen aufweisen… ebenso können nicht-identische Satzinhalte auf der Ausdrucksseite identisch sein oder nicht (*ibid*. p235).

Identical content can be expressed in identical or different forms… similarly, non-identical contents can be expressed using identical forms

Reinke identifies this inability to cope with paraphrase and ambiguity as a limitation in current TMs (see section 6.3) and casts doubt on the accuracy of calculating match percentage based on the similarity or otherwise of source text segments, as this does not necessarily correlate with the amount of editing required on a target text proposal (see section 6.5.6). He also discusses current methods of segmentation and retrieval, which he believes can be improved upon through integration of improved algorithms and linguistic analyses. Reinke concludes with a recommendation for further interdisciplinary research into TMs, and his book provides a reasoned and well researched introduction to the technology that powers current TM tools from the perspective of the technologist rather than the user.

2.6.2 Comparison and Integration of TM and MT

O'Brien's study in 2006 used eye-tracking to measure the relationship between the translator's cognitive load and fuzzy match value, forging an avenue of empirical investigation beyond think-aloud protocols (TAP, when a subject is asked to vocalise his or her thought processes) and keystroke logging. While keystroke logging collects data as the translator discharges material from working memory during the translation, eye-tracking is said to give a "detailed picture of the complex processing involved" during the process (Carl 2009, p1). "Lest it be presented as a panacea," O'Brien cites data loss, expense, and the requirement of controlled lighting as drawbacks of eye-tracking technology (2006, p186).

Her measurements per fuzzy match percentage in this study are particularly relevant to the discussion of how much a translator should discount depending on the match percentage. In this case, given that O'Brien's concern was to establish the effectiveness of eye-tracking as a research methodology and due to

the novelty of the evaluation format, only four professional translators participated in the research. Negative correlations between match percentage and cognitive load were shown in the tests, with the least load caused by exact matches and the highest load caused by non-matches. As the percentage match decreased, the cognitive load was seen to increase, other than in anomalous figures for 74-80% matches.

Cognitive effort and quality in post-edition of 80-90% fuzzy matches and MT output looked to be similar. These results could lead to the suggestion that a similar level of remuneration for both types of match is in order, although a translation discount rate based on the averaged performance of MT rather than the predictive, per-segment approach of TM seems impractically broad. In addition, a further study comparing TM fuzzy matches with SMT proposals by Pelino (2011) found wide discrepancies in comparative effort required for different language pairs. However, as MT is increasingly integrated with TM, O'Brien's work may suggest fuzzy match thresholds below which MT output may prove more valuable than TM fuzzy matches. The methodology had satisfactory results and O'Brien and others (O'Brien and Doherty 2009; Carl 2009) have continued to use eye-tracking for their research.

Guerberof (2008) followed on the work of O'Brien in comparing effort required for post-editing MT output and 80-90% fuzzy matches in her pilot study. She used eight professional translators, measuring post-editing time with an online tool. She writes that, in current enterprise localisation projects, the addition of MT output to pre-translation has made "post-edition... one of the most requested activities in localization as opposed to full translation of new texts" (2008, p1).

O'Brien's study compared cognitive effort in post-edition of fuzzy matches and MT output. However, in this case, Guerberof measures time spent post-editing per word rather than cognitive effort, correcting for the translator's experience and errors. Translators were not made aware of the match value for fuzzy

matches from TM, nor were they made aware whether they had received MT or TM output. Although professional translators were used in this experiment, none were paid, which could affect quality. In addition, the translators were too few to make any generalisations. Nonetheless, Guerberof finds that less experienced translators spent roughly the same time editing MT and 80-90% TM matches. More experienced translators were faster using MT, although some appeared not to actually use the MT output at all. Due to a high level of errors in the texts translated using TM and the evidence of higher productivity using MT, Guerberof was led to question the value of TM fuzzy matches, and suggests that translators may too readily insert TM matches without forethought.

In a more detailed, mixed-methods update of this study carried out with 24 translators and three reviewers, Guerberof again finds that productivity of MT post-editing is close to that of editing 85-94% TM fuzzy matches. She adds the caveats that the MT output must be of high quality (her output was rated at 0.6 using the BLEU automatic metric), the terminology must be correct, and tagging not excessive. Based on her qualitative interviews, she suggests that the "perceived cognitive effort for MT is higher", as translators expect MT output to be of a poor quality and are thus dissatisfied with a pricing scheme that equates pay for post-editing with high TM matches (Guerberof 2012, p242).

Yamada (2011a) also considered the continuing integration of TM and MT in his PhD thesis, comparing the rates of productivity and cognitive load of editing TM matches and MT proposals using student and professional English to Japanese translators. Using the GTM (General Text Matcher) metric, he found that translators are faster post-editing an MT translation above a certain quality threshold (GTM = 0.464) than translating from scratch. At present, he says, the mean GTM score for SMT output is below the threshold. However, in some domains the disparity is small enough that the shortfall will, he predicts, be made up by 2013 or 2014, making the addition of MT to TM valuable, and rendering the technical translator's job one of editing or reviewing existing TM suggestions or machine-translated material without translating from scratch.

A pilot study by Carl et al. (2011) compared gaze time, cognitive effort, speed, and quality of the work of seven translators when either translating from English or post-editing Danish language machine translations of three British newspaper articles. While there were differences in the amount of visual attention given to the target part of the screen (the post-editors tended to look to the target text first), more interestingly, the post-editors saved time and produced better quality (as perceived by seven human evaluators) translations than those translating manually.

Carl et al. stress that this study is of too small a scale to draw firm conclusions, but the research forms part of a growing consensus that MT has a role in the current localisation industry. Zetzsche feels that this sea change occurred when "the concept of quality was replaced by usability" (2012, p51) and, as demonstrated in the research compiled for this section, MT has been shown to produce usable output. Research on TM quality is the focus of the following section.

2.7 TM Quality

A final group of studies surveyed here is concerned primarily with the quality of translations produced using TM data, although they also touch on issues such as translator productivity and translator attitudes to TMs, as discussed previously. Section 2.7.1 reviews research on error propagation in TMs and section 2.7.2 is concerned with comparing the effects of different ways of translating with TM. Finally, section 2.7.3 contains a brief look at quality assurance.

2.7.1 Error Propagation in TMs

Merkel (1996) said that TM should be an ideal tool to promote consistency in translation, as long as the text is repetitive and the source text is of good quality. He carried out a survey of 13 technical translators, providing them with examples of 50 source text segments, each with two different target text translations. Despite a stated preference for consistency (context permitting) among 12 of the translators surveyed, many translators have "difficulty in accepting a previous translation from the translation memory" (*ibid*. p158). While no translator came out against TM, they expressed concern over the translator's changing role, fearing a future where the translator is "reduced to somebody who presses the OK button" (Merkel 1998, p141). Merkel created a tool to check translation consistency and used it on documentation from IBM and Microsoft, finding that texts translated with and without the aid of TM were "more or less inconsistent" (1996, p166). Rather than concluding that this was a problem inherent in translation, Merkel considered that the inconsistencies found may have been the result of a "clash between an established translation culture" and the recently introduced technology (*ibid*. p166). He also concluded that, although translators surveyed preferred consistency, what each translator considered the "'best translation option' varies considerably among translators" (Merkel 1998, p148).

Bowker (2005) carried out a pilot study to evaluate the impact of TM on productivity and quality. Three groups of student translators were given a short translation assignment to complete in 40 minutes, the first group translating without a TM, the second with a TM, and the third using a TM seeded with errors. She found that the translators using TM were not critical enough of proposals offered by the TM system, especially when faced with time pressure. Many of the seeded errors from the TM were propagated in the final translations. Thus, she notes a conflict between productivity and quality, as time saved by using TM may be required to correct poor quality TM proposals. Sentences from different sources or different contexts within the same source text did not always use the same terminology, and she notes that "although it is frequently claimed that TMs improve consistency, this is not always the case" (*ibid*. p18). Another suggested reason for this is that translations "may become inaccurate over time (e.g., if terminology changes) or they may be inappropriate in a given context" (*ibid*. p19). In conclusion, she recommends the use of a "smaller TM containing well-chosen texts" to make the best use of TM. Although the translator may receive fewer suggested matches, prioritising precision over recall, those matches should be "of higher quality and will likely need less revision" (*ibid*. p19).

Rieche (2004) identified factors that, she said, lead to quality problems in TM in her MA thesis. She found inconsistently translated terms in an English to Brazilian Portuguese TM and suggests that multiple users of a single TM may be a cause of error introduction. When a TM contains errors, inaccurate 100% matches are more likely to be propagated when auto-propagation is used. She suggests manual revision of a TM at the end of each project, deletion of duplicated TUs and

> é preciso haver um controle sistemático das memórias, por meio de procedimentos regulares de revisão e manutenção (Rieche 2004, p168).
>
> *systematic control of memory through procedures for regular review and maintenance*.

Furthermore, when TMs are created by many people, she suggests forming a specialised team to assume responsibility for review and quality control of TMs.

In their study of concordance use among six professional translators using SDL Trados TWB 2007, O'Brien, O'Hagan, and Flanagan (2010) found that, despite the TM tool having a terminology plug-in (Multiterm), all participants reported that they regularly use the concordance function to search for terminology. Three participants said that they prefer to use the concordance feature rather than using Multiterm at all, and one translator "was not aware of the existence of Multiterm" (*ibid*. p3). This suggests that the integration of the terminology tool could perhaps be improved (and with the release SDL Trados Studio the software has changed substantially, but was still found to be inadequate by interviewees in section 6.4.4). Nonetheless, the study also found quality improvements when using the concordance feature as a productivity tool to search for sub-segment matches.

The topic of error propagation in TM use is the focus of research by Carlota Ribas López (2007), who believes that the dissemination of errors throughout a project from errors in a TM database is an under-reported phenomenon. In her pilot study, Ribas López seeded an aligned TM with three types of errors, before asking first translation students, then professional translators to translate using this TM. The data gathered was correlated with questionnaire replies detailing respondents' experience and reviewing strategies. The data produced by the students was inconclusive, and a new text was chosen for the professional translators, who were hired based on each having over three years of experience, via the ProZ.com website.

The professional translators were unable to spot all of the errors, and sent an error-filled TM back to the researcher, despite being informed that the TM would be re-used for future work. No correlation was found between errors left uncorrected and error type. The professionals spent a larger percentage of their time editing the texts, but again only one of the four translators reported errors

in the TM. All of the respondents from both groups believed that translators should be paid for all matches, although two of the professional translators responded that payment should be less for exact matches. These findings do little to refute López's hypothesis that TMs propagate errors, but she believes that more subjects, along with a form of screen recording and "retrospective feedback from the active participants", would be necessary for the data gathered to be of value to the industry.

Pym (2004) and Bowker and Barlow (2008) have also written of the dangers of recycling error-filled TMs, with the latter suggesting that this results in clients who are "irritated because the same passages that were corrected last year need to be corrected again. This is not the kind of added value the client was looking for" (2008, p16). García also points out that, due to unfamiliarity with the source text, the reviser of a localised text "may not only miss errors, but actually introduce new ones" (2008, p53). As yet, no full studies have been published (of which we are aware) about the propagation or addition of errors in TMs.

2.7.2 Choices in TM Workflow

A further study of TM quality was carried out by Wallis (2006), who compared the work of translators translating interactively and those using pre-translated texts. As more and more software developers announce server-side TM solutions, progress appears to decree that fewer applications and resources will be held locally in future. This leaves LSPs with the choice of whether to allow translators to work with a TM on the company network, or to provide them with pre-translated texts, populated with exact or fuzzy matches where available. Wallis investigated the effects of this move on productivity, quality, and translator satisfaction. Based on her pilot study using four recently qualified translation students, Wallis's evaluators found no difference in productivity and only a minor qualitative improvement in interactively translated texts, when compared with pre-translated texts. She did, however, find that the translators unanimously

favoured interactive mode, which concurs with reported opinion in García (2006). Wallis's research was, by her own admission, limited by funds available and thus her inability to engage professional translators for her work, but nonetheless provides a possible basis for future research.

A pilot study comparing the effectiveness of two translation styles using TM was carried out by Yamada (2011b). Using eight Masters-level translation students, he compared production speed between those using what he considered freely-translated TM content (obtained from industry) and participants using more literal TM content (that he had adapted from the industry TMs). He found that those using the more literal TM evinced a higher words-per-minute rate, especially when editing 90-99% fuzzy match suggestions, concluding that literal translation with TM is most conducive to optimum productivity.

2.7.3 Quality Assurance

The notion of translation quality in the current localisation industry is reduced to minimisation of errors via quality control, according to Pym. Quality is increased by adding a new layer of control, and if something is considered of inferior quality, it is rooted out. He says that quality is "a question of having enough testing and control procedures in place" and is maximised by "regulating processes" (2011, p6).

Following review, the final quality control procedure in the localisation process is often an automatic quality assurance (henceforth QA) check. Debove, Furlan and Depraetere (2011) find that the use of TM inevitably leads to "occasional erroneous acceptance of fuzzy matches", and that the rapid development of information technology, with shorter time-to-market and increasing volumes, means a "wider range of stumbling stones in the translation process" (p162). In the case of projects with tight deadlines or with tasks shared between many translators, "inconsistency errors become even more likely" (*ibid*. p162).

They compare the performance of five QA tools[4] that check for errors in respect of terminology, consistency, punctuation, formatting, number values, and tags, concluding that these tools, although fallible, are both useful and necessary. All of the tools tested were efficient in identification of inconsistent translations and punctuation errors, but in other categories generated a lot of false positives or incorrectly identified errors. The authors identified false positives with regard to sentence ending punctuation, consecutive spacing, and terminological inconsistency, which "impacts on the time spent on the QA check" as the relevant segments have to be opened (Debove, Furlan and Depraetere 2011, p186). In a test to compare the findings from this study with two commercial automatic QA analysis tools (ApSIC Xbench and QA Distiller) in section 5.6, the results also showed accuracy in the identification of inconsistent target segments but generated false positives. The reason for the incorrectly identified errors is unclear, as these automatic QA tools are proprietary programs and as such their functioning is not made explicit. Nonetheless, Debove, Furlan and Depraetere conclude that automated QA is a "crucial step in the translation workflow" (2011, p191). This concurs with the industrial notion of QA as experienced by the interview participants in Chapter 6 (see QA discussion in section 6.5.7).

[4] QA Distiller. 6.5.8, Xbench 2.8, ErrorSpy 5.0, SDLTrados 2007 QA Checker 2.0 and SDLX 2007 SP2 QA Check.

2.8 Future Direction of TM Systems

This section looks at how TM tools may evolve in the future. New iterations of TM systems have shown increased integration of MT, QA, and project management (PM) with TM functionality. Some recent research has suggested other additions and improvements to TM tools. Section 2.8.1 focuses on the addition of sub-segment matching and section 2.8.2 discusses shared and cloud-based TMs.

2.8.1 Sub-Segment Matching

Suggestions for the future direction of TM are included in papers by Wu et al. (2005) and Mitkov and Corpas (2007). In the former paper, the authors suggest a combination of sentence-level matching with "sub-sentential level matching and pattern-based translation", which they found gave positive results using MT evaluation metrics (2005, p371). In the latter paper, Mitkov and Corpas write of their current research on a 'Third Generation Translation Memory System,' which they say will find matches that do not retain the same semantic structure as the source segment, yet have the same meaning. However, as yet the authors have doubts as to "whether the new technology would deliver in a robust and scalable way", and they also state that it is "still a long way from operating in a practical environment" (2007, p2).

Some research on sub-sentential matching was also carried out by Colominas (2008), who suggests segmentation by noun phrase chunks. Initial research has so far proved inconclusive, as "only pre- or postmodified noun phrases have turned out to be especially adequate for pretranslation tasks" (Colominas 2008, p352). Since the early 2000s, DéjàVu has offered matching at word and phrase level using the 'assemble' function (García 2003). Some degree of sub-segment matching has been integrated into the latest generation of TM tools, such as MemoQ 5 and SDL Trados Studio 2011, although the latter requires 25,000

translation units to use the 'auto-suggest' sub-segment functionality. Zetzsche suggests that, as integration of sub-segment matching with TM is in its infancy, current "sub-segmenting approaches are almost as varied as the number of tools supporting them" (2012, p51).

2.8.2 Cloud-Based and Shared TMs

In 2003, Levitt predicted that, due to online TM sharing, "many freelance translators will no longer need to own a TM tool", and as a result "the localization workflow paradigm is set to change" (Levitt 2003, p41). Cloud-based[5] and online TM sharing first became a reality with the release of Lionbridge's Logoport software in 2006, which allows linguistic assets to be stored in a central repository and accessed online via the Logoport client, and the Lingotek language search engine, which was released in the same year. This was followed in June 2009 by the release of the free online Google Translator Toolkit, which by October 2009 could handle text in 345 languages (Google Blog 2009). This online TM tool allows the user to leverage from a shared or private TM, although by default any new translations are added to the shared 'Global TM'. As a result of this, and because aligned pairs from the tool are used to train Google's SMT engine, there have been some concerns about the confidentiality and ownership of TMs created with or added to the Toolkit, to the extent that the system has been deemed "not practical for use in most real-world professional translation projects" (Drugan and Babych 2010, p6).

TAUS (Translation Automation User Society) believe that combining TMs will allow for greater leverage and reported 3% to 5% increased leveraging in tests carried out by members with an aggregated TM from five sources (TAUS 2009). They have developed a 'Language Search Engine' (LSE), allowing users to search for concordance results within uploaded TMs. While, unlike with Google's

[5] Cloud computing is the delivery of a resource or service over a network.

Translator Toolkit, the LSE is explicit in permitting TM data sharing and reuse, Drugan and Babych have cited concerns relating to both systems with regard to the reuse of translators' work beyond the initial remit (2010, p8). Other companies such as MyMemory/Translated.net and BigTM.net/Digital Silk Road are following the same route, although some LSPs have voiced doubts about aggregated TMs, claiming that smaller "bespoke TMs" provide better results and less noise (Simultrans personal communication, 2009), echoing the opinion of Bowker (see section 2.7.1) and interviewees in the current study (see section 6.4.1). Most TM tools at the time of writing allow TM sharing via a central server and in 2011 Wordfast launched a free, fully-functional, Internet browser-based version of their TM tool that also permits cloud-based TM sharing (Wordfast 2011).

Asia Online carried out an analysis of shared TAUS TMs and the reuse of those TMs for training SMT systems, and found that the TM data contained inconsistencies such as those caused by punctuation, terminology, and letter case, in addition to the appearance of English words in the French target segments. The study found that TM data quality has an effect on the quality of SMT results, that smaller amounts of consistent, high quality data can yield better SMT results than large amounts of 'dirty' data, and that "unmanaged terms result in inconsistent translations", especially when TMs from different sources are merged (Vashee 2009, p48). This finding contradicts the conventional wisdom for SMT that "more data is better data" (Kenny 2012, p5). The increasing demand for low cost and high volume translation means that SMT is becoming more widely used in the localisation industry. This means that inconsistent TM data may have an impact beyond that on recycling using TM tools.

Opinion on the use of smaller product-specific TMs was mixed in McBride's investigation, with one forum poster stating a preference for an "ultra-large" TM to maximise leverage (2009, p127). Aggregated online TMs have not yet become commonplace, but there is increasing integration of online resources into the

localisation workflow. Lagoudaki also sees further movement in that direction. In the near future, cost-cutting and the opportunity presented by the ongoing integration of Idiom and SDL Trados has meant that several parties are rushing to create an open-source integrated TM and Translation Management System (TMS).

As someone involved in one such project, Bergmann (2008) cites MySQL and Linux as inspirations for the potential reward of a successfully deployed open-source localisation solution. He admits that current projects may not "win any commercially interesting market share", but believes that open-source will "sooner or later change the market for TM/GMS systems" (Bergmann 2008, p57). So far Welocalize have chosen to make their Globalsight translation workflow management tool open source (Global Watchtower 2009) and the open source java-based TM tool Omega T is one of several TM tools available free of charge to translators and continues to be developed[6].

The integration of Idiom and SDL, re-engineering of SDL Trados Studio (the most popular TM system at present), and the competition provided by upcoming tools such as the frequently updated MemoQ, is changing the CAT landscape and presenting opportunities for other developers. Whether they can take advantage in straitened financial times while taking account of users' needs remains to be seen. The current gap between a stable user-centric application that makes the best use of currently available technology and commercially available TM systems means that there is room for improvement for developers of CAT tools.

[6] Open source tools are also valuable for research purposes, as in Moran and Lewis' (2011, p6) work adapting Omega T and Globalsight to record "user behaviour within a system".

2.9 Summary

In this chapter we presented a short history of the conceptualisation and introduction of TM tools and reviewed several introductory guides that are largely uncritical, although notably Bowker (2002) introduced problems of inconsistency caused by old or low-quality TMs. Surveys and online fora have revealed concerns of TM tool users and the gap between developers and their customers. Work focussed on evaluation of TM tools has shown that fuzzy matches may be of comparable value to MT below certain thresholds, that character-string-in-bitext and sentence-based approaches have advantages and disadvantages in different situations, and that there may be work to do to improve the accuracy of retrieval in TM systems. Research on quality in TM has uncovered problems with consistency in translations due to changing terminology, poor quality TMs, and the tendency for translators to accept TM proposals without sufficient forethought. In addition, Reinke and Pym both noted that inconsistency may be added to a TM by propagation of an inappropriate 100% match, as ST content is no guaranteed measure of editing effort required on a TM proposal.

While valuable critical research has been focussed on TM, much of the work has been in the form of pilot studies and has used inexperienced student translators as opposed to those working full-time in the industry. Full-time translators' attitudes to TM have revealed the importance of ease-of-use and support without offering concrete suggestions to improve translation using TM. Research on quality in TMs has shown that consistency is an issue, and while there has been work (Whyman and Somers 1998; Gow 2003) on metrics for measuring retrieval, there is no published metric (to our knowledge) for measurement of consistency in TM environments. Published works have highlighted the need for consistency (Austermühl 2001; García 2003; Gow 2003; Pym 2004) and the likelihood of error propagation when using unclean TMs (Pym 2004; Bowker 2005; Ribas López 2008). In an effort to address the dearth of comprehensive

research focussed on TM consistency that uses real-world data, the current study presents a suggested typology for inconsistent TM segments, categorises commonly-found inconsistencies, and then applies those measurements to TMs produced by professional translators. While current plans to aggregate TMs are likely to produce a greater number of matches, they are also likely to exacerbate problems of inconsistency. If these large TMs are to be of value, there needs to be consistency in their proposed translations.

3 Pilot Study

3.1 Introduction

In this chapter we describe a pilot study that, following the literature review, formed the second part of our preparation for the main study. We begin with a profile of the data analysed, then discuss the methodology and some results (previously published as Moorkens 2009). The chapter concludes with some issues that informed our methodology for the main study in Chapter 4.

As outlined in section 2.9, there has been no work of which we are aware that focuses on consistency of TMs. As a result, this pilot study was experimental, a first-pass attempt to interrogate TM data for consistency with a view to finding whether TMs contained inconsistency and of what kind, and of informing a more refined approach to be used for the main study. The pilot study was carried out in the summer of 2009 using TMs received from Symantec, an industrial partner within CNGL.

3.2 Research Data Used In This Study

The pilot study investigated whether recurring segments in the ST were translated consistently: a comparison was made between levels of consistency achieved in a Symantec translation project from 2004 involving 70,276 words in US English that were translated into German, and a later translation from 2008 involving 31,681 words also translated into German. The documents translated were software manual updates for the Norton Ghost product (a backup and recovery tool for Microsoft Windows produced by Symantec).

Between 2004 and 2008 Symantec made some changes to its translation workflow: Machine translation (MT) was introduced to propose translations in cases where the TM in use could not provide matches of at least 85% fuzzy match value. Controlled language (CL) rules for technical writing were introduced in an effort to improve the quality of the MT output thus produced. Technical writers were provided with style guides containing writing rules requesting, among other rules, that writers maintain consistency, avoid the passive voice, and avoid long sentences (the CL limits sentence length to 24 words or less). The source text was then checked using the acrocheck tool[7]. Symantec had also begun to use a bespoke version of the Idiom Worldserver translation management system (TMS) to assist with project management of enterprise localisation jobs. The TMS is used to monitor the translation workflow and although translators may translate interactively within the TMS environment while sharing a central TM, sharing does not take place within the Symantec workflow. Instead, TMs are sent to translators and merged after the translation has been completed.

[7] 'acrocheck' is an automated content consistency and quality tool available from http://www.acrolinx.com/.

3.3 Method for Pilot Study

For the pilot study, collected samples were analysed to investigate whether the use of TMs eliminates inconsistency in translations. TM files were presented in a proprietary .tmw format from SDL Trados Translator's Workbench (TWB). Using SDL Trados TWB 2007, these files were exported to a text format so that they could be opened by non-Trados tools. They were separated into aligned texts in English and German using the UNIX grep command to search for the *<seg L=EN_en>* tag for English and *<seg L=DE_de>* for German within the text editor software Maruo.

The Wordlist tool in the Wordsmith lexical analysis software was used to find instances of repeated phrases in the source text. Within the Wordlist tool, the 'Make/Add to index' command was used to create a tokens file. When this tokens file was opened, the 'compute' command was used to search for the most commonly occurring two and three-word clusters in the source text. The separate files of aligned bilingual text were then loaded into the Paraconc bilingual concordance tool.

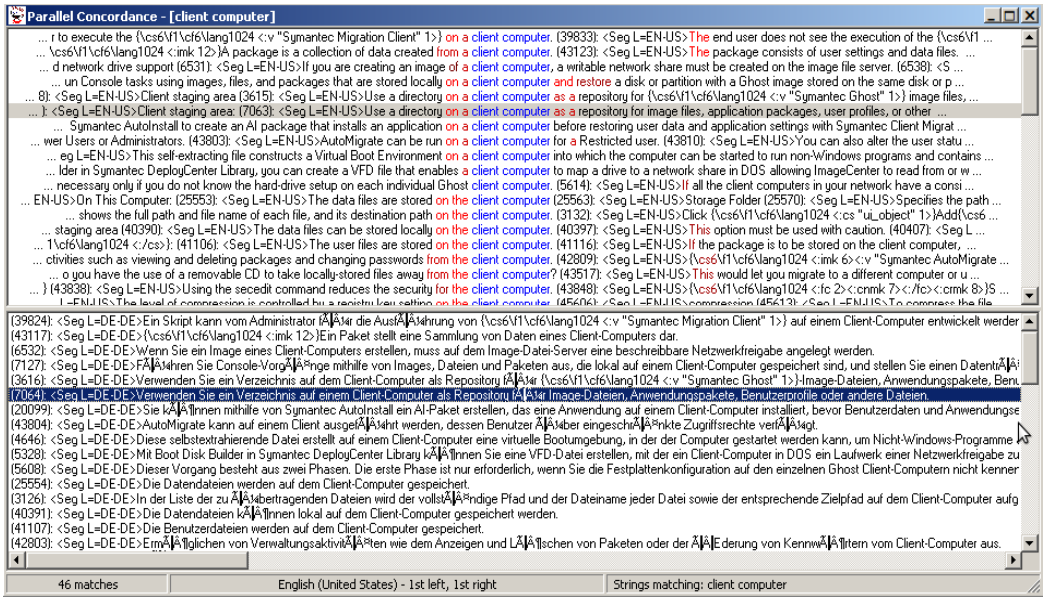


Figure 3.1 A Search for the Words 'Client Computer' in Paraconc

By searching for specific two- and three-word terms in Paraconc, such as 'boot package' and 'client computer', and then sorting the results alphabetically by the words to the left and right of the search term, repetitions could be more easily identified. This was intended as a method for finding identical ST segments, and as these segments were identified in an upper pane, the target text (TT) translations could be compared in the TT pane below. Multiple TT translations of a single ST segment could reasonably be expected to be identical. In cases where they were not identical, TT segments that contained inconsistencies were copied and pasted into an MS Word document and segments that contained similar inconsistencies were grouped together. In this way the initial categorisation of inconsistencies was driven by the inconsistencies found in the TM data.

As the study progressed, categories of inconsistencies became more defined based on which part of speech (e.g. noun phrase, verb phrase, etc.) was inconsistent, whether there was a word order change, or if the inconsistency appeared to be caused by segmentation of the ST. While these categories were appropriate to this TM data, an issue remained in some infrequent cases where inconsistencies fell into more than one category. In this pilot study inconsistencies were measured at the segment level and for each inconsistent TT segment type, we counted one inconsistency. This type of inconsistency was assigned to one of our categories with further categories added and the counting method was revised so as to accommodate TT segments containing more than one type of inconsistency in the main study.

3.4 Results of Pilot Study

From the 2004 TM data 59 inconsistencies were found in a sample of 2066 segments. In the 2008 TM 25 inconsistencies were found in a sample of 4433 segments. The 2008 data contained fewer clauses per sentence, and the average sentence length was found to be shorter. This is possibly due to the implementation of the controlled language rules as referred to in section 3.2. Table 3.1 gives an overview of the categorised pilot study inconsistencies.

Category of Inconsistency	2004 Data	2008 Data
Noun Phrase	24	11
Word Order	17	7
Preposition	12	0
Verb	6	3
Inconsistent ST Segmentation	0	4
Total Segments	2066	4433
Segment Level Inconsistencies	59	25

Table 3.1 Pilot Study TT Inconsistencies

In both sets of data noun phrases (NP) represented the largest category of inconsistency. NP inconsistencies comprised 40% of the total in 2004 (24 inconsistencies) and 44% in 2008 (11 inconsistencies). Verb phrase inconsistencies, on the other hand, accounted for 10% of the total in 2004 (6 inconsistencies) and 12% (3 inconsistencies) in 2008. Examples of NP inconsistencies included the following: the term 'account data' was translated as 'Kontodaten' in two instances, then 'Kontendaten' in another. Of the 46 matches found for the term 'client computer', an inconsistency was found in which the same segment was translated once using a definite article 'auf dem' and once with an indefinite article 'auf einem' for the same noun, as shown in **example 3.1:**

3.1.1s *Use a directory on a client computer as a repository for {\cs6\f1\cf6\lang1024 <:v "[Product Name]" 1>} image files, application packages, user profiles, or other files.*	3.1.1t *Verwenden Sie ein Verzeichnis auf dem Client-Computer als Repository für {\cs6\f1\cf6\lang1024 <:v "[Product Name]" 1>}-Image-Dateien, Anwendungspakete, Benutzerprofile oder andere Dateien.*
3.1.2s *Use a directory on a client computer as a repository for image files, application packages, user profiles, or other files.*	3.1.2t *Verwenden Sie ein Verzeichnis auf einem Client-Computer als Repository für Image-Dateien, Anwendungspakete, Benutzerprofile oder andere Dateien.*

In this case, the only discernible difference between the original segments was the placing of tags, which may suggest that the TM tool used did not discount the tags, and therefore did not suggest the segment as translated previously. Alternatively, one translator decided that his or her preference was to use the indefinite article in one instance despite the 100% match. Other examples from the 2004 data are translations of the phrase 'risk of the user', which became 'das Risiko des Benutzers' and 'Gefahr des Benutzers', and 'summary information', which became 'Sammelinformationen' and 'Übersichtsinformationen'. In another case, a compound noun ('der Paketname') alternated with a genitive construction ('der Name des Pakets').

Within the NP category there were also several cases of inconsistent terminology, and a noticeable amount of influence from the source language. This resulted in the use of an anglicised loan word in one case but not in another. One example is a translation of the English word 'the administrator', which in one case became 'der Administrator' and in another 'der Verwalter'. This influence of source language on one of the inconsistently translated words accounted for 13% (8 cases) and 12% (3 cases) of total inconsistencies found in 2004 TM data and 2008 TM data respectively.

Other inconsistencies found were in word order (29% or 17 inconsistencies in 2004, 28% or 7 inconsistencies in 2008) in the TT. For example the segment 'The selected user profile is restored to the user who is currently logged on.' was translated as in **example 3.2**:

3.2s *The selected user profile is restored to the user who is currently logged on.*	3.2.1t *Das ausgewählte Benutzerprofil wird für den derzeit angemeldeten Benutzer wiederhergestellt.*[8]
3.2s *The selected user profile is restored to the user who is currently logged on.*	3.2.2t *Das ausgewählte Benutzerprofil wird für den Benutzer wiederhergestellt, der derzeit angemeldet ist.*[9]

In some cases, the word order and word choice were both inconsistent. The phrase 'the template that was used to create the user package', for example, was translated as both 'Richtlinien für die Behandlung der zusammengetragenen Dateien' (lit: 'Guidelines for the handling of collated files') and 'Dateibehandlungsrichtlinien für einige der erfassten Dateien', (lit: 'File handling guidelines for some of the captured files') where 'erfassten' replaces 'zusammengetragenen' as modifier of 'Dateien'.

Further inconsistencies appeared to be a result of inconsistent segmentation (4 cases or 16% of the total in 2008). Although TM software by default accepts a single ST sentence ending with a full stop as a single segment, some sentences appeared once as a single ST segment and again as part of a larger segment of two or more sentences. A translator can manually change the segment size during translation, for example if two or more sentences in the ST are to be translated as a single TT sentence. The sentences containing these inconsistencies were not translated to a single TT sentence and, because of the

[8] Lit: 'The selected user profile is for the currently logged on user restored'.

[9] Lit: 'The selected user profile is for the user restored who currently logged on is'.

inconsistent segmentation, a translator would not have received a 100% match although the ST sentences were identical. **Example 3.3** illustrates this phenomenon.

3.3.1s *This is updated at the end of the [Product] process, which might require you to reinsert the first disk in the span set. [Product name] prompts you for the first disk in the span set and for subsequent volumes when restoring…*	3.3.1t *Diese Informationen werden am Ende des Duplizierungsprozesses aktualisiert. Für die Aktualisierung ist es notwendig, dass Sie den ersten Datenträger des verteilten Datenträgersatzes erneut einlegen.*[10]
3.3.2s *This is updated at the end of the [Product] process, which might require you to reinsert the first disk in the span set.*	3.3.2t *Diese Informationen werden am Ende des {Product}-Prozesses aktualisiert, möglicherweise müssen Sie dann den ersten Datenträger des gesplitteten Satzes nochmals einlegen.*[11]

As a result of inconsistent segmentation segment 3.3.1t may not have been suggested to the translator, leading to segment 3.3.2s being translated as 3.3.2t.

An initial finding from the pilot study was the prevalence of NP inconsistencies. Karamanis, Luz and Doherty suggest that many terms are not specified in key term lists (2010, p4). The prevalence of NP inconsistencies may also be due to the presence of segments from different sources and times in the TMs. Also, larger projects may be divided between several translators and their TUs added to the TM thereafter, introducing inconsistency, as mentioned in section 2.6.3.

[10] Lit: 'This information is updated at the end of the duplication process. For the update, it is essential that you insert the first disk of the distributed volume set again'.

[11] Lit: 'This information is updated at the end of the {Product} process, you may have to insert the first disk of the split set again'.

Far fewer inconsistencies were found in the 2008 TM data, when fuzzy matches below the threshold of 85% were not used in the translation cycle.

3.5 Issues Arising from the Pilot Study

A number of methodological issues arose in the pilot study. During the course of the study, the manual search for inconsistencies within Paraconc was found to be laborious, and particularly so when searching among tags. The method would not have scaled up easily to a larger sample of TM data. Some elements of the process could be automated to remove tags and improve the speed and measurement reliability of a study without compromising accuracy. The search for commonly-occurring word clusters in Wordsmith and the follow-up search for those word clusters in Paraconc were unreliable methods of capturing all repetitions in the ST, although they succeeded in the initial aim of discovering whether there were TT inconsistencies present in the TM. A better method was nonetheless required for the main study.

As discussed later in section 4.3.2, a single case study forms a poor basis for generalisation. A decision was thus made to carry out the main study on more than one set of data, to show that the methodology was replicable. It was decided to use English/Japanese data in addition to English/German, to show that the methodology may be applied for non-European languages. Japanese was chosen based on the researcher's language competence.

At the data analysis stage of the pilot study, inconsistencies were counted based on the number of types at TT segment level. In example 3.2, where one ST segment is translated as two different TT segments, we counted two inconsistencies. However, while translating interactively using this TM, the translator would presumably have received one of 3.2.1t or 3.2.2t as a suggested match. The translator's decision to change the suggested TT segment to the second TT translation should count as one inconsistency, if the suggested segment is considered as the 'first' or 'master' segment. For this reason we changed the system of counting for the main study, assigning one TT segment

the status of the master segment in cases where there were multiple types of TT translated from a single ST segment. This is further explained in section 4.4.

Inconsistencies were largely categorised according to part-of-speech as it became clear while searching for inconsistencies that many translations of a single segment differed only by a noun or verb. Categories for inconsistent word order and inconsistent punctuation were added as these became apparent from the data. This pattern continued into the main study. Although this seemed successful, some inconsistent segments fell into more than one category, for example:

3.4s *This chapter includes the following topics:*	3.4.1t *Dieses Kapitel enthält folgende Themen:*[12]
3.4s This chapter includes the following topics:	3.4.2t *In diesem Kapitel werden folgende Themen behandelt:*[13]

In segments such as 3.4.1t and 3.4.2t, with verb and prepositional changes, counting as a single inconsistency would not give the full picture. Making each adjustment takes time for a translator - salient as TM tools are intended to save time and money. As a result, it was decided that inconsistencies would be measured at segment level and again at sub-segment level in the main study, counting each inconsistency found within the segment. Full details of the methods of counting and categories for the main study can be found in section 4.4. It quickly became apparent from examples such as 3.1, and the addition of ST segmentation inconsistency as a cause of TT inconsistency in the 2008 TM, that measuring TT inconsistencies alone did not provide a full picture of the level of consistency within a TM. As a result, the four categories of consistency in section 4.4 were chosen for the main study.

[12] Lit: 'This chapter contains [the] following themes'.

[13] Lit: 'In this chapter are the following themes dealt with'.

The pilot study was an exploratory project and as such could not be expected to explain the reasons behind the prevalence of certain categories of inconsistency. It was felt that for the main study explanatory data, provided by professionals with experience of TM, would be valuable so that consistencies may be minimised in future. This was addressed in the main study by choosing a mixed methods research design as outlined in the following chapter. The pilot study and issues raised within it proved valuable as a framework to be improved upon in the methodology used in the main study.

3.6 Summary

This chapter described a pilot study that was carried out in 2009 in preparation for the main study, the methodology for which follows in Chapter 4. The pilot study was an initial attempt to check for consistency in two English-to-German TMs created during software updates. The methodology involved converting the data from a proprietary format, extracting segments, and searching for commonly occurring word clusters. By searching for these word clusters using a bilingual concordance tool and sorting the results alphabetically, it was possible to see whether segments had been repeated and to check for introduced inconsistency.

The pilot study found inconsistencies, mostly of noun phrase and word order, in the data. Fewer inconsistencies were found in the second TM, possibly as a result of a translation process that involved controlling the source text and using a translation management system. The study did, however, show that there were inconsistencies in the TMs that, using an improved and less laborious methodology, could be measured and categorised. It was decided to use a replicable quantitative approach, supplemented by a series of follow-up qualitative interviews for the main study, updating methods of counting, and increasing the scope to search for source text inconsistencies.

4 Methodology

4.1 Introduction

The methodology for the main study implements changes from and additions to the exploratory pilot study as outlined in the previous chapter and aims to address a shortfall in the existing literature as identified in section 2.9. It begins with a desk-based quantitative analysis of TM data to search for inconsistencies. The second, qualitative stage of the research is based on interviews with translators, QA (quality assurance) specialists, and others who work with TMs. In this chapter we outline the methodology adopted in the main study. We begin by restating our research questions (section 4.2). We then discuss the research paradigm within which our research is situated (4.3) before going on to describe the particular choices made in the design of our study (4.3.1). Section 4.4 outlines how key concepts (consistency, segment-level consistency, etc.) are operationalised. Section 4.5 describes the data used in this study, following which section 4.6 provides an overview of the first quantitative phase of the current research and section 4.7 provides an overview of the subsequent qualitative phase.

4.2 Research Questions

As previously detailed in section 1.1, the research questions motivating this study are as follows:

RQ1: Are TMs consistent?

The assumption behind the use of TMs in the localisation industry is that they produce cheaper, faster, and higher quality translations than translations produced without TM. For this to be the case translators would have to accept exact (100%) or high (eg 98-99%) fuzzy matches in order to retain the internal consistency of the TM (associated with the quality criterion) and to maximise the benefits in time saving and thus cost saving that TM promises.

RQ2: How can consistency be identified and measured in TM data?

As seen in the literature review, a study testing translations or their corresponding TMs for consistency has not yet been published, possibly due to the difficulty in procuring TM data or the amount of time required to carry out such a study. In addition, an accurate method for identifying and measuring consistency in TMs that could be applied beyond the present study would be of value within the localisation industry as a test of the consistency of - and potential error propagation from - a company's TM data. This in turn could lead to a readjustment of a company's expectations in relation to potential gains in time and cost when using TMs. A hypothetical scenario setting out the potential cost of ignoring inconsistency is set out in section 7.4. This test could also be applied to an amalgamated TM from different sources in order to measure the effect that combining TMs has on consistency.

The final set of research questions looks to the wider context of how TM tools are used in localisation, and asks how translation professionals in the industry

view inconsistency in the context of TM tools. Specifically, we are interested in finding out:

RQ3: In the opinion of translators and those working in the localisation industry, is inconsistency a problem?

Inconsistency is considered a problem in the literature reviewed in section 2.7. This question addresses whether the literature reflects the opinions of those working in the translation industry.

RQ4: In the opinion of those same people, what causes inconsistency?

RQ5: What procedures could be implemented to minimise inconsistency?

4.3 Mixed-Methods Research

We have chosen to use a mixed-methods research design in this study. We provide an overview of mixed-methods research in section 4.3.1 and then discuss how it was applied in this study in section 4.3.2, explaining the sequencing of the two-phased study and considering some threats to validity and reliability.

4.3.1 Introduction to Mixed Methods Research

Williams and Chesterman write that quantitative research allows us to see the "generality of a given phenomenon or feature... regularities, tendencies, frequencies, distributions", whereas qualitative research, although not permitting us to generalise, can lead to "conclusions about what is possible" (2002, p64). Creswell and Plano Clark believe that a combination of quantitative and qualitative research in a mixed-methods approach can provide a "more complete picture" by exposing generalisations along with participants' insights (2007, p33).

Tashakkori and Teddlie make explicit the link between mixed methods research and the philosophy of pragmatism, the epistemological basis of the current study, calling it a "deconstructive paradigm that debunks concepts such as "truth" or "reality" and focuses instead on "what works" as the truth regarding the research questions under investigation" (2003, p713). The pragmatic worldview was adopted by philosophers such as C. S. Peirce, William James, and John Dewey, the latter of whom explained that it "does not insist upon antecedent phenomena but upon consequent phenomena; not upon the precedents but upon the possibilities of action" (1931, p32). Mixed methods research based on the pragmatic paradigm "seeks to clarify meanings and looks to consequences", and became popular in the 1990s as several seminal works were published and it emerged as the "third methodological movement" (Cherryholmes 1992, p13;

Teddlie and Tashakkori 2009, p76). Frey et al. write that, although quantitative studies are understood to have high measurement reliability, a combination of qualitative and quantitative studies of the same concept enhances validity and reliability of measurements along with enhancing the "credibility of the conclusions they draw" (1991, p124). At the data analysis stage, Onwuegbuzie and Teddlie state that mixed methods analyses offer "a more comprehensive means of legitimizing findings" than quantitative or qualitative analysis alone (2003, p355).

In choosing mixed methods for the current study it is hoped to provide a more complete picture than a quantitative study alone. Rather than just quantifying inconsistencies in TM, credible and effective results should provide further insight into how inconsistencies are added and propagated, and thus suggest ways to minimise further inconsistency.

4.3.2 Research Design in the Current Study

For this study we applied a sequential, explanatory mixed methods design. Creswell and Plano Clark state that this design is appropriate to "when a researcher needs qualitative data to explain significant results" (2007, p72). There are two variants of this design: the participant selection model, in which the quantitative results are used to identify participants for the qualitative phase, and the follow-up explanations model as used in this study, in which the qualitative data is intended to expand upon the quantitative results (*ibid*. p72). The steps necessary in this research design are shown in table 4.1, adapted from Creswell and Plano Clark (2007, p73).

Phase One:	Quantitative Data Collection	Quantitative Data Analysis	Quantitative Results	Identify Results for Follow-Up
Phase Two:	Qualitative Data Collection	Qualitative Data Analysis	Qualitative Results	Interpretation Quantitative->Qualitative

Table 4.1 Explanatory Design (Follow-Up Explanations Model)

In the first, quantitative phase of the study, four sets of TM data collected from two companies were measured for consistency. In the second phase, qualitative interviews with translators and other translation professionals who work with TMs were conducted to explore their experiences of consistency issues in TM. It was also anticipated that these interviews would allow us to explain some of the findings from the quantitative stage of the study. Primary emphasis was placed on the quantitative phase as prescribed when using the follow-up explanations model variant of the explanatory mixed research method, and as the quantitative phase provided results from which the interview questions were drawn. Having chosen an independent level of interaction between the quantitative and qualitative research data, in that data from the two phases is analysed separately, the analyses are presented consecutively in chapters 5 and 6.

Our choice of a two-phase sequential explanatory mixed methods research design reflects an effort to provide research that may be of utility and follows on from the work on quality and consistency by Bowker (2005), Rieche (2004), and Ribas López (2007). This methodology was chosen following the pilot study, in which it was concluded that interview data could offer reasons for the presence of inconsistencies and suggestions to minimise inconsistency (see interview schedule in section 4.7.3). The pilot study reported on in Chapter 3 proved the feasibility of a quantitative search for inconsistency in TM data and provided an opportunity to address any flaws or issues in the initial iteration of the research

design. This led not only to the choice of a mixed methods research design, but also to the decision to automatically parse segments and remove tags (other than those within the segment). The choice of a mixed-methods approach also mitigates the threat to external validity inherent in a case study.

External validity concerns whether findings can be generalised beyond the scope of the research (Shaughnessy 2009, p231). A single case study is seen to be a poor basis for generalisation (Stake 1995, p7). The current study, however, contains replicated multiple case studies, which according to Susam-Sarajeva tends to give "considerable advantages over single-case studies in terms of the rigour of the conclusions which can be derived from them" (2009, p7). Shaughnessy also writes that partial replication helps establish external validity by showing that the result is replicable (2009, p234). Susam-Sarajeva warns that although no number of case studies may be enough for generalisation, researchers should provide enough contextual information and detail of the data that "other researchers may judge the degree of fit" and whether findings may be relevant to other contexts (2009, p11). Schofield believes that case study findings are made generalisable by studying real-world data, thus making the study ecologically valid. She suggests that the goal of designing research is to "maximize the fit between the research site and what is more broadly [the case] in society" (2000, p77). In this research we tried to counteract threats to external and ecological validity by studying real-life TMs to reflect real-life circumstances.

The choice of a mixed methods design introduces a threat to validity at the stage between the two main phases of research when the researcher must decide what results from the first phase to pursue in the second phase. Creswell and Plano Clark have a series of recommendations for this stage that may mitigate the possibility of bias, such as choosing major themes as the basis for a follow-up, and using only significant results (2007, p148). This was a threat we were wary of in the course of this research, although the significant quantitative results should be apparent in Chapter 5.

Kvale and Brinkmann write that validation should “permeate the entire research process” and beyond, relating to the craftsmanship of the researcher and the historical quality of his or her research (2009, p248). This, they believe, is built up by a policy of checking and questioning of results and maintenance of rigour throughout the investigation. Conversely, they warn of the invalidating effects of overemphasising validity at the expense of creating thorough research, adding that the pragmatic view of validity is that it can only be measured in the effectiveness of research and the actions borne of it after the research has been completed and disseminated (*ibid*. p256). This is a view we will reflect upon in our concluding chapter. Throughout all phases of this study efforts (such as those highlighted in the following sections) have been made to maximise validity and reliability as part of the objective of producing effective research.

4.4 Operationalisation

In order to make a study transparent, operationalisation is a necessary step. This means defining the empirical, observable characteristics of key concepts to be measured in the study. These definitions must be adequate and intelligible, they must be accurate as universally agreed, and they must be clear and replicable by other scholars (Frey et al. 1991, p94). In the current study the concepts of 'consistency' and 'inconsistency' need to be operationalised, and we turn to this issue in section 4.4.1. Operationalisation of segment level and sub-segment level inconsistencies follows in section 4.4.2 and 4.4.3 respectively along with a discussion of counting methods, categorisation of inconsistencies, and threats to reliability.

4.4.1 Operationalising Inconsistency

If there are two or more words or segments that differ when we could reasonably expect to be formally identical, we describe them as 'inconsistent'. We expect TT segments to be formally identical if the corresponding ST segments are formally identical or contain minor non-semantic differences as stipulated later in this section. Inconsistencies are measured at two levels in this study: at the segment level and the sub-segment level. The usual default setting for segmentation in TM software is to segment at the sentence level following an end-of-sentence marker. The translator can choose to segment at another point of punctuation or at paragraph breaks. The segmentation in this case study is performed at carriage returns at the end of headings, at end-of-sentence markers, or at colons. The translator may choose to shrink the size of an individual segment or to expand to include the following segment. The former usually occurs if more than one source text segment will be translated as a single target text sentence; the latter if a section of a source text segment is to be translated as a single target text sentence. In chapter 5, our findings are

described using the terms 'TUs' (translation units) for the total number of TUs (including repetitions) in a TM or in a category, 'segment level inconsistencies', and 'sub-segment level inconsistencies'. The latter two terms will now be explained in detail.

4.4.2 Segment Level Inconsistencies

Segment-level inconsistency is observed where two segments that one could reasonably expect to be formally identical differ from each other in some way. Source segments may be expected to be formally different if their meanings differ. We thus use the term 'inconsistent source segments' to refer to cases where there are very minor formal differences between two source segments and such differences do not reflect any semantic differences between the segments in question. Such minor formal differences include differences in:

> capitalisation, tags, punctuation, spaces, character formatting, and spelling (where a segment may be inconsistent with another segment simply because of a misspelling, inconsistent use of British or US English spelling, or a typographical error in one of the segments).

In the case of target segments, it appears reasonable to expect segments that are translations of 'the same' source segment (i.e. segments that are translations of different tokens of the same source type) to be formally identical, especially in a translation memory scenario where the goal is to reuse existing translations for already encountered source segments.

Where there are two different translations (and thus two different target segments) for a single source segment type, we speak of target segment-level inconsistency. The differences between the target segments in question can be very minor formal differences (as defined above), but they can also be more substantial, in extreme cases even leading to semantic differences between the two segments.

At segment level, the following four categories are possible:

1. inconsistent source segments are translated as inconsistent target segments

2. inconsistent source segments are translated as consistent target segments

3. consistent source segments are translated as inconsistent target segments

4. consistent source segments are translated as consistent target segments

The current study is primarily interested in categories 1 and 3, but also in the possibility of consistency being introduced during the process of computer-assisted translation (category 2). Category 4 may be seen as the ideal in specialised translation, whereby the TM has provided the best possible leverage and thus saved the maximum possible amount of time and money. An example of each category from our TM data is given in table 4.2.

	Source Text	Target Text
Category 1	Callouts window	Fenster "Callouts"
inconsistent source-> inconsistent target	Callouts Window	Callouts-Fenster
Category 2	Plane, perspective	Ebenen, perspektivische
inconsistent source-> consistent target	Planes, perspective	Ebenen, perspektivische
Category 3	Camera button	Kameraknopf
consistent source-> inconsistent target	Camera button	Kamera- Schaltfläche
Category 4	text background	Texthintergrund
consistent source-> consistent target	text background	Texthintergrund

Table 4.2 Examples of TU Categories

Segment-level inconsistencies are counted as follows:

For each set of inconsistent target segments we identify the number of types n. The number of segment-level inconsistencies is the type count minus one (n-1). Thus in the case of a source segment that occurs 4 times in the TM, if there are 3 separate translations in the TT (one of which appears twice), then the number of target segment inconsistencies is 2 (or 3-1). Thus we give a special status (of 'master' or 'reference' segment) to one of the target segments, and treat the other two segments as inconsistent with that reference segment. The reference segment is the one which appears first in our sorted list, and which a translator could have, but did not reuse in unchanged form.

For example, the following four translations for 'Click an empty part of the drawing area.' appear in the TM data:

1a. Klicken Sie auf der freien Zeichenfläche.
1b. Klicken Sie auf einen freien Bereich der Zeichenfläche.
1c. Klicken Sie auf einen freien Bereich der Zeichenfläche.
1d. Klicken Sie auf einen beliebigen freien Bereich auf der Zeichenfläche.

Although there are four TT tokens, there are only three types: 1a, 1b, and 1d. If we assign the status of reference segment to segment 1a, the segments that are inconsistent with the reference segment are 1b (repeated for 1c) and 1d: thus we count two segment-level inconsistencies. When we have three types (n=3), and since our count is of type (n - 1), we count two inconsistencies.

4.4.3 Sub-segment Level Inconsistencies

At segment-level, source or target segments are either consistent or formally differ and are thus inconsistent. However, there may be more than one inconsistency within these segments. For this reason we also count and categorise inconsistencies found within inconsistent segments.

These inconsistencies are categorised mostly by part of speech, as in the pilot study, aside from those where word order has been changed. If there are more than three inconsistencies within a target segment, we consider that segment to have been wholly retranslated. These categorised inconsistencies may be further subcategorised; for example nominal inconsistencies can be further divided into those that exhibit lexical differences and those that differ in number (singular/plural).

These sub-segment level inconsistencies are counted in the same way as segment-level inconsistencies: we identify the number of types n, assign one the status of master or reference sub-segment, then count the types that are inconsistent with the part-of-speech or word order in the reference sub-segment. Thus the count is n minus the reference sub-segment (n-1). Again, the reference sub-segment is the one which appears first, and which a translator could have, but did not reuse in unchanged form.

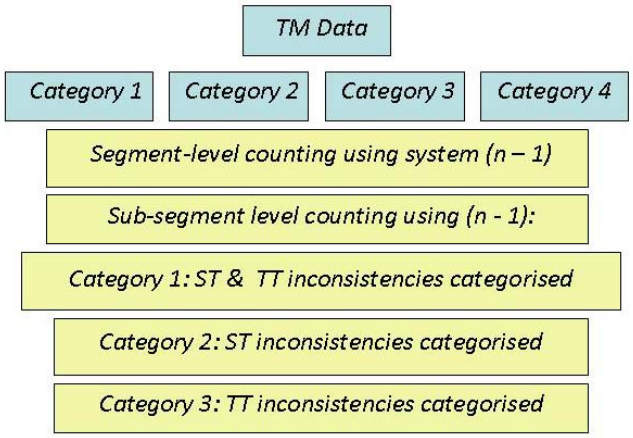


Figure 4.1 Schema of Categorisation of Data (Blue Measured in TUs)

Measurement validity concerns whether research measures what it is intended to measure. This study measures inconsistency as defined in this section, consistently applying the categorisation of TUs, segments, and sub-segment inconsistencies as stipulated in the course of evaluation and classification of issues in the TM data. Counting and categorisation of inconsistencies were double or triple-checked in an attempt to limit measurement reliability concerns and minimise errors, be they procedural or human errors.

4.5 Research Data Used In This Study

Access to data for use in this study was subject to protracted discussion and, in one case, the signing of a non-disclosure agreement. Sharing of data is a hot topic generally, but is particularly so in the localisation industry. As noted in section 2.5.1, this is due to financial value being attributed to TMs and the lack of a clear precedent with regard to ownership of TM data (see also sections 2.5.2 and 2.8.2). In addition, no LSP wants to be identified as having poor quality TM data. As a result, companies involved in localisation are protective of the TM data they possess and accessing a large TM corpus for research purposes presents a challenge. The data used for this study was made available thanks to connections between the Centre for Next Generation Localisation (CNGL) at Dublin City University, and the companies that donated data. This factor, along with the tools used and post-translation processes applied to the TM data, has an effect on the reliability and validity of the research. We look at the data in detail in this section, beginning with the format of the corpus as a whole, then looking at the four individual TMs, and finally considering some threats to validity.

4.5.1 Corpus Compilation and Format

The corpus for this research was TM data received from companies within the localisation industry. These TMs are aligned bilingual texts from commercial localisation of software documentation with English as the source language and either German or Japanese as the target language, where language pairs were chosen to fit the researcher's language competence. The data was presented in the TMX 1.4 (Translation Memory eXchange) format, an XML-based open standard compatible (to an extent) with many TM systems. The STs had been translated using SDL Trados Translator’s Workbench 2007 and Idiom Worldserver.

The TMX standard format was defined by the Localisation Industry Standards Association (LISA 2005) and consists of aligned segments of text in two or more languages, known as translation units (TUs). A TMX file begins by stating the version of XML with which it complies, followed by the <tmx> tag detailing the version of TMX in use, all in lower case. All TMX data is included between the opening and closing of the <tmx> tag. Details of the tool used to create the file and segmentation type, along with the edit language, creation time, and creation date following ISO standards, are given in the header of the TMX file along with notes and Unicode version information. All TMX files must be in Unicode. The next section of the TMX file is the body of the text, within the <body> tag.

The body section is made up of translation units, tagged <tu>. These are assigned an ID number (tuid), dates of last usage, and information about the domain of the translation. The aligned text in each language is within the translation unit variant tag <tuv> preceded by listings for the creation date, creator's ID, and language, along with information about any subsequent edits. The segments of text themselves, as identified by a Python script (see Appendix C) in phase one of the study (see section 4.6), are enclosed within segment tags <seg>. Following the last translation unit variant, the </tu> tag marks the ending of that translation unit. Finally, the </body> tag closes the body of the file and </tmx> marks the end of the TMX file. All files used in this study followed this standard.

An example of a TU in TMX format follows:

<tu srclang="en" creationid="[Translator's Name]"
creationdate="20061211T191159Z" changeid="[Translator's Name]"
changedate="20061222T193026Z" tuid="64594480--1">
<prop type="x-idiom-source-ipath">/English/aug_00.xml</prop>
<prop type="x-idiom-target-ipath">/German/aug_00.xml</prop>
<prop type="x-idiom-tm-uda-Import Job Id">1776</prop>
<tuv xml:lang="en">

<seg><ph x="1">{1}</ph>Current location and how often to check for new online content.</seg></tuv>
<tuv xml:lang="de">
<seg><ph x="1">{1}</ph>Aktueller Standort und Häufigkeit der Abfrage von neuen Online-Inhalten.</seg></tuv>
</tu>

4.5.2 Data Profile

The data used in the first phase of this study is four sets of TM corpora obtained from two companies. TMs A and B are from documentation for an updated iteration of a technical drawing tool, and TMs C and D are from a 2010 update of another technical drawing tool. All four TMs were presented in the TMX format, parsed using a Python script[14], copied into the LibreOffice[15] Calc spreadsheet tool and categorised as per section 4.4.

As shown in table 4.3, TM A is an English-to-German TM containing 22,691 TUs of aligned segments of which 188 contain only numbers, dates, or punctuation symbols. The remaining 22,503 TUs were categorised as specified in section 4.4.

TM B is an English-to-Japanese TM from the same project as TM A, containing 18,799 TUs. After removing those segments that contain only numbers, dates, or punctuation symbols, 18,650 TUs remained to be categorised.

TM C is an English-to-German TM containing 301,583 TUs. After removing those that contain only numbers, dates, or punctuation symbols, 293,924 TUs remained.

[14] The full script is contained in Appendix C.

[15] Version 3.4.5 of a cross-format open source office suite available from www.libreoffice.org.

TM D is an English-to-Japanese TM containing 298,700 TUs from the same project as TM C. After removing the TUs that contain only numbers, dates, or punctuation symbols, 292,258 TUs were left.

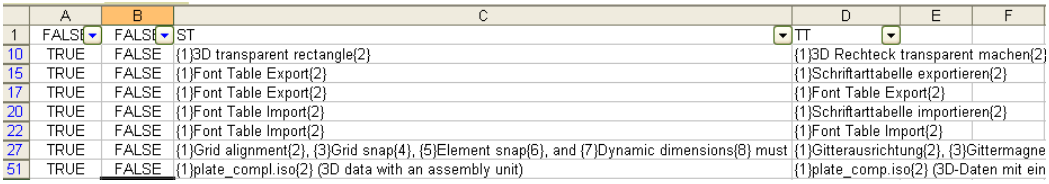
	TM A	TM B	TM C	TM D
Target Language	German	Japanese	German	Japanese
Total TUs	22961	18799	301583	298700
Total less TUs containing only numbers or punctuation	22503	18650	293924	292258

Table 4.3 Overview of TM Data

For TMs C and D, a sample of 50,000 TUs was studied. In order to confirm homogeneity between the sample and the full TM corpus in each case, a chi test was carried out using Microsoft Excel 2010 (version 14.0) software. The test was based on comparative measurements of corpus statistics as measured using Wordsmith Wordlist software (as used in chapter 3) and comparative measurement of the frequency of category 3 inconsistencies. The corpus statistics used were types (distinct words), standardised type-token ratio, and mean word length (in characters).

Category 3 inconsistencies were measured for the whole corpus within the Libreoffice Calc spreadsheet software by entering ST segments in column C and the TT segments in column D. In column A we searched for ST repetitions using the formula "=EXACT(C1,C2)", similarly in column B we searched for TT repetitions. By filtering for column A='TRUE', our search was limited to repeated ST or categories 3 or 4. Then by further filtering for column B='FALSE', limiting the search to exact ST repetitions with inconsistent TT, only category 3 TUs remained. As the final repetition was eliminated (due to not being an exact

match of the following segment), this followed our counting system of (n-1). Results can be seen in figure 4.2.



	A	B	C	D	E	F
1	FALS	FALS	ST	TT		
10	TRUE	FALSE	{1}3D transparent rectangle{2}	{1}3D Rechteck transparent machen{2}		
15	TRUE	FALSE	{1}Font Table Export{2}	{1}Schriftarttabelle exportieren{2}		
17	TRUE	FALSE	{1}Font Table Export{2}	{1}Font Table Export{2}		
20	TRUE	FALSE	{1}Font Table Import{2}	{1}Schriftarttabelle importieren{2}		
22	TRUE	FALSE	{1}Font Table Import{2}	{1}Font Table Import{2}		
27	TRUE	FALSE	{1}Grid alignment{2}, {3}Grid snap{4}, {5}Element snap{6}, and {7}Dynamic dimensions{8} must	{1}Gitterausrichtung{2}, {3}Gittermagne		
51	TRUE	FALSE	{1}plate_compl.iso{2} (3D data with an assembly unit)	{1}plate_comp.iso{2} (3D-Daten mit ein		

Figure 4.2 Screenshot of Filtering for Category 3 TUs

The chi test found no significant difference between the sample and the full TM. Full details are in sections 5.4.1 and 5.5.1.

4.5.3 Threats to Validity

Although random sampling is recognised as the best method for generalisation (Frey et al. 1991, p134), sampling for the current study was restricted to data made available to Dublin City University and the researcher and thus non-random, introducing a threat to external validity. The scarcity of TM data made available for research purposes means that access to the data used in this study was both opportune and unusual. Following translation, there are usually review stages, after which the TM is updated with reviewed and possibly edited segments. This was the case with the data used in the current study. Given that post-translation reviewing is usual practice within the bulk localisation industry, and that we use data from two discrete sources, it is hoped that we may reach generalisable conclusions.

A threat to internal validity identified by Frey et al. is the threat due to the subject of research. In the quantitative phase of the study this pertains to the data selected for analysis. Prior to translation, the ST data was written according to style guides that specified conventions, dictionaries, glossaries, and fonts to be used in the drafting of technical manuals. The TT data was created using SDL Trados TWB and the TM function in Idiom Worldserver. While each TM tool

essentially does the same job, they may tag and segment data in different ways. In our study we were able to overcome problems caused by such differences as the Python script we used stripped away tags and formatting, retaining only the segments of text between the <seg> segment tags, in the same text format in the ST and corresponding TT.

4.6 Quantitative Phase

This section begins with an overview of the first phase of the research and introduces our hypothesis that the TM will contain inconsistencies. We then look at how the research was carried out in practice in section 4.6.2.

4.6.1 Overview

The first phase of this research is a quantitative study in two stages. It is a naturalistic empirical study in that the aim is to derive new information from real-life data (Williams and Chesterman 2002, p62). The first stage is the corpus analysis stage. A corpus is a naturally occurring body of text, usually held in electronic form. Corpus-based research uses software to assist in the processing of linguistic data and the analysis of particular linguistic phenomena based on these corpora. The corpus in the current study consists of several aligned bilingual translation memories from the computing domain containing ST segments and corresponding TT segments as translated by professional translators (detailed in section 4.5). With the assistance of corpus processing software, we can search the aligned TM data for ST repetitions and see if these repetitions have been uniformly translated in order to investigate RQ1. The assumption that forms the basis of the prevalence of TM in the localisation industry is that these ST segments will have been consistently translated. Our hypothesis in this study, having found inconsistencies in the pilot study, is that the translations of repeated ST segments will contain inconsistencies.

4.6.2 Procedure

The aim of the quantitative study was to identify translation units that fall into the four categories as specified in section 4.4, and repeated here for convenience:

1. inconsistent source segments are translated as inconsistent target segments

2. inconsistent source segments are translated as consistent target segments

3. consistent source segments are translated as inconsistent target segments

4. consistent source segments are translated as consistent target segments

This was done by using a Python script to extract the ST and TT segments from the <seg> tags. Procedure validity and reliability require that research is conducted accurately and consistently (Frey et al. 1991, p126). Thus, in the present study, the parsing of and extraction of segments from TM data was consistently applied to all TMs using the same Python script before they were examined manually. The extracted, aligned segments were pasted into a spreadsheet using the Libreoffice software and sorted alphabetically so that repeated segments would appear consecutively. When repeated ST segments were found (or those containing the minor inconsistencies as specified in section 4.4), the corresponding TT was checked for consistency. We also searched the TT for repeated segments and where they were found, checked the corresponding aligned ST segments for consistency, to see if consistency was introduced via translation using TM as per category 2. TUs extracted in this way were copied to a new spreadsheet and classified according to whether they belonged to category 1, 2, or 3.

TUs in each of these spreadsheets were then categorised using the categories of sub-segment inconsistency from the pilot study: noun, verb, adverb, punctuation, preposition, word order, tag inconsistency, typographical error, or complete retranslation. In the case of category 1, TUs were further categorised by ST inconsistency: capitalisation, tags, punctuation, or typographical error. The topics chosen for the follow-up qualitative study were based on these results.

Examples used in the data analysis are presented in the following format:

4.0s *Consistent ST segment.*	4.0.1t *Inconsistent TT segment with inconsistency highlighted.*
4.0s *Consistent ST segment.*	4.0.2t *Second variant of inconsistent TT segment with inconsistency highlighted.*

4.7 Qualitative Phase

This section begins by outlining the broad methodology adopted in the second phase of the research, providing details of the interview type and sampling method. Thereafter we examine how the research was carried out in practice and list the interview questions. Details of the interview transcription process are in section 4.7.4, followed by a discussion of the method of data coding chosen in section 4.7.5. Potential threats to validity and reliability are highlighted at each stage.

4.7.1 Overview

The second phase of this research is a series of qualitative interviews with translators and others in the localisation industry with experience of using TM tools in order to find the causes of inconsistency and methods of minimising inconsistency. These are in the form of face-to-face personal interviews or, where this was not possible, telephone interviews, seeking opinions on results and conclusions reached in the quantitative phase of the study. Interviewees who are translators are usually native speakers of the target language who may also review and edit draft translations done by others for quality assurance purposes[16].

Semi-structured interviews, usually allowing probes once the interviewee has begun to answer, are the "most common qualitative strategy used in mixed method design" (Morse and Niehaus 2009, p127). We have chosen what Quinn Patton calls a standardised open-ended interview technique, with scripted questions and prompts, as this means the interview is highly focussed and makes responses easy to compare (2002, p346). It was hoped that this would minimise

[16] Rassmussen and Schjoldager (2011) found that amongst their survey respondents "most revisers are qualified translators, who spend much of their time translating".

the effect of the interviewer, but we have allowed for the interviewer to remain active, that is, to prompt or ask for further explanation if necessary, although the approach does not permit pursuit of unanticipated topics (*ibid*. p347). It is also important that responses are not supplied by the interviewer so that the interviewee can express his or her perspective in an open-ended response. We have scripted some prompts from the interviewer so as not to introduce any further bias, although even a tightly scripted interview cannot be devoid of input from both parties. All participants are "inevitably involved in making meaning" (Gubrium and Holstein 2003, p78).

There are two ways in which a researcher may threaten the internal validity of interview data. The researcher personal influence effect occurs when the researcher's characteristics (such as sex or ethnicity) or behaviour influences the behaviour of the subject. Johnson and Turner suggest remaining non-judgemental to interviewee responses as a method of minimising bias (2003, p305). The researcher unintentional expectancy effect occurs when the researcher inadvertently influences the subject. These two effects can be minimised by using research assistants or by standardising research procedures (Frey et al. 1991, p126). In this study, the former was not feasible, so the latter method was employed. Questions for the qualitative study were scripted so as not to influence the subjects, and although interview venues or methods (face to face or by phone) varied, all possible efforts were made to keep the interview procedure consistent in application of the interview structure, recording, transcription, and coding of the interview material.

For interviewee selection, purposive sampling was used. This means that subjects who would provide the most detailed information about the research questions were chosen, emphasising their depth of knowledge rather than seeking a large sample of respondents. Purposive sampling is associated with qualitative research and provides narrative data. Researchers using purposive techniques tend to minimise the sample size, selecting only cases that might "best illuminate and test the hypothesis of the research team" (Kemper,

Stringfield and Teddlie 2003, p279). The subtype of purposive sampling in this case was homogeneous cases sampling, which aims to gather opinions from people who are "demographically, educationally, or professionally similar" (*ibid*. p282). In this case, those sampled had worked as translators or with TM data for at least five years and were considered professionally similar.

Selection of subjects may influence the validity of research, but in this case we felt that translators (or those who work with TM) would be most accustomed to searching for inconsistencies in translations that had been overlooked by other translators, and best able to describe phenomena found in the TM data. There may be reactive effects from respondents who, aware that they are targets of study, may be defensive, dishonest or extreme in their responses (Johnson and Turner 2003, p302). In our interviews we believed that the respondents would have no reason to state anything other than their honest responses to the questions, particularly as they were not made aware of the sources of inconsistent data and also since their anonymity was assured. Inter-subject bias results from when subjects influence each other; however in this case subjects were all interviewed individually and worked on a freelance basis or were drawn from different language service providers, thus minimising the potential for subjects to confer.

4.7.2 Procedure

Thirteen interviewees were chosen based on the homogeneous cases sampling subtype of purposive sampling (see section 4.7.1; Kemper, Stringfield and Teddlie 2003, p279). Calls for potential interviewees were circulated via email and Twitter, and translators were approached at several industry events. Some of the interviews took place in the interviewees' workplaces, for the purpose of convenience and for the added benefits of seeing their habitat and having the chance to meet others mentioned in explanations and accounts (Lindlof and Taylor 2002, p185). Where this was not possible, interviews were carried out

remotely via Skype (version 2.2 for Linux) and recorded using Skype-Recorder version 0.8[17].

Each interview was limited to one hour due to constraints on respondents' time. Interviews were recorded to .m4a format using an iPhone 4 and the Voice Memo digital recording app, or to .mp3 via Skype-Recorder. The digital recording app was tested and the phone charged prior to meeting with interviewees so as to minimise adjustments during the course of the interview, as this has been reported to inhibit interviewees (*ibid*. p188). Despite inhibitions about the recording of interviews, it is a standard procedure, recommended for understanding the role of all parties and for completeness and accuracy (*ibid*. p188; Elliott 2005, p33). Notes were also taken manually as a backup. Interviewees were first given an explanation of the research and were told the purpose of the interview. They were told of potential uses of the research findings (such as inclusion in this thesis and in journal articles), how their identities and that of their company would be anonymised, and assured of secure storage of the data. Their consent was requested for recording the interview on this basis (see Appendices A and B containing the Plain Language Statement and Informed Consent Form as required by the Dublin City University Research Ethics Committee).

4.7.3 Interview Schedule

The interviewees were asked questions based on the findings from the quantitative study. These questions were:

1. What is your job?
2. What TM tool do you use, and what version of that tool?
3. What do you consider the benefits and disadvantages of TM?

[17] Skype is a voice over IP telephony service available from www.skype.com, Skype-Recorder is a tool for recording Skype conversations available from http://atdot.ch/scr.

4. You mentioned/did not mention consistency. What is the effect of TM on consistency?
5. Have you come across source text inconsistency?
    - *Scripted prompt: For example, letter case, punctuation, or script inconsistency?*

5a. If the answer to 5 is 'yes': Do you find ST inconsistencies frequently, and do you have processes to deal with ST inconsistency? What effect do these have on target texts?

6. Have you seen similar examples to any of the following examples? What do you consider their effect on translation quality?

Question 6, example A: Target text term inconsistency (ST influence)

4.1s *All lines that have been converted using the {1}Create surface borders{2} function can be recognized easily since they are drawn with the {3}Border{4} pen.*	4.1.1t *Alle Linien, die mit der Funktion {1}Flächenränder anlegen{2} konvertiert wurden, können Sie leicht erkennen, da sie mit dem Stift mit der Bezeichnung {3}Border{4} gezeichnet werden.*
4.1s *All lines that have been converted using the {1}Create surface borders{2} function can be recognized easily since they are drawn with the {3}Border{4} pen.*	4.1.2t *Alle Linien, die mit der Funktion {1} Flächenränder anlegen{2} konvertiert wurden, können Sie leicht erkennen, da sie mit dem Stift mit der Bezeichnung {3}Rand{4} gezeichnet werden.*

4.2s *A new layer group filter can be nested only under another group filter.*	4.2.1t 新しいレイヤグループ フィルタは、他のグループ フィルタに対してのみネストできます。
4.2s *A new layer group filter can be nested only under another group filter.*	4.2.2t 新しい画層グループ フィルタは、他のグループ フィルタに対してのみネストできます。

Question 6, example B: Source text format retention

4.3.1s *(Do Not Add to Workspaces or Add to Woprkspaces)*	4.3.1t *(Zu Arbeitsbereichen Nicht Hinzufügen oder Zu Arbeitsbereichen Hinzufügen)*
4.3.2s *(Do Not Add to Workspaces or Add to Workspaces)*	4.3.2t *(Zu Arbeitsbereichen nicht hinzufügen oder Zu Arbeitsbereichen hinzufügen)*

4.4s *Accesses Dimensioning mode*	4.4.1t 寸法記入モードにします
4.4s *Accesses Dimensioning mode*	4.4.2t 寸法記入モードにします。

Question 6, example C: Alternated phrases propagated throughout the TM

4.5.1s *At the Command prompt, enter subtract.*	4.5.1t *Geben Sie in der Befehlszeile differenz ein.*
4.5.2s *At the command prompt, enter subtract.*	4.5.2t *Geben Sie an der Eingabeaufforderung DIFFERENZ ein.*

4.6s *Binding ISO Elements to XML Elements using Object Info*	4.6.1t オブジェクト情報を使用して ISO エレメントを XML エレメントに拘束する
4.6s *Binding ISO Elements to XML Elements using Object Info*	4.6.2t オブジェクト情報を使用して ISO エレメントを XML エレメントに関連付ける

Question 6, example D: Target text explicitation

4.7s *Selecting*	4.7.1t エレメントの選択
4.7s *Selecting*	4.7.2t コールアウトエレメントの選択
4.7s *Selecting*	4.7.3t アセンブリの選択
4.7s *Selecting*	4.7.4t 多角形の選択
4.7s *Selecting*	4.7.5t 線の選択
4.7s *Selecting*	4.7.6t 楕円の選択
4.7s *Selecting*	4.7.7t 選択
4.7s *Selecting*	4.7.8t 長方形の選択
4.7s *Selecting*	4.7.9t ベジエ曲線の選択
4.7s *Selecting*	4.7.10t めねじの選択
4.7s *Selecting*	4.7.11t おねじの選択

7. How could TM tools be improved?

The first two questions were about the interviewee's background. Questions three and four were intended to be quite broad, seeking the interviewee's opinion of the benefits and disadvantages of TM, leading to the effect of TM on consistency and whether they felt consistency was important, in addition to their own experiences of consistency issues. Kvale and Brinkmann suggest that such broad introductory questions may yield spontaneous and rich descriptions of the

interviewee's experience of the phenomena investigated (2009, p135). Question five related to ST inconsistencies such as those found in categories 1 and 2 in section 4.4. Question six concerned specific types of inconsistency:

- TT noun or term inconsistency - the largest category of sub-segment inconsistency found in the study, also the prevalence of Anglicisation in the target language, and suggestions to improve terminological consistency
- Inconsistencies of ST format retention in the TT
- Alternation of whole phrases throughout the TM
- Explicitation in Japanese TT

Finally, question seven gave the interviewee an opportunity to suggest ways of improving how TM tools deal with consistency issues in future.

The interviewer practised active listening throughout the interview (Kvale and Brinkmann 2009, p138), consciously analysing replies and offering affirmation in order to create a rapport with the interviewee. In the case of an ambiguous response from the interviewee, the interviewer endeavoured to interpret the statement to the interviewee's satisfaction, so that incorrect interpretations could be ruled out. Remaining ambiguities were cleared up by email contact with the interviewee at the analysis stage. At the end of the interviews, interviewees were asked if they had any other comments or suggestions, then offered transcripts of the interview, not only as a matter of courtesy, but also in case they wanted to change or add to an answer. Only one interviewee asked for a transcript and no interviewees requested changes.

4.7.4 Transcription

Interviews took place between December 9th 2011 and February 14th 2012 and recordings were transferred to a password-protected PC in Dublin City University as stipulated by the University Ethics Committee. Interview data was then

transcribed to a document following playback with VLC (version 2.0.0) media player software[18] or via direct playback from the iPhone. Transcription is a significant stage in processing interview data, transforming the narrative mode from oral to written discourse and de-contextualising the interview conversation (Kvale and Brinkmann 2009, p178). This step necessarily involves interpretation of meaning and associated choices, such as where to place punctuation, that can substantially change the content. This again presents a threat to reliability and validity, however "there is no true, objective transformation from the oral to the written mode" (*ibid*. p186). This was a threat we were sensitive to in the transcription process and as a result ambiguous sections were listened to repeatedly in the hope of retaining the meaning and intent of the speech as expressed by the interviewee.

As the process is also time consuming (in this case one hour of interview required eight hours of transcription) some decisions on transcription had to be made before the process began. The transcriptions were largely verbatim, but omitted pauses shorter than 4 seconds (longer pauses were represented with […]) and coughs. Laughs were included as they were considered to affect the tone of the content. Inaudible sections were marked by [???]. Confidential details such as names that we had been requested to omit were substituted with {brackets} and words that were emphasised were italicised in transcription. Transcriptions of interviews can be found in Appendix D in the print version of this thesis.

4.7.5 Coding

Coding is a key element of interview analysis and can mean assignment of a keyword to a section of transcribed text or categorisation of a section of that text (Kvale and Brinkmann 2009, p202). In this study, we took the latter

[18] An open source media player tool available from www.videolan.org/vlc.

interpretation, applying top-down or concept-driven coding, in that the responses were categorised by questions which were based on the five prescribed themes identified prior to the interviews: general opinions on TM, opinions of inconsistency, ST inconsistency, TT inconsistency, and the future of TM. For sections that digressed from the initial themes we applied bottom-up or data-driven coding, allowing the interview data to set the theme. In such cases, sections were labelled according to their topics (Richards 2005, p88).

Coding was done using NVivo 9 qualitative analysis software[19] in several steps: Transcripts were first imported into NVivo, these were then coded by interviewee, by question, and finally by themes that emerged over the course of the interviews. Each interview was assigned attributes signifying the interviewee's job, gender, first language, and main TM tool so that queries could be refined using these attributes.

In attempting to glean data-driven themes from the interview material, we used the method that Kvale and Brinkmann term 'bricolage', reading through the interviews to get an overall impression, to generate meaning, and to capture key understanding (*ibid*. p234). As is typical when coding with Nvivo, emergent themes were gathered as free or open codes (or nodes, to use the Nvivo terminology). These open codes were then sorted into a hierarchy of branching "tree nodes" to reflect the "structure of the data" (Bazeley 2007, p100). Aside from adding organisation to the open codes, the sorting stage is also said to prompt the user to code thoroughly, to improve conceptual clarity, and to help to identify patterns and connections (*ibid*. p104).

In coding of interviews, some comments were subdivided into positive or negative opinions expressed. This system was used to draw some quantitative data from the interviews. Moreover, this coding of responses along with

[19] Available from www.qsrinternational.com.

meaning condensation, and abridgement of responses to a central theme, made those responses easier to compare. The NVivo software also permitted word queries and cross-referencing of codes, to see if emergent themes were more common among a particular section of the cohort of interviewees.

4.8 Summary

This chapter set out the methodology used to measure and categorise inconsistencies in TM data, along with the qualitative process used to verify and enrich the quantitative findings. It began with a reiteration of the research questions pursued in this study, followed by an overview of the mixed methods research paradigm and how that paradigm was applied in the current research. Section 4.4 contained operationalisation of key concepts in this study such as 'inconsistency', 'segment level inconsistency', and 'sub-segment level inconsistency'. Thereafter followed details of the data used in this study, how it was procured and processed prior to analysis. The final sections concerned the two phases of the mixed methods study, explaining the methodology for each stage along with potential threats to validity and reliability of the research.

5 Quantitative Phase: Data Analysis

5.1 Introduction

In this chapter analysis of each TM is presented in turn, subdivided into the four categories of repeated segments. The first TU (sorted alphabetically by source segment) that contains a segment repeated in another TU (or in the case of category 1, similar to one in another TU) is considered the reference TU. Each subsequent TU containing the repeated segment and that contains an aligned segment that is not consistent with the reference is considered a segment-level inconsistency. Thereafter, inconsistencies within the segments are measured, as per section 4.4. Section 5.6 briefly compares the results of this study with the results of a QA check using current commercial QA tools, following which the main findings are summarised. The findings from this chapter are explored further in the qualitative interviews and discussion in Chapter 6.

5.2 TM A

As profiled in section 4.5.2, TM A is a corpus of English-to-German software documentation containing 22,691 TUs of aligned segments. Eligible TUs (22,503) in TM A were categorised as specified in section 4.4. In this section we provide the results of our data analysis for each category in turn.

5.2.1 Category 1 TUs (Inconsistent ST → Inconsistent TT)

Category 1 ST		Category 1 TT	
Letter case	60	Noun	84
Punctuation	48	Verb	45
Tags	42	Word Order	41
Space	20	Adverb	21
Typo	4	Article	15
		Preposition	12
		Space	6
		Others	16
		Completely rewritten	12
Total Segments	370	Total Segments	370
Segment level Inconsistencies	174	Segment level Inconsistencies	174
Total sub-segment inconsistencies	174	Total sub-segment inconsistencies	240

Table 5.1 TM A Category 1 Inconsistencies

370 TUs contained minor ST inconsistencies (as specified in section 4.4) and inconsistent translations. Table 5.1 gives an overview of these TUs, broken down according to the kind of inconsistency evidenced on the source and target sides.

ST Inconsistencies

Figure 5.1 gives an overview of category 1 sub-segment inconsistencies, broken down according to the ST feature that causes segment-level inconsistency.

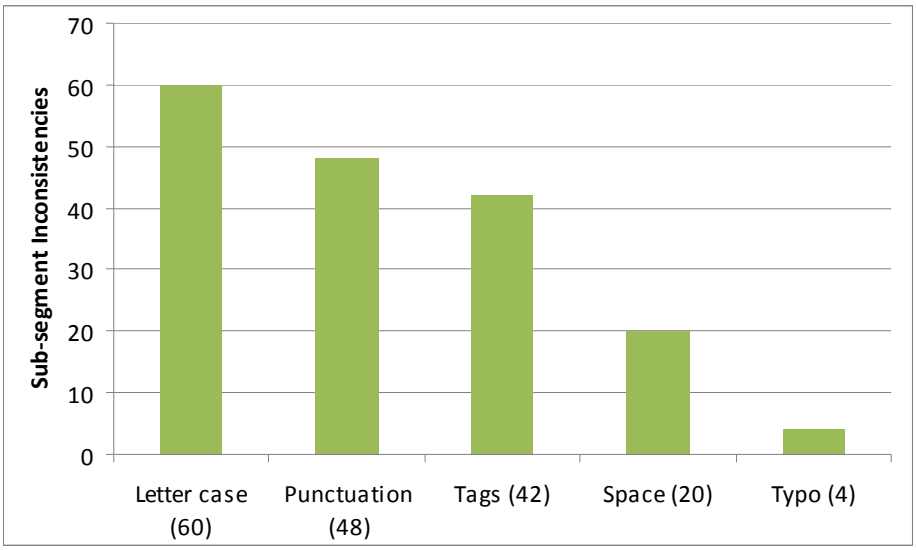


Figure 5.1 TM A Category 1 ST Inconsistencies

Within the 370 category 1 TUs we found 174 ST inconsistencies. One sub-segment inconsistency was found for each ST segment, so the number of segment level and sub-segment inconsistencies is equal. These ST inconsistencies were mostly in letter case, as in **example 5.1** (inconsistencies highlighted):

5.1.1s *Add the Macro Created in Exercise 1 to a Toolbar*	5.1.1t *Makro aus Übung 1 in eine Symbolleiste aufnehmen*
5.1.2s *Add the macro created in exercise 1 to a toolbar*	5.1.2t *In Übung 1 erzeugtes Makro in eine Symbolleiste aufnehmen*

Within the 174 segment level inconsistencies, 60 (34%) contained inconsistent letter case, 48 (28%) contained inconsistent punctuation, 42 (24%) contained tags in one instance that were replaced by quotation marks in another, 20 (11%)

contained inconsistent numbering or placement of tags, 20 (11%) contained inconsistent spacing between words or tags (as in example 5.5 on p106; nine of these had an extra white space at the end of the ST segment), and 4 (2%) contained typographical errors in the ST as per **example 5.2**:

5.2.1s *As noted earilier, you do not use {1}{2}Edit {3}{4}Preferences{5}{6} in {7}In I's{8} to set preferences for JT import.*	5.2.1t *Wie bereits erwähnt, verwenden Sie in {7}{8} nicht {1}{2}Bearbeiten{3}{4}Vorgaben{5}{6}, um Voreinstellungen für den JT-Import festzulegen.*
5.2.2s *As noted earlier, you do not use {1}{2}Edit {3}{4}Preferences{5}{6} in {7}In I's{8} to set preferences for JT import.*	5.2.2t *Wie bereits erwähnt, verwenden Sie nicht {1}{2}Bearbeiten {3}{4}Vorgaben{5}{6}, um in {7}In I's{8} Vorgaben für den JT-Import festzulegen.*

TT inconsistencies

Figure 5.2 gives an overview of category 1 TT sub-segment inconsistencies.

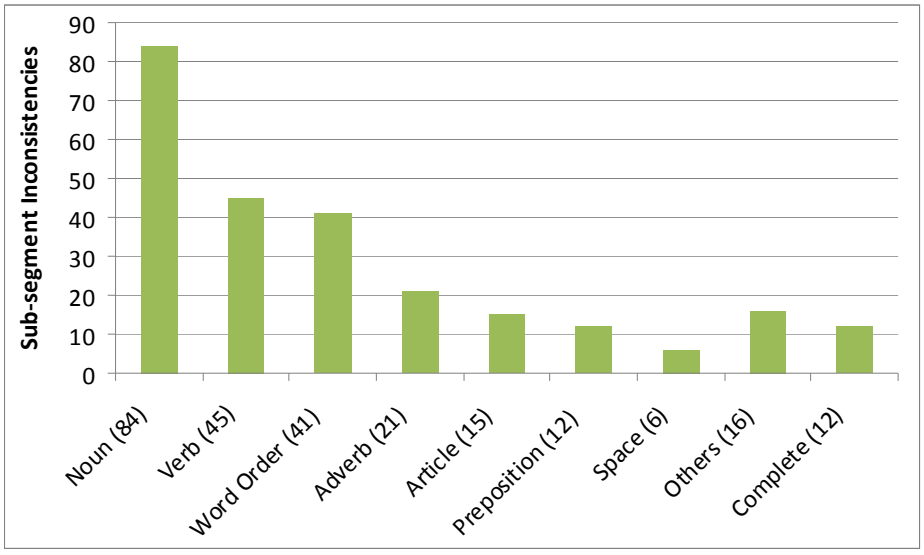


Figure 5.2 TM A Category 1 TT Inconsistencies

Some of the inconsistent TT segments contained more than one inconsistency (as can be seen in this section at example 5.13 on p112). These sub-segment inconsistencies (as operationalised in section 4.4) were counted up to a maximum of three, beyond which we considered that the TT segment had been completely rewritten. Within the 370 category 1 TUs, 240 sub-segment inconsistencies were found along with a further 12 TT segments that had been changed completely. The following **example 5.3** contains a completely rewritten TT segment:

5.3.1s *Each time you click {1}OK{2} in the {3}3D projection{4} window, a new file window will be created containing a "snapshot" of the required situation.*	5.3.1t *Jedes Mal, wenn Sie im Fenster {3}3D Projektion{4} auf {1}OK{2} klicken, wird ein neues Dateifenster mit einem "Schnappschuss" der gewünschten Situation erstellt.*
5.3.2s *Each time you click {1}OK{2} in the {3}3D projection{4} window, a new file window will be created containing a snapshot of the required situation.*	5.3.2t *Jeder Klick auf {1}OK{2} im Dialogfenster {3}3D-Projektion{4} erzeugt ein neues Dateifenster mit einem "Foto" der gewünschten Situation.*

The most commonly occurring TT inconsistency was inconsistently translated nouns, of which there were 84 (35% of the total). An example of noun inconsistency can be seen in **example 5.4,** in which a letter is incorrectly left out in segment 5.4.1t:

5.4.1s *as wireframe*	5.4.1t *Als Drahtmodel*
5.4.2s *As Wireframe*	5.4.2t *Als Drahtmodell*

Of these 84, 14 showed the influence of the source language in one instance, as per **example 5.5** (the ST inconsistency is spacing between tags 3 and 4):

5.5.1s *Help Center includes both web-based HTML and PDF versions of the information and is accessed from the {1}{2}Help{3} {4}Help Center{5}{6} menu.*	5.5.1t *Das Help Center enthält webbasierte HTML- und PDF-Versionen der Informationen und kann über das Menü{1}{2}Hilfe{3} {4}Help Center{5}{6} aufgerufen werden.*
5.5.2s *Help Center includes both web-based HTML and PDF versions of the information and is accessed from the {1}{2}Help{3}{4}Help Center{5}{6} menu.*	5.5.2t *Das Hilfe-Center beinhaltet webbasierte HTML- und PDF-Versionen der Informationen und kann über das Menü{1}{2}Hilfe{3}{4}Help Center{5}{6} aufgerufen werden.*

In two of the 84 noun inconsistencies there was an inconsistency between the singular and plural forms of the noun.

45 cases (19% of the total) of inconsistently translated verbs were found in the TT, as per example 5.5 previously, which contains the verb 'to include' translated as 'enthält' in one case and 'beinhaltet' in another. Of these 45 verb inconsistencies, five contained an active verb in one instance (5.6.1t) and a passive verb in another (5.6.2t), as in **example 5.6**:

5.6.1s *The drawing must be depicted in {1}HLR{2} display mode to do this.*	5.6.1t *Die Zeichnung muss dazu im Darstellungsmodus {1}Ohne verdeckte Kanten{2} vorliegen.*
5.6.2s *The drawing must be depicted in "{1}HLR{2}" display mode to do this.*	5.6.2t *Die Zeichnung muss dazu im Darstellungsmodus {1}Ohne verdeckte Kanten{2} angezeigt werden.*

The TT word order was changed in 41 cases (17%), although the same words were chosen. In a change unrelated to the ST letter case inconsistency, this can be seen in **example 5.7**:

5.7.1s *Working Outside the Major Axes*	5.7.1t *Außerhalb der Hauptachsen arbeiten*
5.7.2s *Working outside the major axes*	5.7.2t *Arbeiten außerhalb der Hauptachsen*

5.2.2 Category 2 TUs (Inconsistent ST → Consistent TT)

613 TUs fell into category 2, where the translation process has introduced consistency. One sub-segment inconsistency was found for each ST segment, so the number of segment level and sub-segment inconsistencies is equal. Table 5.2 gives an overview of the ST inconsistencies to be found in these TUs.

Category 2 ST	
Letter case	140
Space	95
Punctuation	67
Word (Noun, verb, adverb, article)	11 (3, 1, 3, 1)
Typo	2
Word Order	1
Segment-level Inconsistencies	316
Total Segments	613

Table 5.2 TM A Category 2 Inconsistencies

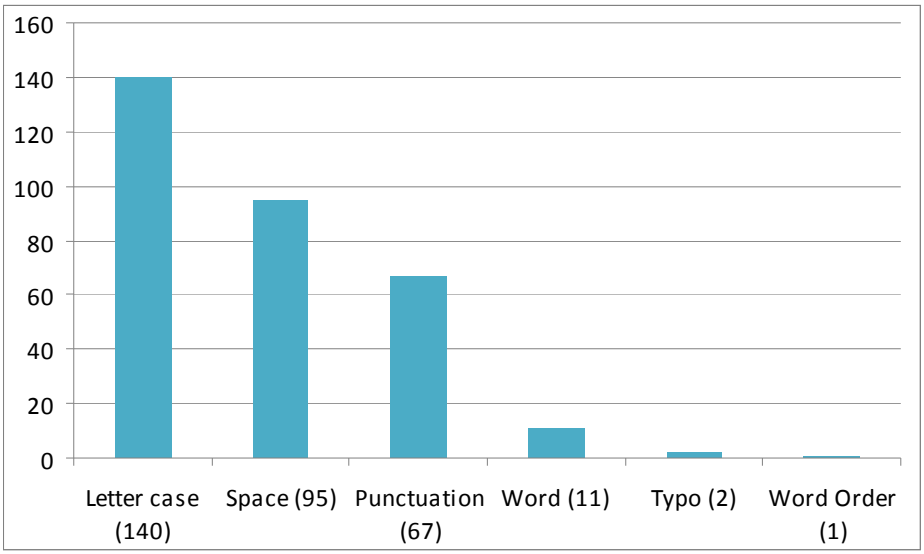


Figure 5.3 TM A Category 2 Inconsistencies

In TM A, as with all of the other TM data in this study, the largest proportion of category 2 inconsistency was letter case inconsistency. We found 140 (44% of the total) letter case inconsistencies, such as in **example 5.8**:

5.8.1s *dimensions bar*	5.8t *Maßleiste*
5.8.2s *Dimensions bar*	5.8t *Maßleiste*

Between categories 1 and 2 there are 200 letter case inconsistencies in the ST. 140 of those (in category 2) are associated with consistent TT. In our tests using Omega T, matches containing letter case inconsistent with our translatable segment are suggested as 100% or 99% matches, so this introduced consistency is unsurprising. Nonetheless, 60 inconsistencies (those in category 1; 30% of the ST letter case inconsistencies in TM A) are associated with TT segments containing non-letter case inconsistency.

Category 2 contains 95 space inconsistencies (30%). Of these, 70 contain one trailing white space at the end of the segment. Nine category 1 trailing space inconsistencies are associated with inconsistent TT segments. This means that in

the whole of TM A, nine out of 79 TUs (or 11%) with a space inconsistency at the end are associated with inconsistent TT segments.

Category 2 TUs contain 67 punctuation inconsistencies (21%) as in **example 5.9**:

5.9.1s *All element parts which do not lie within the area of the mask are "hidden".*	5.9t *Alle Elementteile, die nicht innerhalb der Maskenfläche liegen, werden ausgeblendet.*
5.9.2s *All element parts which do not lie within the area of the mask are hidden.*	5.9t *Alle Elementteile, die nicht innerhalb der Maskenfläche liegen, werden ausgeblendet.*

22 category 2 TUs (7%) contained lexical differences in the ST. **Example 5.10** contains a lexical change, although the different nouns are semantically similar and thus 5.10.1s and 5.10.2s were translated in the same way (5.10t).

5.10.1s *If you have selected {1}Find structures in data{2} a check mark indicates all assemblies and/or elements in the file.*	5.10t *Wenn Sie {1}Strukturen in Daten finden{2} ausgewählt haben, sind alle Baugruppen und/oder Elemente in der Datei mit einem Häkchen versehen.*
5.10.2s *If you have selected {1}Find structures in data{2} a tick indicates all assemblies and/or elements in the file.*	5.10t *Wenn Sie {1}Strukturen in Daten finden{2} ausgewählt haben, sind alle Baugruppen und/oder Elemente in der Datei mit einem Häkchen versehen.*

5.2.3 Category 3 TUs (Consistent ST → Inconsistent TT)

As explained in section 4.4, TUs in category 3 contain consistent ST segments that are translated as inconsistent TT segments. These are the TT inconsistencies we searched for in the pilot study (chapter 3) and they were the initial focus of this study. The consistency of the ST segments should have had the consequence that, once the first ST segment had been translated, the translator would have received a 100% match for repetitions of that segment and the software should have suggested the existing translation to the translator. In TM A from 390 TUs, 137 inconsistencies were found at the segment level. 174 inconsistencies were found at the sub-segment level and three segments were considered to have been completely rewritten. Table 5.3 gives an overview of these inconsistencies.

Category 3 TT	
Noun	81
Verb	34
Adverb	17
Punctuation	12
Preposition	7
Article	6
Tags	5
Word Order	5
Others	7
Completely rewritten	3
Total Segments	390
Segment Level Inconsistencies	137
Total Sub-segment Inconsistencies	174

Table 5.3 TM A Category 3 Inconsistencies

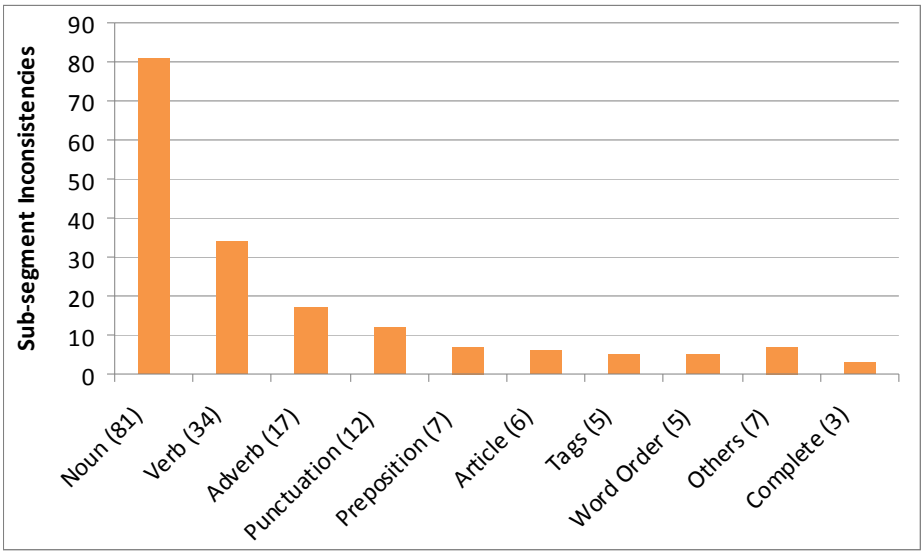


Figure 5.4 TM A Category 3 Inconsistencies

As in category 1, the most prevalent category of TT inconsistency was noun inconsistency. There are 81 inconsistently translated nouns (47% of the inconsistencies) of which 18 show influence from the English source language in one instance as in **example 5.11**:

5.11s *All lines that have been converted using the {1}Create surface borders{2} function can be recognized easily since they are drawn with the {3}Border{4} pen.*	5.11.1t *Alle Linien, die mit der Funktion {1}Flächenränder anlegen{2} konvertiert wurden, können Sie leicht erkennen, da sie mit dem Stift mit der Bezeichnung {3}Border{4} gezeichnet werden.*
5.11s *All lines that have been converted using the {1}Create surface borders{2} function can be recognized easily since they are drawn with the {3}Border{4} pen.*	5.11.2t *Alle Linien, die mit der Funktion {1} Flächenränder anlegen{2} konvertiert wurden, können Sie leicht erkennen, da sie mit dem Stift mit der Bezeichnung {3}Rand{4} gezeichnet werden.*

'Border' is used three times in TM A to translate 'border'. In all other instances 'Rand' is chosen as the translation for the English word 'border'. Looking at the

metadata, all six inconsistent segments (three each containing 'Border' and 'Rand' as translations of the same ST segment) were saved on February 20th 2009. The three containing 'Border' were saved initially at 16.01 and changed at 19.54, one minute after the segments with 'Rand' were saved. It may be that the translator chose to edit the suggested match from the TM to use the term 'Rand' and settings in the TM tool were such that revised segments were added to the TM rather than over-writing TUs already in the TM (see interviewees opinions on whether to overwrite or append new translations to a TM in section 6.5.2). A further seven TT nouns were singular in one instance and pluralised in another.

34 cases (20% of the inconsistencies found) of inconsistent TT verbs were found, of which 11 varied between active and passive as had been seen previously in example 5.5 (on p106), and in five cases a word had been translated as a verb in one instance (5.12.2t) and nominalised in another (5.12.1t) as in **example 5.12**:

5.12s *Export*	5.12.1t *Export*
5.12s *Export*	5.12.2t *Exportieren*

Unlike in category 1, the next most prevalent category of inconsistency was adverbs. 17 inconsistently translated adverbs fell into category 3 as per **example 5.13**, which contains the adverbs ‘hier’ (here) and ‘wieder’ (again):

5.13s *Now release the mouse button.*	5.13.1t *Hier lassen Sie die Maustaste los.*
5.13s *Now release the mouse button.*	5.13.2t *Lassen Sie die Maustaste wieder los.*

As previously mentioned, the numbers of inconsistencies at segment level and sub-segment level differed, as TT segments that were inconsistent at segment level occasionally differed in more than one respect. This can be seen in **example**

5.14, which contains examples of inconsistent adverbs, punctuation, articles, and verbs:

5.14s *The following dialog box appears.*	5.14.1t *Das folgende Dialogfenster erscheint:*
5.14s *The following dialog box appears.*	5.14.2t *Es erscheint ein Dialogfenster.*
5.14s *The following dialog box appears.*	5.14.3t *Ein Dialogfenster wird angezeigt.*

The ST segment 5.14s was repeated exactly in each TU. We have assigned segment 5.14.1t the status of reference translation. In segment 5.14.2t, the adverb 'folgende' has been removed, the article 'das' has been replaced by the indefinite article 'ein', the verb has moved from the end of the segment, and the colon at the end has been replaced by a full stop. As this segment contains more than three inconsistencies, it is considered to be completely rewritten, although some of the vocabulary remains the same. Other segments that we deemed completely rewritten differed from the 'reference translation' in a similar way. In segment 5.14.3t, the article has become an indefinite article, the adverb has been removed, and the verb has changed lexically and in voice from active to passive.

While there are similarities between TT segments in the following **example 5.15**, it seems time-consuming and impractical for the translator to make so many adjustments should he or she have a match (such as the segment we have assigned as reference translation) suggested:

5.15s *You can switch thumbnail image generation on or off for an ISO file using the {1}Generate preview{2} option on the {3}Preferences: misc{4} page.*	5.15.1t *Die Erstellung von Miniaturansichtsbildern für eine ISO-Datei kann über die Option {1}Vorschau erzeugen{2} auf der Seite {3}Vorgaben: diverse{4} aktiviert oder deaktiviert werden.*
5.15s *You can switch thumbnail image generation on or off for an ISO file using the {1}Generate preview{2} option on the {3}Preferences: misc{4} page.*	5.15.2t *Sie können die Erstellung von Miniaturansichten für eine ISO-Datei ein- oder ausschalten, indem Sie die Option {1}Generate preview{2} auf der Seite {3}Preferences: misc{4} wählen.*

The reasons for these changes become no clearer despite searching through the metadata. The <tu> tag in the TMX file tells us that the same translator ID was recorded for both TT segments. The TT segment 5.15.1t was confirmed and saved to the TM on June 16th at 15.25, and the segment 5.15.2t was saved to the TM on July 1st at 07.23.

5.2.4 Category 4 TUs (Consistent ST → Consistent TT)

The remaining TUs in this TM data (98.2% of the 22,503 TUs) either contain ST segments that are not repeated or fall into category 4, containing consistent ST and TT segments. In total 6674 TUs contain ST segments that were repeated at least once and were translated consistently. If we add the total number of category 3 TUs (390 with consistent ST and inconsistent TT) to the number of category 4 TUs (6674), we can see that the total number of TUs with repeated ST is 7064. Of these, 390 (or 5.5%) contained inconsistency that was introduced in the TM translation process.

5.3 TM B

TM B is an English-to-Japanese TM corpus from the same company and translation project as TM A containing 18,799 TUs of aligned segments. The TUs in TM B were categorised and analysis of each category in turn is presented in this section.

5.3.1 Category 1 TUs (Inconsistent ST → Inconsistent TT)

65 TUs contained minor ST inconsistencies as specified in section 4.4 and inconsistencies in the corresponding TT, and were thus category 1 TUs. Table 5.4 gives an overview of these TUs, broken down according to source and target inconsistencies.

Category 1 ST		Category 1 TT	
Letter case	13	Noun	15
Tags	10	Tags	7
Space	8	Punctuation	5
Punctuation	3	Verb	3
Typo	1	Particle	1
		Typo	1
		Word Order	1
		Completely rewritten	5
Total Segments	65	Total Segments	65
Segment-level Inconsistencies	32	Segment-level Inconsistencies	32
Total Sub-segment Inconsistencies	35	Total Sub-segment Inconsistencies	33

Table 5.4 TM B Category 1 Inconsistencies

ST Inconsistencies

Figure 5.5 gives an overview of category 1 ST inconsistencies, broken down according to sub-segment inconsistency.

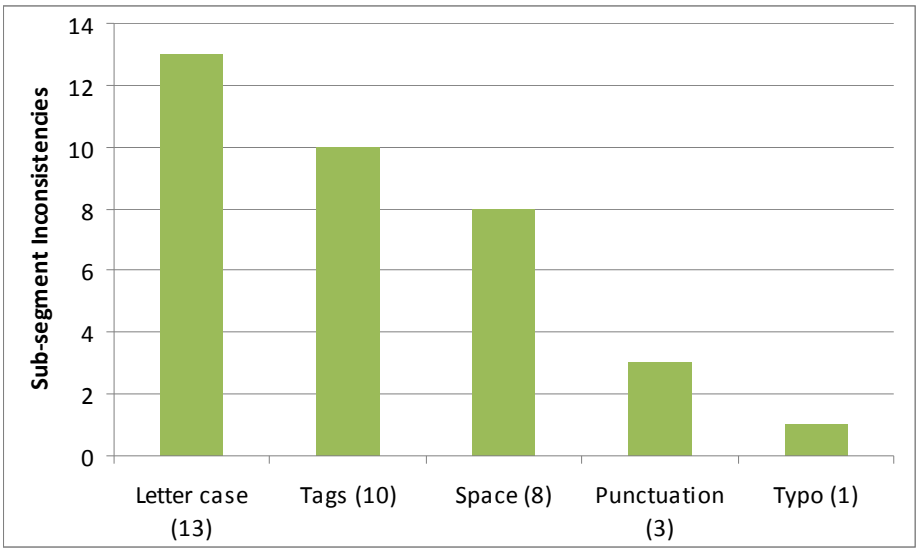


Figure 5.5 TM B Category 1 ST Inconsistencies

As with TM A, the most commonly occurring category of ST inconsistency is inconsistent letter case, as in **example 5.16**:

5.16.1s *Find the Appropriate Ellipse for the Axis*	5.16.1t 軸に対して適切な楕円の検索[20]
5.16.2s *Find the appropriate ellipse for the axis*	5.16.2t 特定の軸に適合する楕円[21]

[20] 'Jiku ni taishite tekisetsuna daen no kensaku'. Translated: 'Search for appropriate ellipse for the corresponding axis' (our translation).

[21] 'Tokutei no jiku ni tekigou suru daen'. Translated: 'Ellipse that is compatible for a particular axis' (our translation).

The only ST change was in capitalisation, yet the TT segments appear to have been completely rewritten, aside from the same term having been used for axis - 軸 or jiku − and for ellipse − 楕円 or daen. From the metadata for these TUs it is clear that the ST segments were translated by different translators. Segment 5.16.1t was saved to the TM on December 22nd 2008 yet not used in the subsequent translation of 5.16.2t on March 24th 2009 in which the imperative ST verb 'find' was not translated.

Within the 32 category 1 TUs there are 35 sub-segment inconsistencies: 13 ST segments (37%) contained inconsistent letter case, two of those also contained inconsistent placement of tags. 10 ST segments (29%) contained inconsistent tags, eight (23%) contained inconsistent spacing (four of these with an extra white space at the end of the ST segment), three differed in punctuation, and one contained a typographical error. The latter was the same typographical error as in example 5.2 in TM A. See **example 5.17**:

5.17.1s *As noted earilier, you do not use {1}{2}Edit {3}{4}Preferences{5}{6} in {7}In I's{8} to set preferences for JT import.*	5.17.1t 前述のとおり、JT インポートの環境設定では、{1}{2} の{3}{4}「編集」{5}{6}「環境設定…」{7}{8}メニューは使用しません。[22]
5.17.2s *As noted earlier, you do not use {1}{2}Edit {3}{4}Preferences{5}{6} in {7}In I's{8} to set preferences for JT import.*	5.17.2t 前述のとおり、JT インポートの環境設定では、{7}{8} の{1}{2}「編集」{3}{4}「環境設定…」{5}{6}メニューは使用しません。

In the Japanese example above, there has been no formal change in the text. The only difference between 5.17.1t and 5.17.2t is in re-ordering of tag positions. The tag order is identical in both ST segments.

[22] 'Zenjutsu no toori, JT inpooto no kankyousettei de wa,{7}{8} no {1}{2}「henshuu」{3}{4}「kankyousettei...」{5}{6}menyuu wa shiyou shimasen'.

TT Inconsistencies

Within the 65 category 1 TUs, we found 33 TT sub-segment inconsistencies, an overview of which is in figure 5.6.

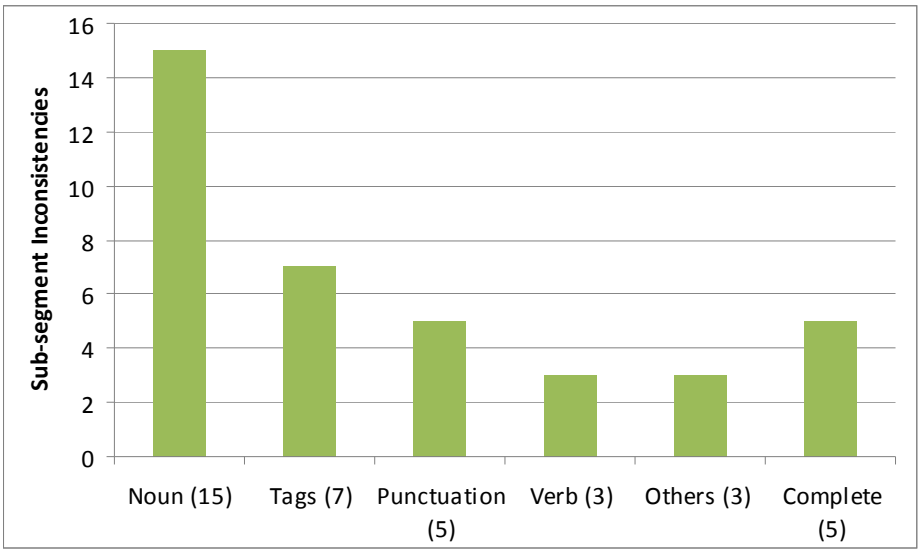


Figure 5.6 TM B Category 1 TT Inconsistencies

15 category 1 TT inconsistencies (45%) are noun inconsistencies as in **example 5.18**:

5.18.1s *basic settings*	5.18.1t 設定の基礎[23]
5.18.2s *Basic Settings*	5.18.2t 基本設定[24]

Five of the 15 noun inconsistencies showed the influence of the source language in one instance, as per **example 5.19**:

[23] 'Settei no kiso' : 'Basic settings' (Literally 'Basics of settings'. Our translation).

[24] 'Kihon settei' : 'Basic settings'. (Our translation).

5.19.1s *Format*	5.19.1t 形式
5.19.2s *format*	5.19.2t フォーマット

Segment 5.19.1t contains the word 形式 or keishiki, a Japanese word used for file format. 5.19.2t on the other hand uses フォーマット or foomatto, a word borrowed from English and written in the katakana alphabet used for loan words or the transcription of words from foreign alphabets into Japanese.

Other TT inconsistencies in category 1 were seven cases of inconsistent tags (as seen in example 5.17), five inconsistencies in punctuation or formatting, three verb inconsistencies, and five completely rewritten TT segments (as in example 5.16).

5.3.2 Category 2 TUs (Inconsistent ST → Consistent TT)

914 TUs fell into category 2, containing inconsistent ST segments associated with consistent TT. Table 5.5 gives an overview of the inconsistencies in these TUs.

Category 2 ST	
Letter case	219
Word (Noun, verb, adverb, article)	98 (86, 6, 4, 2)
Space	96
Punctuation	20
Others	10
Whole ST segment inconsistent	7
Total Segments	914
Segment Level Inconsistencies	450
Total Sub-segment Inconsistencies	443

Table 5.5 TM B Category 2 Inconsistencies

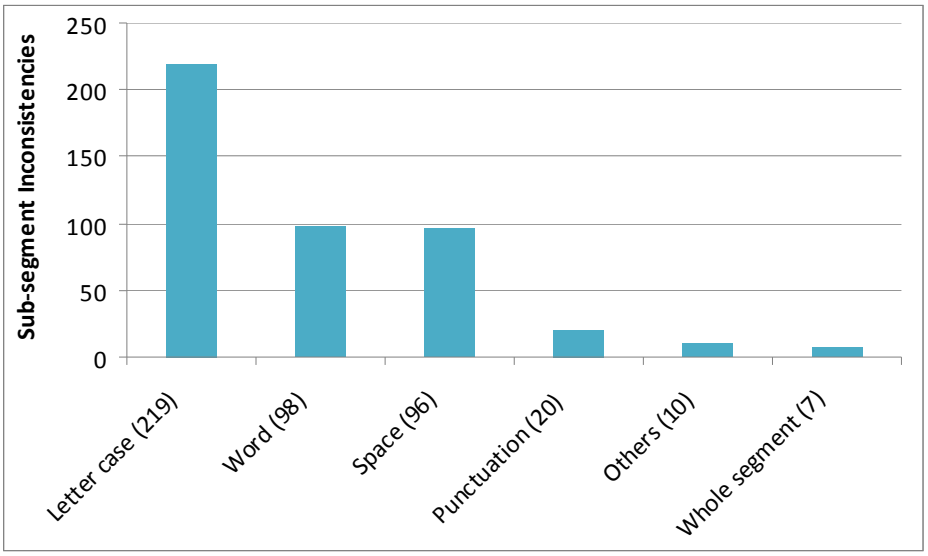


Figure 5.7 TM B Category 2 Inconsistencies

In TM A only 613 TUs fell into category 2, fewer than the 914 from TM B. There are a number of reasons why ST inconsistencies, such as those caused by inconsistent letter case or the presence of plural instead of singular in the ST, may be correctly ignored by an English to Japanese translator who chooses instead to accept a fuzzy match. As in the previous TM, a high percentage (49%) of category 2 inconsistencies is attributable to letter case inconsistencies. As Japanese characters do not vary by letter case, this type of ST inconsistency cannot be carried across into Japanese. **Example 5.20** shows an instant of inconsistent letter case being eliminated in Japanese:

5.20.1s *Case sensitive*	5.20t 大文字と小文字の区別[25]
5.20.2s *Case Sensitive*	5.20t 大文字と小文字の区別

Unlike in German, there is no distinction between singular and plural form in Japanese – numbers are given explicitly or are implicit in context. Of 86 cases of inconsistent nouns in the ST segments of category 2 TUs (19% of the total), 44

[25] 'Ōmoji to komoji no kubetsu' [The distinction between upper case and lower case letters]

differ in number: singular in one case, plural in another. These differences are eliminated in the Japanese TT, as in **example 5.21**:

5.21.1s *Dimension*	5.21t 寸法[26]
5.21.2s *Dimensions*	5.21t 寸法

In another instance, the words 'place' and 'placing' were both translated as 割り付け 'waritsuke' [allocation].

There are 96 space inconsistencies in the category 2 TUs (22% of the total) of which 81 contain a trailing space at the end of ST segments. In total, 85 ST segments in TM B contain an inconsistent trailing space. Of those 85, four (the white space inconsistencies from category 1; 5%) were aligned with inconsistent TT. 20 ST inconsistencies (5%) contain inconsistent punctuation. The following example contains an inconsistent comma in segment 5.22.2s and also contains an explicitation at the end of the same segment.

[26] 'Sunpou'.

5.22.1s *In the case of a placed drawing/illustration however, clicking {1}OK{2} in the {3}3D projection{4} dialog box displays the projection again, in its entirety, in the {5}{6} window.*	5.22t ただし、割り付けられたドローイングまたはイラストレーションの場合、{1}「3D プロジェクション」{2}ダイアログボックスの{3}「OK」{4}をクリックすると、開かれているファイルの{5}{6} ウィンドウに投影が再び完全な形で表示されます。[27]
5.22.2s *In the case of a placed drawing/illustration, however, clicking {1}OK{2} in the {3}3D projection{4} dialog box displays the projection again in its entirety in the {5}{6} window of the opened file.*	5.22t ただし、割り付けられたドローイングまたはイラストレーションの場合、{1}「3D プロジェクション」{2}ダイアログボックスの{3}「OK」{4}をクリックすると、開かれているファイルの {5}{6} ウィンドウに投影が再び完全な形で表示されます。

Six verbs were inconsistent in the English ST, two of which were in the future tense in one instance, present tense in another. Since there is no future tense in Japanese (actions to take place in the future are understood as such based on their context) these verbs were translated as present tense as in example 5.23:

[27] 'Tadashi, waritsukerareta dorooingu mata wa irastoreeshon no baai, {1} “3D purojekushon”{2} daiarogubokusu no {3}”OK”{4} o kurikku suru to hirakareteiru fairu no {5}{6} uindou ni touei ga futatabi kanzenna katachi de hyouji saremasu'.

5.23.1s *The following dialog box appears:*	5.23t 次のダイアログボックスが表示されます。[28]
5.23.2s *The following dialog box is displayed.*	5.23t 次のダイアログボックスが表示されます。
5.23.3s *The following dialog box will appear:*	5.23t 次のダイアログボックスが表示されます。
5.23.4s *The following dialog box will then appear:*	5.23t 次のダイアログボックスが表示されます。

The TT segment 5.23t uses the verb 表示される (hyouji sareru) [to be displayed (present tense passive)] in each TT translation.

5.3.3 Category 3 TUs (Consistent ST → Inconsistent TT)

239 category 3 TUs were found in TM B containing 102 segment level inconsistencies and 129 sub-segment inconsistencies. Seven TT segments were considered to have been completely rewritten. Table 5.6 gives an overview of the category 3 TUs.

[28] 'Tsugi no daiarogu bokkusu ga hyouji saremasu'.

Category 3 TT	
Noun	49
Verb	40
Particle	15
Punctuation	8
Explicitation	6
Word Order	4
Others	7
Completely rewritten	7
Total Segments	239
Segment Level Inconsistencies	102
Total Sub-segment Inconsistencies	129

Table 5.6 TM B Category 3 Inconsistencies

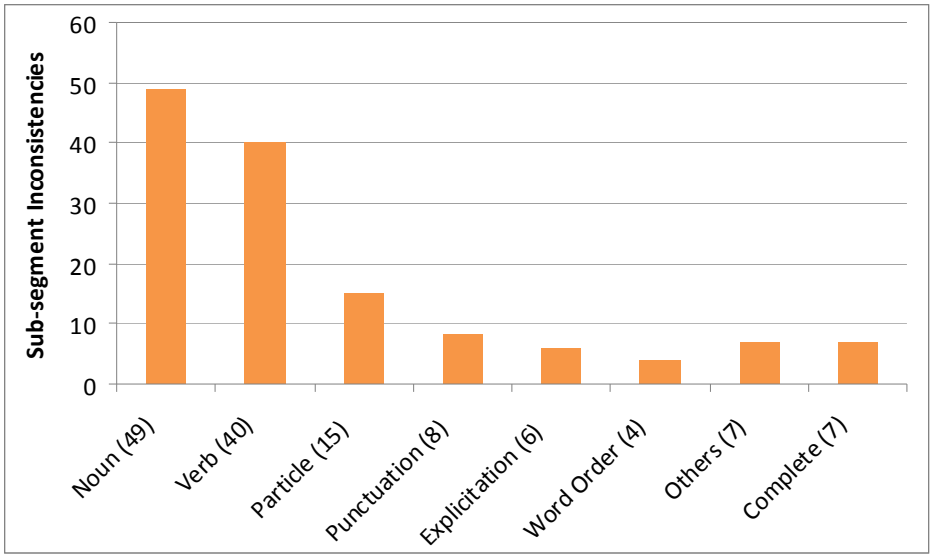


Figure 5.8 TM B Category 3 Inconsistencies

As in TM A, the largest category of inconsistency was noun inconsistency. We found 49 noun inconsistencies (38% of the total). Nine of those showed the

influence of the source language in one instance as shown in example 5.19. Ten inconsistencies involved translations of the word 'selecting', as shown in example 5.24:

5.24s *Selecting*	5.24.1t エレメントの選択
5.24s *Selecting*	5.24.2t コールアウトエレメントの選択
5.24s *Selecting*	5.24.3t アセンブリの選択
5.24s *Selecting*	5.24.4t 多角形の選択
5.24s *Selecting*	5.24.5t 線の選択
5.24s *Selecting*	5.24.6t 楕円の選択
5.24s *Selecting*	5.24.7t 選択
5.24s *Selecting*	5.24.8t 長方形の選択
5.24s *Selecting*	5.24.9t ベジエ曲線の選択
5.24s *Selecting*	5.24.10t めねじの選択
5.24s *Selecting*	5.24.11t おねじの選択

In the example above 5.24.1t is taken as the reference translation as it appeared first in the TM data. Each TT segment contains the noun 選択 (sentaku) meaning 'selection' but most involve further explicitation, using the particle の to make the genitive case. 5.24.1t is エレメントの選択 or 'selection of elements'. Segment 5.24.2t is コールアウトエレメントの選択 or 'selection of callout elements'. They continue with アセンブリの選択 (5.24.3t): 'selection of assembly'; 多角形の選択 (5.24.4t): 'selection of polygon'; 線の選択 (5.24.5t): selection of a line; 楕円の選択 (5.24.6t): 'selection of ellipse'; 選択 (5.24.7t): 'selection'; 長方形の選択 (5.24.8t): 'selection of rectangle'; ベジエ曲線の選択 (5.24.9t): 'selection of Bezier curve'; めねじの選択 (5.24.10t): 'selection of female screw'; おねじの選択 (5.24.11t): 'selection of male screw'. While this explicitation may make the TT segments clear and understandable, it would have

a negative effect on TM leverage. It may be in this case that the first translation contained added detail that was not appropriate for the subsequent translations or that the translators felt that more detail was necessary in the context of the finished document. We see the same pattern for translations of the ST segment 'Generation', translated using the verb 描く (kaku) [to draw]. The TT segments include the thing that is to be drawn, as seen in **example 5.25**:

5.25s *Generation*	5.25.1t イメージエレメントの作成[29]
5.25s *Generation*	5.25.2t 線を描く
5.25s *Generation*	5.25.3t 長方形を描く
5.25s *Generation*	5.25.4t 楕円を描く
5.25s *Generation*	5.25.5t ベジエ曲線を描く
5.25s *Generation*	5.25.6t めねじを描く
5.25s *Generation*	5.25.7t 多角形を描く
5.25s *Generation*	5.25.8t 回転面の作成

After noun inconsistencies, the next most prevalent category was verb inconsistencies. Of the 40 verb inconsistencies (31% of the total), 18 were accounted for by alternation between the verb 拘束する (kousoku suru) [to bind or restrict] in one case, and the verb 関連付ける (kanren tsukeru) [to relate] in another. For **example**:

[29] 5.25.1t: Generation of image element; 5.25.2t: Drawing a line; 5.25.3t: Drawing a rectangle; 5.25.4t: Drawing an ellipse; 5.25.5t: Drawing a Bezier curve; 5.25.6t: Drawing a female screw; 5.25.7t: Drawing a polygon; 5.25.8t: Generation of a surface of revolution. (Our translations).

5.26s *Binding ISO Elements to XML Elements using Object Info*	5.26.1t オブジェクト情報を使用して ISO エレメントを XML エレメントに拘束する
5.26s *Binding ISO Elements to XML Elements using Object Info*	5.26.2t オブジェクト情報を使用して ISO エレメントを XML エレメントに関連付ける

Segment 5.26.1t translates as 'bind ISO elements to XML elements using object information', segment 5.26.2t as 'relate ISO elements to XML elements using object information'. These verbs are alternated in eight other TUs. Looking through the metadata, each verb choice is not attributable to a single user ID as several translators chose one or the other, but the translations using 拘束する were all saved to the TM at the same time on April 22nd 2009. Two uses of 関連付ける were also saved then, but all others were dated from the 7th of May in 2009. However, as both were saved to the TM on April 22nd, both were propagated as 100% matches thereafter.

The category of particle inconsistency, of which we found 15 instances (12% of the total), is particular to the Japanese TMs. A particle or postposition in Japanese signals the role of the word that precedes it. In the following **example 5.27**, the particle alternates between は, marking the sentence topic, and では, meaning 'by means of'.

5.27s *Files in this format cannot be exported by {1}IsoDraw CADproces{2}.*	5.27.1t {1}IsoDraw CADprocess{2} は、このフォーマットのファイルをエクスポートできません。
5.27s *Files in this format cannot be exported by {1}IsoDraw CADproces{2}.*	5.27.2t {1}IsoDraw CADprocess{2} では、このフォーマットでファイルをエクスポートすることはできません。。

Also highlighted in example 5.27 is the change in verb: both mean 'to be unable to', but segment 5.27.2t contains a more complicated version of the same expression.

Examples 5.24 and 5.25 illustrate a tendency in this English to Japanese TM data towards explicitation. The 'explicitation' category was added for English to Japanese data to measure this tendency. Six TUs fell into this category, including **example 5.28**:

5.28s *A tick indicates that the assembly will be imported.*	5.28.1t チェックマークが表示されていれば、そのアセンブリがインポートされます。
5.28s *A tick indicates that the assembly will be imported.*	5.28.2t チェックマークの付いたアセンブリがインポートされます。
5.28s *A tick indicates that the assembly will be imported.*	5.28.3t チェックマークの付いたアセンブリがインポートされます。

The word 'tick' has been translated as チェックマーク [check mark], with 5.28.1t [if the check mark is indicated, the assembly will be imported] differing from the repeated segments 5.28.2t and 5.28.3t [the assembly with the check-mark attached will be imported]. A search through the metadata shows that 5.28.1t was created on March 9th 2009 and both subsequent translations were added to the TM on March 24th 2009, despite, presumably, a 100% match suggestion from the TM.

5.3.4 Category 4 TUs (Consistent ST → Consistent TT)

The remaining TUs in TM B are not repeated or fall into category 4, containing consistent ST and TT segments. 4502 TUs contain ST segments that were repeated at least once. Of these, 239 (5.3%) fell into category 3, leaving 4263

category 4 TUs. The number of repeated ST segments is noticeably lower than in TM A (in which there are 7064 repeated segments), but the rate of inconsistency introduced in the TM translation process (5.5% for TM A, 5.3% for TM B) is very similar. Some consequences of this level of inconsistency are discussed in sections 7.3 and 7.4.

5.4 TM C

As profiled in section 4.5.2, TM C is an English-to-German TM corpus from the same company and translation project as TM D. After removing the TUs that contain only numbers or punctuation, 293,924 TUs remained. In this section we first show that our sample is representative of the data as a whole, then analyse results for each of the categories of TUs.

5.4.1 Testing the Data Sample

Due to the difference in scale between the first set of TMs and the second, we chose a sample of the first 50,283 TUs (sorted alphabetically) for analysis and categorisation as specified in section 4.4. Within this sample, 826 TUs (1.64%) fell into category 3. In the TM as a whole, 4527 TUs (1.54%) fell into category 3. More details can be seen in table 5.7.

TM C	Sample	Total
Number of TUs	50,283	293,924
Category 3 (percentage in brackets)	826 (1.64%)	4527 (1.54%)
Standardised TTR (ST)	14.88	15.75
Standardised TTR (TT)	16.84	17.71
Mean sentence length in ST (words)	16.70	17.26
Mean sentence length in TT (words)	17.15	17.33
Mean ST word length (chars)	4.37	4.68
Mean TT word length (chars)	5.21	5.56

Table 5.7 Comparison of Corpus Analysis Statistics for Sample & Whole of TM C

The standardised type/token ratio (TTR) measures the average number of unique words (types) per block of 1000 words (tokens) in the text. A chi square test,

comparing the sample measured in this study and subsequent 50,000 TU sections (sorted alphabetically) with the whole of TM C, found no significant differences in the number of category 3 TUs[30], standardised TTR [ST: p = 0.65, TT: p = 0.55], mean sentence length, or the mean word length [ST: p = 0.87, TT: p = 0.88]. These findings support the hypothesis that the measurements for the sample shown in table 5.7 do not differ significantly from the whole [p > 0.05].

5.4.2 Category 1 TUs (Inconsistent ST → Inconsistent TT)

995 TUs fell into category 1, with minor ST inconsistencies (as specified in section 4.4) and inconsistent translations. An overview of the ST and TT inconsistencies can be seen in table 5.8.

Category 1 ST		Category 1 TT	
Letter case	314	Noun	323
Punctuation	60	Preposition	138
Tags	46	Word Order	81
Space	37	Verb	66
Typo	17	Punctuation	57
Word Order	1	Space	38
		Others	27
		Completely rewritten	10
Total Segments	995	Total Segments	995
Segment Level Inconsistencies	475	Segment Level Inconsistencies	475
Total Sub-segment Inconsistencies	475	Total Sub-segment Inconsistencies	730

Table 5.8 TM C Category 1 Inconsistencies

[30] A chi test between the sample and the whole for category 3 TUs, applying Yates' Correction, found that $X^2 = 3.85$, where 3.84 equates to a probability (p) of 0.05.

ST Inconsistencies

Figure 5.9 gives an overview of category 1 ST inconsistencies, broken down according to the feature that causes segment-level inconsistency.

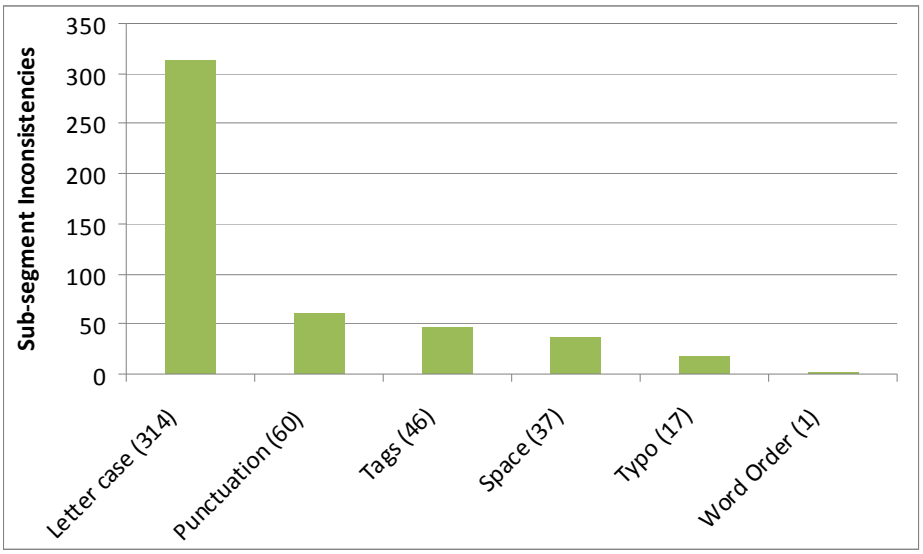


Figure 5.9 TM C Category 1 ST Inconsistencies

As with the previously analysed TMs, the most commonly occurring category of ST inconsistency was letter case. Letter case inconsistencies accounted for 314 (66%) of category 1 ST inconsistencies. Inconsistent use of case is illustrated in **example 5.29**:

5.29.1s *(SHIFT+right-click the drawing area.)*	5.29.1t *(Klicken Sie bei gedrückter UMSCHALTTASTE mit der rechten Maustaste in den Zeichenbereich.)*
5.29.2s *(Shift+right-click the drawing area.)*	5.29.2t *(Klicken Sie bei gedrückter UMSCHALTTASTE mit der rechten Maustaste in den Zeichnungsbereich.)*

In both TT segments the ST word 'shift' has been translated as 'Umschalttaste' (shift key) and capitalised. This would suggest that a TM match was used despite

the change of case in 5.29.2s. However, the German translation of 'drawing area' was changed from 'Zeichenbereich' to 'Zeichnungsbereich'. According to the metadata, segment 5.29.1t was created on December 22nd 2006 and last changed two years later on December 7th 2008. Segment 5.29.2t was created by a different translator on January 15th 2009 and last changed one year later on January 18th 2010. None of the TT segments aligned with ST segments that contain inconsistencies in letter case themselves contain instances of inconsistent letter case; rather the TT segments in question evince other kinds of inconsistencies. Statistical significance tests did not support the suggestion of a causal relationship between ST and TT inconsistencies, or a correlation between the levels of ST and TT inconsistency.

60 ST segments (13%) contained inconsistent punctuation and a further 37 contained inconsistent placement of spaces (8% of the total, 26 of these were inconsistent spaces at the end of the ST segment). 46 ST segments (10%) contained inconsistent placement of tags and 17 contained a typographical error in one of the otherwise matching ST segments, as highlighted in **example 5.30**:

5.30.1s *{1}Realistic:{2} Plots objects with the Realistic visual style applied regardless of the way the objects are displayed on the screen.*	5.30.1t *{1}Realistisch:{2} Plottet Objekte mit der Render-Voreinstellung Realistisch, unabhängig davon, wie die Objekte auf dem Bildschirm angezeigt werden.*
5.30.2s *{1}Realistic:{2} Plots objects with the Relaistic visual style applied regardless of the way the objects are displayed on the screen.*	5.30.2t *{1}Realistisch:{2} Plottet Objekte mit dem visuellen Stil Realistisch, unabhängig davon, wie die Objekte auf dem Bildschirm angezeigt werden.*

TT Inconsistencies

Figure 5.10 gives an overview of category 1 TT sub-segment inconsistencies.

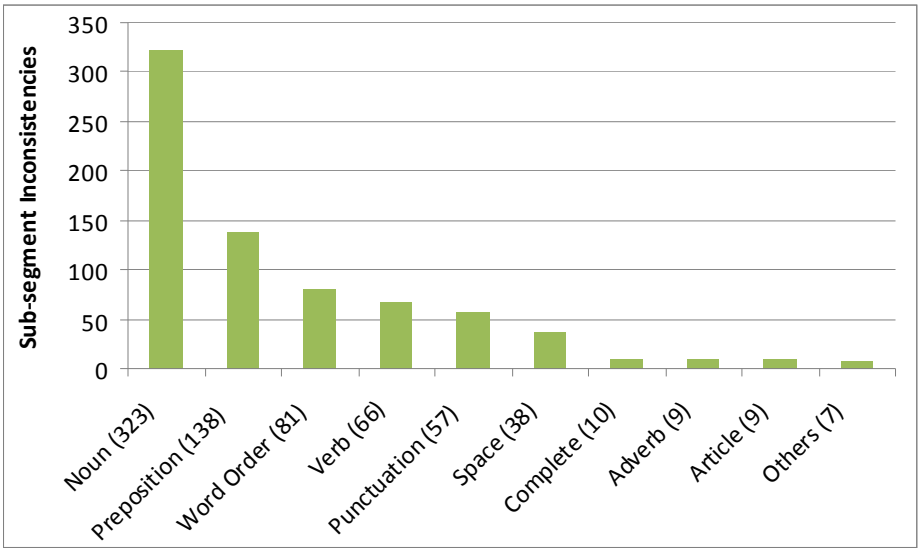


Figure 5.10 TM C Category 1 TT Inconsistencies

In the category 1 TUs we found 730 sub-segment inconsistencies in the TT and a further 10 TT segments that had been completely rewritten. As with the other TMs in this study, we found a large proportion of noun inconsistencies, such as in example 5.29. As table 5.8 above shows, there were 323 noun inconsistencies (44% of the total) of which 12 showed the influence of the source language in one instance but not in another, and 83 contained inconsistencies of capitalisation or case, as may be seen in **example 5.31.** In this example, the noun '(die) Differenz' appears in lower and upper case, rather than with just the first letter capitalised as is required for German language nouns.

5.31.1s *At the Command prompt, enter subtract.*	5.31.1t *Geben Sie in der Befehlszeile differenz ein.*
5.31.2s *At the command prompt, enter subtract.*	5.31.2t *Geben Sie an der Eingabeaufforderung DIFFERENZ ein.*

Example 5.31 also displays a phenomenon that resulted in a high prevalence of preposition inconsistencies. We found 138 preposition inconsistencies in category 1, just under 19% of the total. This is a far higher proportion than found in the previously studied TMs and is the result of a repeated inconsistency in this

TM. In a pattern that can also be seen in category 3, 126 of these preposition inconsistencies (and thus 17% of the total) are the result of alternation between the phrases 'in der Befehlszeile' [in the command line] and 'an der Eingabeaufforderung' [at the command prompt]. While these preposition inconsistencies are thus secondary to the noun inconsistency between 'Befehlszeile' and 'Eingabeaufforderung', there is also an instance of the phrase 'an der Befehlszeile'. This shows that the change in preposition was not necessary, but was nonetheless propagated throughout the TM.

Approximately 11% of the TT inconsistencies in category 1 were due to word order changes. Of these 81 word order inconsistencies, 64 showed a repeated pattern whereby the first word in the 'reference' segment appeared in parentheses at the end of the second segment as in **example 5.32**:

5.32.1s *3D Modeling Tab (Options Dialog Box)*	5.32.1t *Registerkarte 3D-Modellierung (Dialogfeld Optionen)*
5.32.2s *3D Modeling tab (Options dialog box)*	5.32.2t *3D-Modellierung (Registerkarte im Dialogfeld Optionen)*

In cases where there were no existing parentheses, they were added as in **example 5.33**:

5.33.1s *Action Macro Dialog Box*	5.33.1t *Dialogfeld Aktionsmakro*
5.33.2s *Action Macro dialog box*	5.33.2t *Aktionsmakro (Dialogfeld)*

Both 5.32.1t and 5.33.1t were created by the same translator in December 2006 and October 2007 respectively and changed by user 'System' after three years. Segments 5.32.2t and 5.33.2t are attributed to user 'System' on July 16th 2009.

The next most prevalent types of TT inconsistencies in category 1 are: verb inconsistencies (66 or 9%), punctuation inconsistencies (57 or 8%), and

inconsistent spacing (38 or 5%). The verb inconsistencies tended to be changes in verb prefix rather than the whole verb as in **example 5.34**:

5.34.1s *{1}Swap the Colors.{2}*	5.34.1t *{1}Farben austauschen{2}*
5.34.2s *{1}Swap the Colors{2}*.	5.34.2t *{1}Farben tauschen{2}*

Of the 57 punctuation inconsistencies, 17 were due to missing full-stops and in 11 instances the difference between the reference and the subsequent sentences was that one used a full-stop and the other used a colon as in **example 5.35**:

5.35.1s *{1}Front and back clipping planes{2}.*	5.35.1t *{1}Vordere und hintere Zuschneideebenen{2}:*
5.35.2s *{1}Front and back clipping planes{2}.*	5.35.2t *{1}Vordere und hintere Zuschneideebenen{2}*.

5.4.3 Category 2 TUs (Inconsistent ST → Consistent TT)

In TM C, 2077 TUs fell into category 2. Table 5.9 lists all category 2 inconsistencies, which, by definition, appear only in ST segments.

Category 2 ST	
Letter case	480
Space	287
Word (Noun, verb, adverb, article)	146 (70, 47, 16, 13)
Punctuation	89
Typo	19
Others	22
Total Segments	2077
Segment Level Inconsistencies	1016
Total Sub-segment Inconsistencies	1043

Table 5.9 TM C Category 2 Inconsistencies

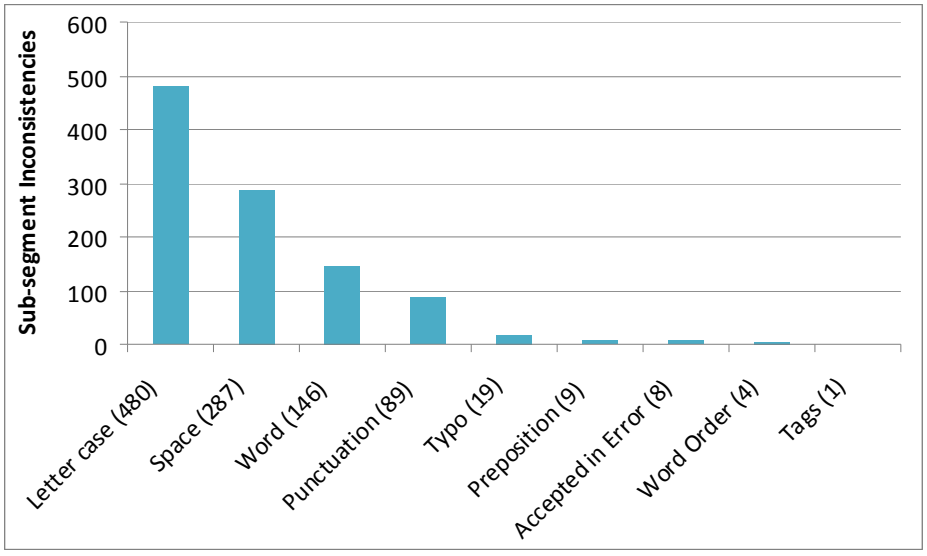


Figure 5.11 TM C Category 2 Inconsistencies

Within the ST segments there are 1043 sub-segment inconsistencies. The majority of these (480 or 46%) are due to inconsistent letter case in the ST. **Example 5.36** shows a typical example of this ST inconsistency.

5.36.1s *Action Macro*	5.36t *Aktionsmakro*
5.36.2s *action macro*	5.36t *Aktionsmakro*

Although there are instances of the ST letter case being retained in the TT in this TM, particularly if the ST segment is in upper case, capitalisation of the first letter of German language nouns means introduced consistency would be expected in example 5.36. However, we have already seen (in the previous section) that 306 inconsistencies of ST letter case in category 1 were associated with further TT inconsistencies.

Inconsistent spacing (287 inconsistencies) accounts for 28% of the category 2 inconsistencies. Of those 287, 245 are inconsistent white spaces at the end of one of the ST segments. **Example 5.37** contains punctuation inconsistencies, of which there are 89 (9%). Four variations in the ST are translated as the TT segment 5.37t. Assigning the segment 5.37.1s the status of 'reference' segment, this counts as three inconsistencies.

5.37.1s *At the Command prompt, enter {1}bconstruction.{2}*	5.37t *Geben Sie in der Befehlszeile {1}bkonstruktion{2} ein.*
5.37.2s *At the Command prompt, enter {1}bconstruction{2}*	5.37t *Geben Sie in der Befehlszeile {1}bkonstruktion{2} ein.*
5.37.3s *At the Command prompt, enter {1}bconstruction{2}.*	5.37t *Geben Sie in der Befehlszeile {1}bkonstruktion{2} ein.*
5.37.4s *At the Command prompt, enter {1}bconstruction>{2}*	5.37t *Geben Sie in der Befehlszeile {1}bkonstruktion{2} ein.*

There are 70 noun inconsistencies (7% of the total) in the category 2 ST segments. Of these 70, 19 differed in number and 14 alternated between hyphenated and unhyphenated nouns. **Example 5.38** shows introduced noun consistency from what may be an error in the ST.

5.38.1s *At the Command prompt, enter {1}thicken{2}.*	5.38t *Geben Sie in der Befehlszeile {1}DICKE{2} ein.*
5.38.2s *At the Command prompt, enter {1}thickness{2}.*	5.38t *Geben Sie in der Befehlszeile {1}DICKE{2} ein.*

Example 5.39 shows one of the eight occasions in our TM C corpus when we believe a match was wrongly accepted, as there is a semantic difference between 5.39.1s and 5.39.2s.

5.39.1s *(CTRL not pressed)*	5.39t *Fläche verschoben*
5.39.2s *(CTRL pressed twice)*	5.39t *Fläche verschoben*

5.4.4 Category 3 TUs (Consistent ST → Inconsistent TT)

TM C contains 826 TUs in which consistent ST segments were translated to inconsistent TT segments. From these 826 TUs, 408 inconsistencies were found at the segment level of which 9 were considered to have been completely rewritten. These inconsistent segments contained 570 sub-segment inconsistencies. Table 5.10 gives an overview of these inconsistencies. The same information is provided in chart form in figure 5.12.

Category 3 TT	
Noun	282
Preposition	112
Space	61
Punctuation	44
Verb	30
Word Order	17
Adverb	8
Others	16
Completely rewritten	9
Total Segments	826
Segment Level Inconsistencies	408
Total Sub-segment Inconsistencies	570

Table 5.10 TM C Category 3 Inconsistencies

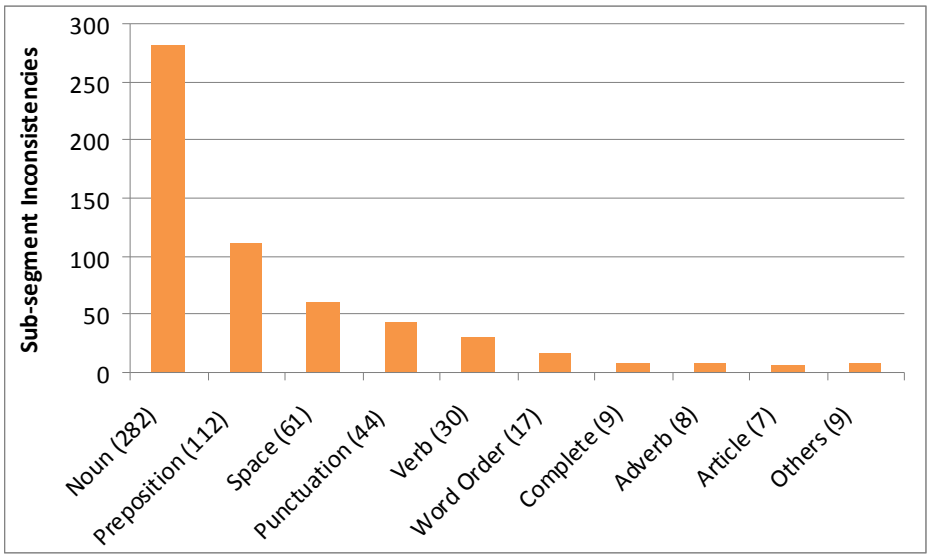


Figure 5.12 TM C Category 3 Inconsistencies

Again, the largest category of TT inconsistency was noun inconsistency. Of 282 noun inconsistencies (49% of the total), 53 contained inconsistent letter case and 30 showed source language influence in one instance as in **example 5.40** below where 'Polyarc-Segment' is used in 5.40.1t and 'Polybogen-Segment' in 5.40.2t.

5.40s *{1}Example: Apply a Diameter Constraint{2}Specify a diameter constraint by picking a circle, arc, or polyarc segment.{3}*	5.40.1t *{1}Beispiel: Anwenden einer Abhängigkeit für den Durchmesser{2}Legen Sie eine Abhängigkeit für den Durchmesser fest, indem Sie einen Kreis, einen Bogen oder ein Polybogen-Segment auswählen.{3}*
5.40s *{1}Example: Apply a Diameter Constraint{2}Specify a diameter constraint by picking a circle, arc, or polyarc segment.{3}*	5.40.2t *{1}Beispiel: Anwenden einer Abhängigkeit für den Durchmesser{2}Legen Sie eine Abhängigkeit für den Durchmesser fest, indem Sie einen Kreis, einen Bogen oder ein Polyarc-Segment auswählen.{3}*

The noun 'Polybogen' appears 18 times in the German TT, 'Polyarc' only once, suggesting that Polybogen is the term chosen usually. The alternation between 'in der Befehlszeile' [in the command line] and 'an der Eingabeaufforderung' [at the command prompt], as seen previously in example 5.31, occurs 93 times in category 3 TUs, causing over a third of the noun inconsistencies and giving us a high incidence of preposition inconsistencies. While these preposition inconsistencies are secondary to the noun inconsistencies, they do not appear to be necessitated by the change of noun, as TM C also contains three uses of the phrase 'an der Befehlszeile'. TM C contains 112 preposition inconsistencies in category 3 TUs – 20% of the total.

After noun and preposition inconsistencies, the next most prevalent category is spacing (61) and punctuation (44) inconsistencies, together 18% of the total. These inconsistencies are difficult to spot: differences in types of hyphens or

dropping of full stops may not make a TT segment inaccurate, but leave inconsistent TM data that may be further propagated, as in **example 5.41**.

5.41s *{1}vlx{2})—a compiled set of one or more LSP and/or dialog control language (DCL) files.*	5.41.1t *{1}vlx{2}) – ein kompilierter Satz einer oder mehrerer LSP- und/oder DCL (Dialog Control Language)-Dateien.*
5.41s *{1}vlx{2})—a compiled set of one or more LSP and/or dialog control language (DCL) files.*	5.41.2t *{1}vlx{2}): ein kompilierter Satz einer oder mehrerer LSP- und/oder DCL (Dialog Control Language)-Dateien.*

30 TT verb inconsistencies are contained in the category 3 TUs (5% of the total). Five of these were differences of verb prefix as in example 5.34, but more often there was a completely different verb chosen as in **example 5.42**.

5.42s *Attaches the palette to an anchor tab base at the left or right side of the drawing area.*	5.42.1t *Hängt die Palette an eine Verankerungsleiste auf der linken oder rechten Seite des Zeichenbereichs an.*
5.42s *Attaches the palette to an anchor tab base at the left or right side of the drawing area.*	5.42.2t *Weist die Palette einem Fixierungsanker auf der linken oder rechten Seite des Zeichenbereichs zu.*

Both of these TT segments are attributed to the same translator in the metadata. 5.42.1t was created in October of 2007 and last changed in January of 2008. 5.42.2t was created in December of 2006 and last changed in December of 2009. Both are commented “from previous releases” but further comments suggest that they originated from different XML files. When these sources were combined inconsistency in the TM was caused.

5.4.5 Category 4 TUs (Consistent ST → Consistent TT)

The other TUs in TM C either contain ST segments that are not repeated or fall into category 4, containing consistent repeated ST and TT segments. In total 18,343 TUs contain ST segments that were repeated at least once and were translated consistently and are thus category 4 TUs. If we add the number of category 3 TUs (826 TUs with consistent ST and inconsistent TT) to the number of category 4 TUs (18,343), we can see that the total number of TUs with repeated ST are 19,169. Of these, 826 (or 4.3%) contained inconsistency that was introduced in the TM translation process.

5.5 TM D

TM D is an English-to-Japanese TMX file that was created using Idiom Worldserver during the same localisation project as TM C and contains 298,700 TUs. After removing the TUs that contain only numbers, dates, or punctuation, 292,258 remained. In this section we ensure that our sample is representative of the data as a whole, then present analysis results for each category of TUs.

5.5.1 Testing the Data Sample

Due to the difference in scale between the first set of TMs and the second, we chose a sample of the first 50,131 TUs (sorted alphabetically) for analysis and categorisation as specified in section 4.4. Of these 50,131 TUs, 1,713 or 3.42% fell into category 3. Within the whole TM, 9,840 or 3.37% fell into category 3. More details are in table 5.11.

TM D	Sample	Total
Number of TUs	50131	292258
Category 3 (percentage in brackets)	1713 (3.42%)	9840 (3.37%)
Standardised TTR (ST)	15.09	15.83
Mean sentence length in ST (words)	16.43	17.15
Mean ST word length (chars)	4.44	4.69

Table 5.11 Comparison of Corpus Analysis Statistics for Sample & Whole of TM D

The Wordsmith Wordlist software used to measure corpus statistics (such as type/token ratio, mean sentence length) is unable to discern individual words or accurate sentence lengths in Japanese - whole sentences and phrases are incorrectly recognised as words. For this reason, Wordlist cannot compare the

similarity of target language segments between the sample and the entire corpus. As in section 5.4.1, the standardised type/token ratio measures the average number of unique words (types) per block of 1000 words (tokens) in the text.

A chi square test, comparing the sample measured in this study and subsequent 50,000 TU sections (sorted alphabetically) with the whole of TM D, found no significant differences in the number of category 3 TUs [p = 0.19], standardised TTR [p = 0.83], mean sentence length [p = 0.81], or the mean word length [p = 0.90]. These findings support the hypothesis that the measurements shown for the sample in table 5.11 do not differ significantly from the whole [p > 0.05].

5.5.2 Category 1 TUs (Inconsistent ST → Inconsistent TT)

1980 TUs contained minor ST inconsistencies (as specified in section 4.4) and inconsistent translations. Table 5.12 gives an overview of these TUs, broken down according to the kind of inconsistency evidenced on the source and target sides.

Category 1 ST		Category 1 TT	
Letter case	753	Noun	616
Tags	137	Tags	147
Punctuation	73	Space	141
Space	68	Explicitation	132
Typo	11	Punctuation	129
		Verb	75
		Particle	42
		Others	9
		Completely rewritten	12
Total Segments	1980	Total Segments	1980
Segment Level Inconsistencies	1042	Segment Level Inconsistencies	1042
Total Sub-segment Inconsistencies	1042	Total Sub-segment Inconsistencies	1291

Table 5.12 TM D Category 1 Inconsistencies

ST Inconsistencies

Figure 5.13 gives an overview of category 1 ST inconsistencies, broken down according to sub-segment level inconsistency.

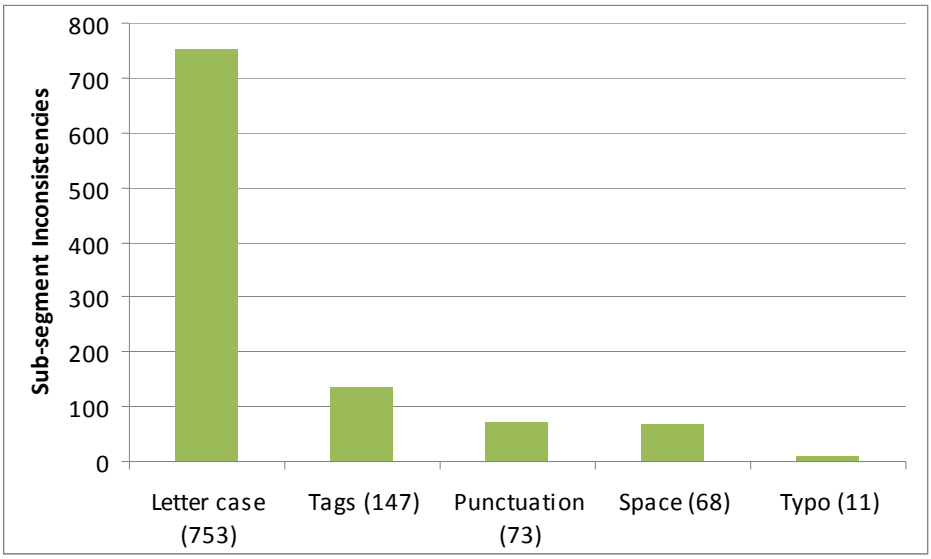


Figure 5.13 TM D Category 1 ST Inconsistencies

As with the other three TMs in this study, the most commonly occurring category of ST inconsistency was letter case or capitalisation of words, which accounted for 753 (72%) of ST inconsistencies in category 1. Inconsistent use of case is illustrated in **example 5.43**:

5.43.1s *[Yes/No] <N>: {1}Enter {2}y{3} or {4}n{5} or press {6}ENTER*	5.43.1t [はい(Y)/いいえ(N)] 〈いいえ〉: {1}{2}y{3} または {4}n{5} と入力するか、{6}[Enter]を押します。
5.43.2s *[Yes/No] <N>: {1}Enter {2}y{3} or {4}n{5} or press {6}Enter*	5.43.2t [はい(Y)/いいえ(N)] <N>: {1}{2}y{3} または {4}n{5} と入力するか、[Enter]を押します。{6}

The ST segments display inconsistent letter case as highlighted. In the aligned segment 5.43.2t, the letter N encased in tags has been left unaltered in the TT, but in 5.43.1t it has been replaced by いいえ [no]. 137 ST tag inconsistencies (13% of the total) were found, of which 65 contain an extra tag at the end as in **example 5.44**:

5.44.1s *{1}Home tab{2}3D Modeling panel{3}Sphere{4}*	5.44.1t {1}[ホーム]タブ {2} [モデリング] パネル {3} [球]{4}
5.44.2s *{1}Home tab{2}3D Modeling panel{3}Sphere{4}.{5}*	5.44.2t {1}[ホーム]タブ{2}[###3D モデリング]パネル{3}[球]{4}{5}

The aligned TT contains two inconsistencies that again are of a different kind to the inconsistency in the ST. 5.44.1t contains a space after the word タブ [tab] and 5.44.2t has been marked with a repeated hash or number sign, a pattern that occurs in 31 category 1 TUs. Propagation of these segments may add further inconsistency throughout the corpus.

Again, there is a high incidence of inconsistent placing of the space character in this TM. Many of these spaces may only be noticed by automatically comparing an ST segment with the following, seemingly identical, ST segment, as 54 of the 68 space inconsistencies are at the end of the segment following a full stop. As a result, match values would be reduced from 100% for seemingly identical ST segments. Once again, the aligned TT segments contain other kinds of inconsistencies, as in **example 5.45**:

5.45.1s *{1}lsp{2} file.*	5.45.1t {1}lsp{2} ファイルから自動的にロードされます。
5.45.2s *{1}lsp{2} file.* [Contains extra space]	5.45.2t {1}lsp{2} ファイルは変更しないでください。

The ST segment contains a space inconsistency, while the aligned TT segments differ by particle (は 'wa' and から 'kara'), 5.45.1t has the additional adjective 自動的 'jidouteki' [automatic], and the verbs differ semantically and in form. These TT segments also provide examples of explicitation in Japanese TT, a phenomenon that appears throughout both English-to-Japanese TMs. The translation of a ST noun has become a sentence in the Japanese TT, containing

detail not present in the ST. Thus we have [It will be automatically loaded from the lsp file.] in segment 5.45.1t and [Please do not change the lsp file.] in segment 5.45.2t. Some explanations for target text explicitation such as this may be found in section 6.5.5.

The TUs contain only one ST segment aligned with one TT segment, showing that two ST segments were not merged to form the more explicit TT. Both 5.45.1t and 5.45.2t were created by the same user on December 12th 2006. However, 5.45.1t was changed three years later (a change credited to user 'System'). Segment 5.45.2t is similar to segments repeated throughout the TM, but the 2009 edit of 5.45.1t may be the result of a translator choosing not to accept a 100% match, and instead editing a fuzzy match, such as the segment that precedes it sequentially in the TM (and which could have been found using a concordance search), which also contains the formation 'ファイルから自動的にロードされ' [You may load automatically from the file.].[31]

TT Inconsistencies

Figure 5.14 gives an overview of category 1 TT sub-segment inconsistencies.

[31] *ST*:{1}lsp{2} file and the NEWCMD command from the {3}newcmd{4}. *TT*: {1}lsp{2} ファイルから自動的にロードされ、コマンド NEWCMD が {3}newcmd{4}.

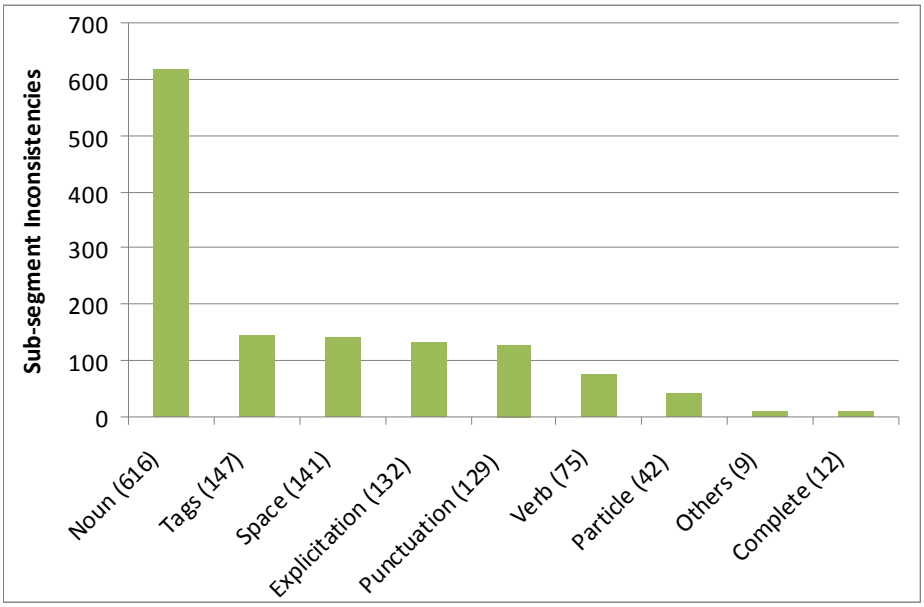


Figure 5.14 TM D Category 1 TT Inconsistencies

In the category 1 TUs we found 1291 sub-segment inconsistencies in the TT and a further 12 TT segments that had been completely rewritten. As in the TT of all other TMs in this study, we found a large proportion of noun inconsistencies. As Table 5.12 shows, there were 616 noun inconsistencies (48% of the total) of which 318 showed the influence of the source language in one instance but not in another. A typical noun inconsistency can be seen in example 5.46:

5.46.1s *Annotational constraints are useful when you want dimensional constraints to have the following characteristics:*	5.46.1t 注釈拘束は、寸法拘束に次の特性を持たせたい場合に便利です。
5.46.2s *Annotational constraints are useful when you want dimensional constraints to have the following characteristics:{1}*	5.46.2t 注釈拘束は、寸法拘束に次の特性を持たせたいときに便利です。{1}

The tag inconsistency from the ST has been applied to the TT, but the word 場合 'baai' (case), written in kanji in 5.46.1t, has been replaced by とき 'toki' (occasion, time) in 5.46.2t. Not only is this a change of noun, the translator has also chosen to write 'toki' using the hiragana syllabary とき rather than in kanji 時.

Unlike in the other TMs, the next most common type of category 1 TT inconsistency is related to tags. The prevalence of TT tag inconsistency is attributable to the following repeated pattern within the TM: the reference segment is replicated followed by an <so> tag, and the content of the reference segment is then repeated syllabically in hiragana. This pattern is repeated 111 times, each counted as a tag and explicitation/interjection inconsistency (as the hiragana version of the segment in each case has been added) and can be seen below in **example 5.47**. In total there are 147 TT segments evincing tag inconsistency (11% of the total) and 132 explicitation/interjection inconsistencies (10% of the total).

5.47.1s *Action Macro Manager Dialog Box* 5.47.2s *Action Macro Manager dialog box*	5.47.1t［アクション マクロ管理］ダイアログ ボックス 5.47.2t ［アクション マクロ管理］ダイアログ ボックス<so>あくしょんまくろかんりだいあろぐぼっくす

The loan words 'action macro' and 'dialog box' are written in katakana, the syllabary for loan words, in 5.47.1t and the first part of 5.47.2t. They are then repeated in hiragana, the syllabary for Japanese words and verb or adjective inflection, in the second part of 5.47.2t. Occasionally hiragana are written above kanji characters as a reading aid, in a system known as furigana. However, furigana are not usually written above katakana characters as they also represent syllables.

Of the 129 punctuation inconsistencies, 31 contain three hash symbols (“###” as shown in 5.44.2t), possibly as an indicator that the translator should revisit or review a segment, but are otherwise identical. 34 feature inconsistency in whether or not the ST punctuation symbols were retained. These were mostly commas and full stops, but in **example 5.48**, we see inconsistent use of the question mark:

5.48.1s *{1}Question:{2} How do I edit an enterprise CUI file if it is read-only in the CUI editor?*	5.48.1t {1}質問：{2}CUI ファイルの共有 CUI ファイルが読み取り専用ファイルである場合、そのファイルを編集するにはどうすればよろしいですか。
5.48.2s *{1}Question:{2} How do I edit an enterprise CUI file if it is read-only in the CUI Editor?*	5.48.2t {1}質問：{2}CUI ファイルの共有 CUI ファイルが読み取り専用ファイルである場合、そのファイルを編集するにはどうすればよろしいですか?

This is interesting as the question mark is not usually used in Japanese. Rather, the role of the preceding statement as a question is marked by the sentence-ending particle か 'ka'. か is written in 5.48.1t and 5.48.2t, marking the sentences as questions, yet in 5.48.2t the question mark is added. The use of the polite adjective よろしい 'yoroshii' (good, alright) in the TT segment is very formal, yet 5.48.2t retains the ST punctuation even though it is not appropriate to formal written Japanese.

5.5.3 Category 2 TUs (Inconsistent ST → Consistent TT)

1801 TUs contain inconsistent ST segments but consistent translations. Table 5.13 gives an overview of the ST inconsistencies to be found in these TUs.

Category 2 ST	
Letter case	505
Space	247
Word (Noun, verb, plurals, article, others)	194 (77, 39, 30, 13, 35)
Punctuation	153
Typo	21
Tag	3
Total Segments	1801
Segment Level Inconsistencies	922
Total Sub-segment Inconsistencies	1123

Table 5.13 TM D Category 2 Inconsistencies

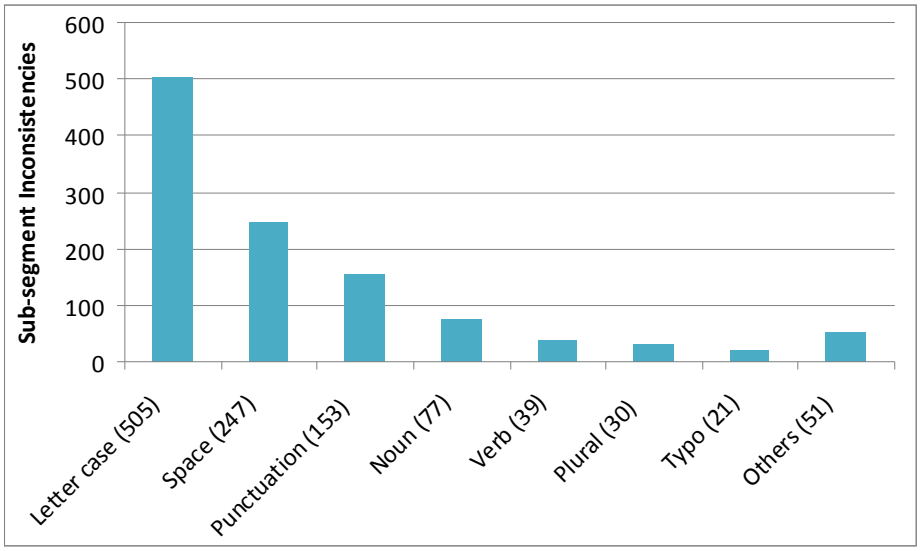


Figure 5.15 TM D Category 2 Inconsistencies

1123 category 2 sub-segment inconsistencies were found in TM D. As in all other data in this study, the bulk of these ST inconsistencies (45%) contained inconsistent letter case as in **example 5.49**:

5.49.1s *At the Command prompt, enter options.*	5.49t コマンド プロンプトに対して、options と入力します。
5.49.2s *At the command prompt, enter options.*	5.49t コマンド プロンプトに対して、options と入力します。
5.49.3s *At the command prompt, enter OPTIONS.*	5.49t コマンド プロンプトに対して、options と入力します。

While 753 category 1 ST segments containing letter case inconsistencies are aligned with TT segments containing further inconsistencies, we can see that in this case three variants of letter case in the ST have been made consistent in the TT. This is to be expected in Japanese as kanji and kana have no letter case, but in 5.49t the Latin alphabet (ローマ字 or 'romaji') has been used for 'options' and the letter case has been normalised.

We again found many space inconsistencies (247 or 22%) in the category 2 ST segments, most of which (191 or 17% of the total) were at the end of the segment. Of the other space inconsistencies, many were added before or after tags, as in the extra space after '{10}' in **example 5.50**, 5.50.1s:

5.50.1s *{1}View{2} {3} {4}Display{5} {6} {7}ViewCube{8} {9} {10}[] On{11}*	5.50t {1}[表示]{2} {3} {4}[表示設定]{5} {6} {7}[ViewCube]{8} {9} {10} [表示]{11}
5.50.2s *{1}View{2} {3} {4}Display{5} {6} {7}ViewCube{8} {9} {10}On{11}*	5.50t {1}[表示]{2} {3} {4}[表示設定]{5} {6} {7}[ViewCube]{8} {9} {10} [表示]{11}

194 (17%) inconsistencies contained inconsistent ST words. Of these, 77 were related to noun choice, 39 to verb choice, 30 to inconsistently pluralised nouns, and 13 to inconsistent articles. In **example 5.51**, the prepositional phrase 'to [company name redacted]' is added in 5.51.2s, and then an important adverb removed in 5.51.3s. The translator has chosen to stick with the meaning of 5.51.1s for the translation.

5.51.1s *All information is sent anonymously to maintain your privacy.*	5.51t すべての情報は、プライバシーを保護するために匿名で送信されます。
5.51.2s *All information is sent to [company name redacted] anonymously to maintain your privacy.*	5.51t すべての情報は、プライバシーを保護するために匿名で送信されます。
5.51.3s *All information is sent to [company name redacted] to maintain your privacy.*	5.51t すべての情報は、プライバシーを保護するために匿名で送信されます。

Example 5.52 contains an inconsistent bracket in 5.52.2s, an example of the 153 (14%) punctuation inconsistencies, but also shows that the TT segments marked with hash symbols (as in example 5.44.2t, here with added comments) may be propagated in the TT. The translator (in the comment marked with '■3-(B037)') has explained that he chose the term 塗り潰し色 'nuritsubushiiro' for filling in colours and this comment has been included with the rest of the accepted matching segment in aligning 5.52t with 5.52.2s.

5.52.1s *{1} If None (Color) is selected as the Map Type then the color needs to be selected.*	5.52t {1} [マップの種類]として[###塗り潰し色]を選択した場合は、色を選択する必要があります。■3- (B037)「なし」という選択肢はなく、「塗り潰し色」という選択肢が表示されるので、このようにしました。(Koizumi 06/11/21)
5.52.2s *{1} If None (Color)) is selected as the Map Type then the color needs to be selected.*	5.52t {1} [マップの種類]として[###塗り潰し色]を選択した場合は、色を選択する必要があります。■3- (B037)「なし」という選択肢はなく、「塗り潰し色」という選択肢が表示されるので、このようにしました。(Koizumi 06/11/21)

5.5.4 Category 3 TUs (Consistent ST → Inconsistent TT)

In TM D from 1713 TUs 887 inconsistencies were found at the segment level (of which we considered 35 to have been completely rewritten) along with 1035 inconsistencies at the sub-segment level. Table 5.14 and figure 5.16 give an overview of these inconsistencies.

Category 3 TT	
Noun	365
Space	272
Punctuation	183
Verb	59
Particle	57
Tags	33
Explicitation	32
Word Order	16
Pronoun	7
Others	11
Completely rewritten	35
Total Segments	1713
Segment-level Inconsistencies	887
Total Sub-segment Inconsistencies	1035

Table 5.14 TM D Category 3 Inconsistencies

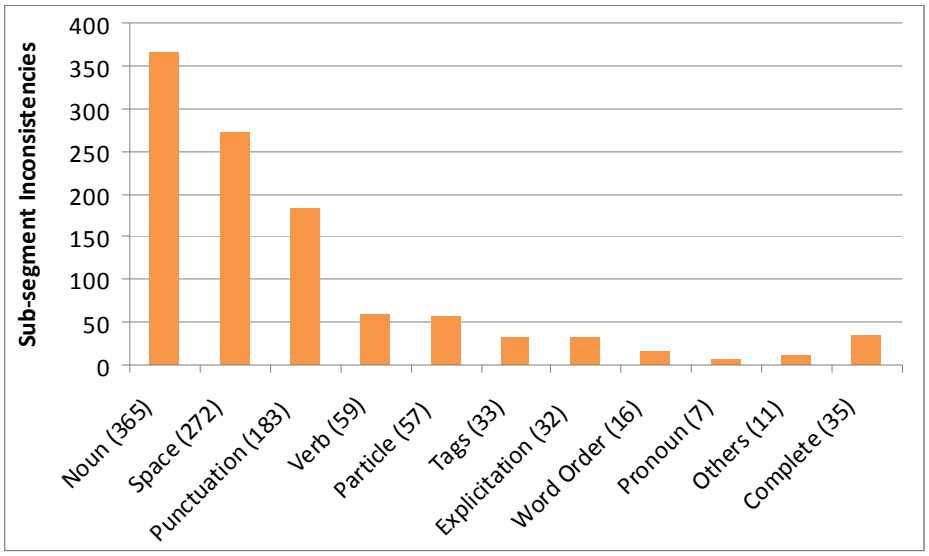


Figure 5.16 TM D Category 3 Inconsistencies

Again, the largest proportion of sub-segment inconsistency in category 3 is noun inconsistency. Of the 365 (35% of category 3 total) noun inconsistencies, 77 show the influence of the source language in one instance and not in another, and 41 follow the pattern of alternating between the borrowed English word レイヤ and the native Japanese word 画層 'gasou' as a translation for the word 'layer', as in **example 5.53**:

5.53s *A new layer group filter can be nested only under another group filter.*	5.53.1t 新しいレイヤグループ フィルタは、他のグループ フィルタに対してのみネストできます。
5.53s *A new layer group filter can be nested only under another group filter.*	5.53.2t 新しい画層グループ フィルタは、他のグループ フィルタに対してのみネストできます。

Looking at the metadata for TUs such as those in example 5.53, the use of レイヤ seems attributable to a single translator who made changes to these TUs on July 11th 2008. The TUs containing 画層 were from various dates between 2006 and 2009 and were translated by two other translators. This would suggest that 画層 is the usual translated term for 'layer' in TM D.

272 (26%) space inconsistencies were found in the category 3 TUs, of which 54, although they looked identical, had an inconsistent trailing space at the end of the segment. These would not necessarily affect the finished translation, but such segments do demonstrate introduced inconsistency. A further 183 (18%) contained punctuation inconsistencies, of which 23 were marked with hash symbols as in example 5.44. Many of these punctuation inconsistencies showed some indecision as to whether to retain the ST punctuation or to add punctuation appropriate for the TL as in example 5.48 or **example 5.54**. Some explanation of this type of format inconsistency may be seen in section 6.5.2.

5.54s *Accesses Dimensioning mode*	5.54.1t 寸法記入モードにします
5.54s *Accesses Dimensioning mode*	5.54.2t 寸法記入モードにします。

The data also contains 59 verb inconsistencies, 57 particle inconsistencies, and 32 examples of explicitation in translation in one of the segments. All of these can be seen in **example 5.55**:

5.55s *A translucent object transmits light, but light is also scattered within the object.*	5.55.1t 半透明のオブジェクトは光を通しますが、同時に光はオブジェクト内で拡散されます。
5.55s *A translucent object transmits light, but light is also scattered within the object.*	5.55.2t 半透明なオブジェクトは光を通しますが、オブジェクト内でも光が分散します

半透明 'hantoumei' [semi-transparent] is followed by the particle の in 5.55.1t to describe the object (オブジェクト) in the first part of the segment. In 5.55.2t 半透明 is used as an adjective and is thus followed by な to describe the object. In 5.55.1t, explicitation comes in the addition of 同時に 'douji ni' [at the same time] 光は 'hikari wa' [light is...]. This added use of the particle は changes the topic of the sentence. The verb in 5.55.1t is the passive 拡散されます 'kakusan saremasu' [may be diffused], and in 5.55.2t 分散します 'bunsan shimasu' [is dispersed] was chosen, as both verb and verb form were changed. The particle も 'mo' [also] is added in 5.55.2t, and finally 5.55.2t features no full stop. The number of changes here result in 5.55.2t being classified as a completely rewritten translation.

Some changes of particle may be a secondary effect of a change in verb or verb form. In the following example, a change in particle is required by verb choice

with 表現 'hyougen' [scale representation] taking the particle が 'ga' when the verb ある 'aru' [to exist] is used, and the direct object particle を 'wo' with 持つ 'motsu' [to hold].

5.56s *Annotative objects may have multiple {1}scale representations{2}.*	5.56.1t 異尺度対応オブジェクトには複数の{1}尺度表現{2}がある場合がありま す。
5.56s *Annotative objects may have multiple {1}scale representations{2}.*	5.56.2t 異尺度対応オブジェクトには、複数の{1}尺度表現{2}を持つものもあり ます。

5.5.5 Category 4 TUs (Consistent ST → Consistent TT)

The remaining repeated ST segments in TM D (25,541) were translated consistently, as is expected when translators use a TM tool. By adding the number of TUs in category 3 (1,713) and category 4, we can see that there were 27,254 TUs that contained repeated ST segments. Of these, 6.3% fell into category 3, in which inconsistency was introduced in the process of translation using TM.

5.6 Comparative Analysis with QA Tools

Section 2.7.3 introduced QA (quality assurance) tools that can be used to measure consistency in TMs. In this section we analyse some of the TM data from this study using two commercial tools – ApSIC Xbench and QA Distiller from Yamagata Europe – to see what level of detail is available from these tools and to compare the results with our findings.

5.6.1 ApSIC Xbench

Two popular tools which are used by several of the interview participants in Chapter 6 are ApSIC Xbench[32]and QA Distiller[33]. The former, as part of its QA functionality, can run a check on a TM, listing categories such as: inconsistency in source (category 2 in this study); inconsistency in target (category 3); and key term mismatch. QA Distiller may be used to search for TM inconsistencies, omissions, formatting problems, and terminology errors.

As a comparison, a check on TM A using Xbench 2.9 appeared to find all category 2 and 3 inconsistencies at segment level (the tool does not provide figures for the number of inconsistencies of each type found). Possibly because a search for inconsistent TT aligned with ST containing minor inconsistencies is difficult to automate, the software cannot search for category 1 inconsistencies. Two additional TT inconsistencies were found in which each segment contained a dash '-' that looked identical, but due to Xbench not being Unicode-enabled, appeared garbled in one instance. This may be because the dashes were added using different locale settings. The lack of Unicode compliance is a bigger

[32] ApSIC Xbench is a programme presented as a 'reference tool' and available from http://www.apsic.com/en/products_xbench.html.

[33] QA Distiller is a translation QA tool available from http://www.qa-distiller.com/.

problem for Japanese TMs, as Japanese text is not compatible with Xbench unless the locale for the computer is changed to Japan[34].

While Xbench is efficient at finding category 2 and 3 inconsistencies at a segment level, there is no further detail of the type of segment or sub-segment inconsistency produced. A QA check on TM A also found 94 false positives for repeated words in the TT. These were all due to an article occurring twice in succession in a segment, however the first occurrence in each case was a relative pronoun referring back to a previous noun. An example is in the following sentence: "Markieren Sie die Elemente, die die Kugel durchdringen sollen."
The first 'die' after the comma is a relative pronoun referring to 'die Elemente', as a translation of 'which' from the ST segment: "Select the elements which are to penetrate the sphere". Similarly the sentence:

> Halten Sie die Maustaste gedrückt, und bewegen Sie sie von dem Element weg, auf das Sie geklickt haben

contains the polite second person pronoun 'Sie' followed by the relative pronoun 'sie', referring to the mouse button or 'die Maustaste'.

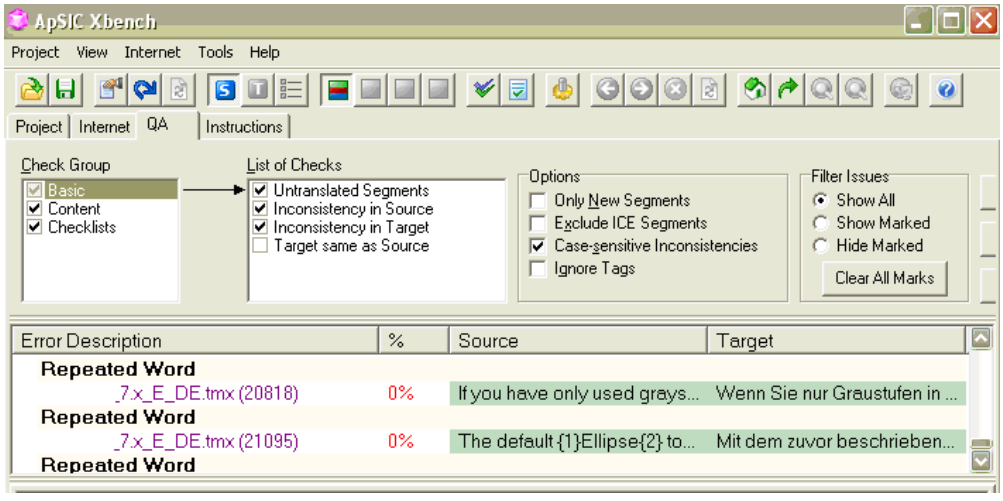


Figure 5.17 Screenshot of the QA Function in ApSIC Xbench

[34] The locale may also be changed using Applocale with the Microsoft XP operating system.

5.6.2 QA Distiller

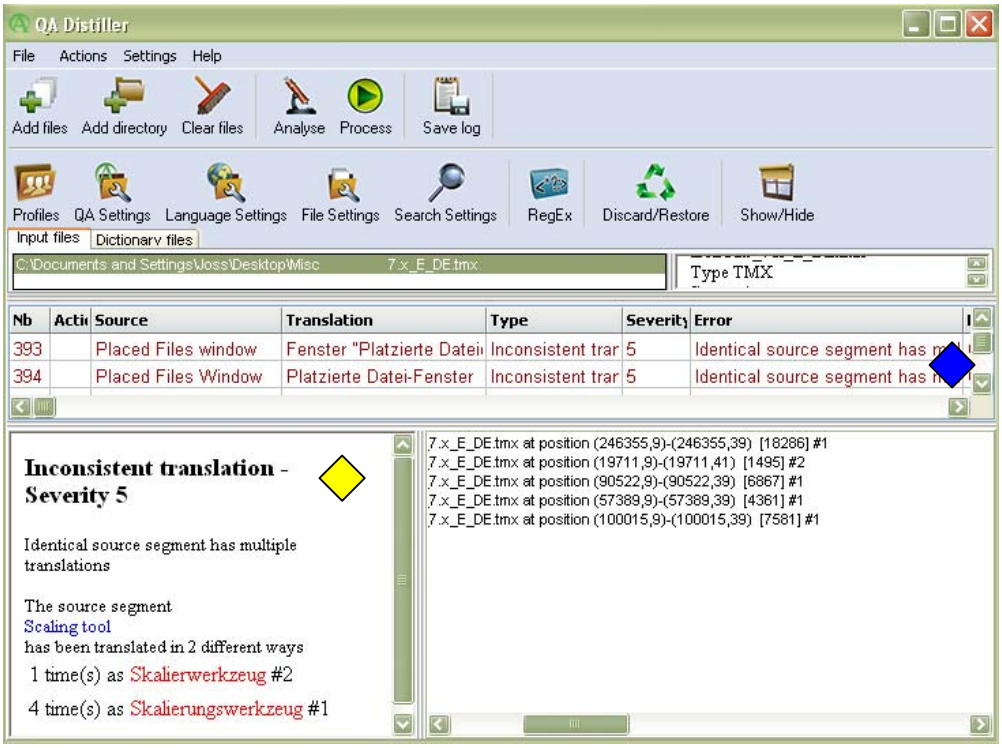


Figure 5.18 Screenshot of the QA Function in QA Distiller

A QA check on TM A using QA Distiller 8.5 returns detailed results. Inconsistencies are graded in severity, and the number of occurrences of a repeated inconsistency is provided (see yellow diamond shape in figure 5.18). The program presents inconsistencies in a spreadsheet format up to a total of 10,000 (see blue diamond shape in figure 5.18). As this QA check exceeded the limit, no total number was given. Double-clicking on an error message in QA Distiller brings up the internal X-Editor tool, wherein the user may edit or delete segments (see figure 5.19).

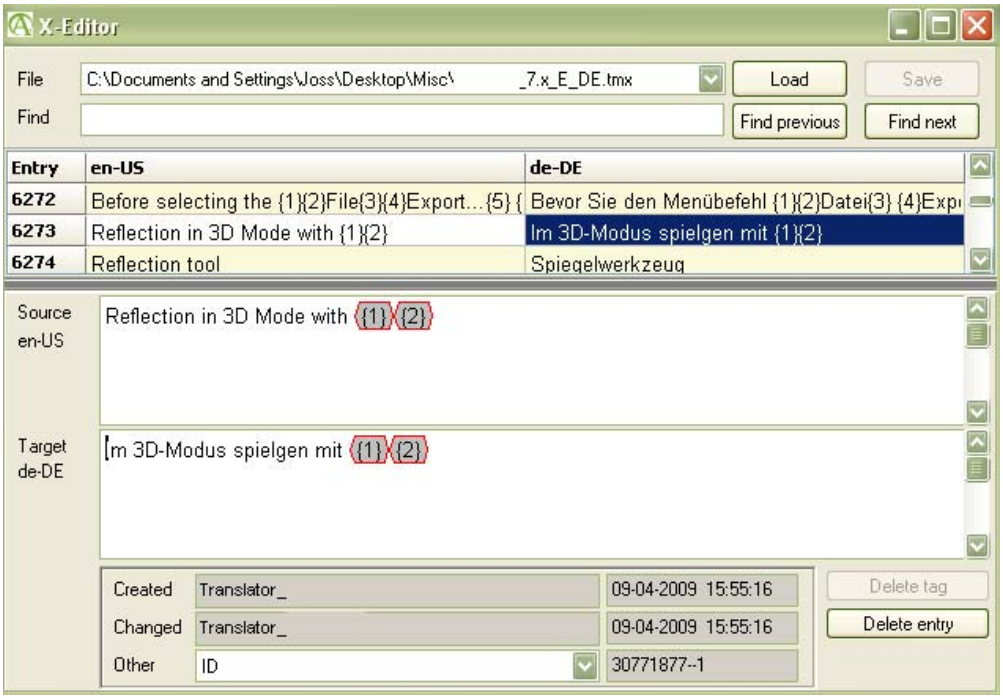


Figure 5.19 Screenshot of the X-Editor Window in QA Distiller

The program appeared to be efficient at finding segment-level inconsistencies in categories 2 and 3, giving each error a type such as 'inconsistent translation', 'consecutive spaces', 'leading/trailing spaces', and 'capitalisation'. The results also contained a large number of false positives. It incorrectly identified 270 'corrupt character' errors due to an accented character in a repeated term appearing (correctly) in the German TT and also misidentified 201 number value errors. An example is in the translation of:

> Options for ends allow you to select dot diameters which are from two to ten times the line thickness.

This was translated as

> In den Listen Anfang und Ende unter Ausführung der Enden können Sie Punktdurchmesser auswählen, die das Zwei- bis Zehnfache der Linienstärke betragen.

Numbers that are written or spelled with punctuation appear to often be misidentified, as in the example provided. Other false positives of this type included ‘two’ translated as ‘both’, and problems with the use of the term ‘3d’.

Both of these tools are clearly useful and may provide valuable information about TMs. However, this section demonstrates the contribution of this study in measurement of inconsistency at a sub-segment level and in using more detailed categorisation. Furthermore, neither tool provides sub-segment categorisation of inconsistencies, or the frequency of category 1 TUs, although both calculations are difficult to automate. Both tools also returned a number of false positives in their error reports. Nonetheless, they quickly provide a snapshot of the state of consistency in a TM and, in the case of QA Distiller, error detail and a straightforward maintenance interface. The additional functionality is reflected in the cost of the tools – Xbench is free of charge at the time of writing (as it is under beta test) and QA Distiller licenses currently cost from €249 to €2500.

5.7 Summary

This chapter demonstrated the practical application of the first phase of the mixed methods study as set out in Chapter 4. It contained an analysis of each TM, presented in turn and subdivided into the four categories of repeated segments. Section 5.6 contains the results of a QA check using two current commercial QA tools so that the results of an automated QA scan may be compared with the results of the current study. The findings in this chapter begin to answer our research questions as specified in section 4.2. Addressing question 1, we can say that, in this case study, TMs are not consistent. Across the four TM corpora, the rate of introduced inconsistency as shown in table 5.15 represents lost leverage and time spent editing previously accepted translations. As such, all of our stated general assumptions about TM (consistency, cost savings, and time savings) are affected by introduced inconsistency. Section 6.5 gives interview participants' opinions as to whether this inconsistency is found more generally in TMs.

	TM A	TM B	TM C	TM D
Category 3 TUs	390	239	826	1713
Category 4 TUs	6674	4263	18343	25541
Total TUs with repeated ST segments (Category 3+4)	7064	4502	19169	27254
Percentage of TUs with introduced inconsistency	5.5%	5.3%	4.3%	6.3%

Table 5.15 Introduced Inconsistency in all TMs

Each TM corpus in this study shows a high proportion of introduced noun or term inconsistency in category 3. Many noun inconsistencies demonstrate influence from the source language, and different translation decisions have been propagated throughout the TM, such as the alternation between レイヤ 'laya' and 画層 'gasou' [layer] from TM D (example 5.53), or between the

alternated whole phrases 'in der Befehlszeile' [in the command line] and 'an der Eingabeaufforderung' [at the command prompt] in TM C (example 5.31). Verb and punctuation changes are also common throughout the corpora. These sub-segment inconsistencies are included in interview questions in the following chapter and discussed further in section 6.5.

Comments and markings, possibly as an indicator that the translator should revisit or review a segment, have been propagated in TMs C and D. In each of the TM corpora there appears to be a lack of clarity as to whether ST punctuation and formatting should be replicated or replaced by that native to the TT (see examples 5.48 and 5.54), leaving a combination of both in the TM data. The English-to-Japanese TM data (B and D) in particular also show evidence of explicitation in the TT. These further questions are also to be addressed in the qualitative study, contained in chapter 6.

Our results also show that TM source texts are not consistent. By comparing category 1 and 2 results, we can see the rate at which minor ST inconsistencies were corrected (or, inversely, further inconsistencies propagated) between the four TM corpora in table 5.16.

	TM A	TM B	TM C	TM D
Category 1 TUs (inconsistent → inconsistent)	370	65	995	1980
Category 2 TUs (inconsistent → inconsistent)	613	914	2077	1801
Total TUs with minor inconsistency in ST segments	983	979	3072	3781
Percentage of those TUs made consistent	62.4%	93.4%	67.6%	47.6%

Table 5.16 Minor Inconsistency in all TMs

A benefit of TM may be seen from the number of TUs in category 2, containing consistency introduced in the TM translation process. English-to-Japanese TM B aside, many of the TUs with inconsistent ST evinced further inconsistency in the aligned TT. ST segments featuring inconsistent letter case or extra trailing spaces at the end are often aligned with TT segments containing further inconsistency, for example, 60% of TUs with inconsistent ST letter case in TM D were associated with further TT inconsistency. However, when the number of inconsistent ST segments that are aligned with TT segments featuring further inconsistencies were tested for correlation using SPSS, the result was not statistically significant, so there was no clear correlation between the rate of inconsistent ST and introduced inconsistency in the TT[35]. Our interviewees' experiences of source text inconsistency will be discussed in section 6.3.

Our second research question asks how consistency can be identified and measured in TM data. The methodology described in chapter 3 and the results in the current chapter offer one possible way of identifying and measuring consistency. These results show that the typology and categorisation operationalised in section 4.4 brought quantifiable results from the TMs studied. Chapter 6 will offer our interviewees opinions as to whether these phenomena appear more widely in TMs.

[35] A strong but statistically insignificant correlation was found between source and target inconsistencies where r = .945, p = .055.

6 Qualitative Phase: Data Analysis

6.1 Introduction

In this chapter we present results from the second phase of our study: the qualitative interviews conducted following and based on the results of the quantitative analysis in Chapter 5. We begin by profiling the interviewees, explaining the coding process, and discussing some themes that emerged during coding. The interviewees all described a simplified translation workflow, as in figure 6.1. We have chosen to structure this chapter to follow the workflow, discussing emergent themes at each stage, beginning with source text issues such as ambiguous and non-native ST and the influence of the translation client.

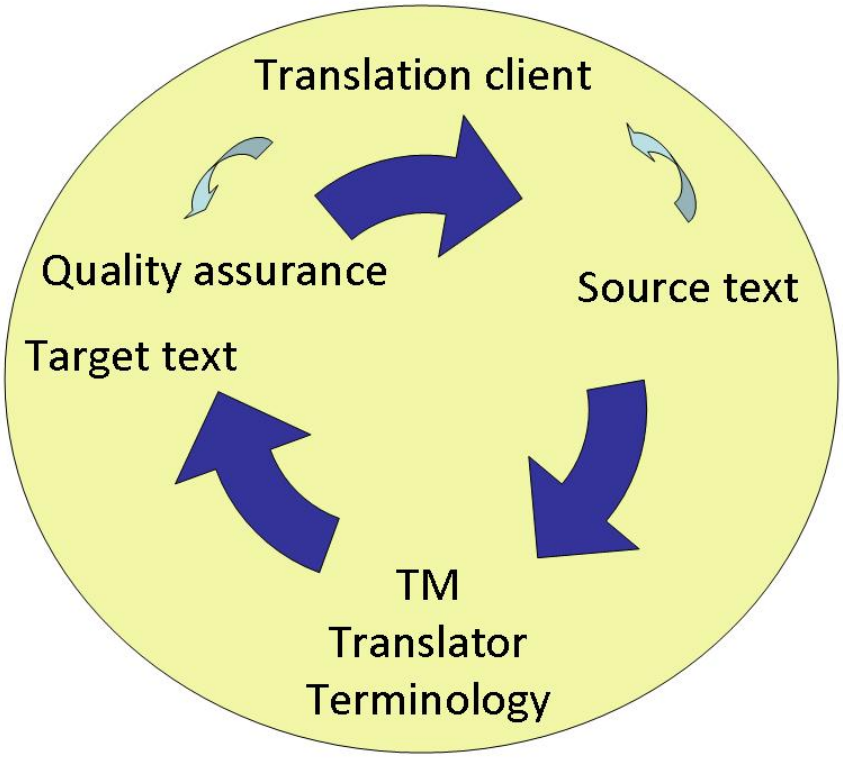


Figure 6.1 Computer-Aided Translation Workflow

Interview responses on the translation process are grouped into sections on TMs, TM tools, translators, and terminology management. Thereafter we return to expectations of the translation client, this time for target texts, and canvas the interviewees on the phenomena found in the quantitative study such as source

text format retention, target text term or phrase inconsistency, and target text explicitation. We also consider where inconsistency may be desirable in target texts and collate current methods of quality assurance. We finally present our interviewees' wish list for TM tools in section 6.6.

6.2 Qualitative Interviews

In this section we present profiles of the interview participants. In section 6.2.2 we look at coding of interview data and introduce themes that we drew from the data. We also signal which interview questions were being responded to when those themes emerged. Details of the research design for this phase are in section 4.7. Full transcripts of the interviews are in Appendix D.

6.2.1 Participant Profiles

Thirteen people were interviewed during this phase: five translators, two quality assurance (QA) specialists, two project managers, one workflow manager, one software localisation engineer, one language technology consultant, and one chief operating officer (COO). A list of interviewees, their job titles, and their mother-tongue language can be seen in table 6.1. Interviewees A, B, and C described themselves simply as translators. Interviewee D works primarily as a translator, but also does some project management (PM) and has started up an agency for video game localisation. Interviewee E does translation, review, and some PM work for a large multinational company. One of the QA specialists described his job as “linguistic QA” for a medical device translation company, the other works as a third party on behalf of the final client. Interviewee H oversees the choice of tools used per project and per translator for a language service provider (LSP).

	Job Title	First Language
Interviewee A	**Translator**	**Brazilian Portuguese**
Interviewee B	**Translator**	**German**
Interviewee C	**Translator**	**Spanish**
Interviewee D	**Translator**	**French**
Interviewee E	**Translator**	**Brazilian Portuguese**
Interviewee F	**QA Specialist**	**English**
Interviewee G	**QA Specialist**	**Spanish**
Interviewee H	**Language Technology Consultant**	**English**
Interviewee I	**Project Manager**	**German**
Interviewee J	**Project Manager**	**Japanese**
Interviewee K	**Chief Operating Officer**	**Spanish**
Interviewee L	**Workflow Manager**	**French**
Interviewee M	**Software Localisation Engineer**	**Spanish**

Table 6.1 List of Interviewees

Interviewee K looks after operations and production for a small LSP, which she says encompasses project management, project evaluation and preparation, resource management, and QA. Interviewee L is in charge of translation management systems at a multinational LSP, and interviewee M specialises in project preparation and analysis with various CAT tools. Six of the interviewees (four translators, a QA specialist, and a project manager) work on a freelance basis, the rest work directly for an LSP. Interviewees' first languages are Spanish (four people), Brazilian Portuguese, English, French, German (two each), and Japanese (one person).

Six interviewees said that the TM tool that they use most often is SDL Trados (of those only two use the more recent Studio 2009 version, the rest use SDL Trados Translator's Workbench (TWB) 2007 and one uses both SDL Trados TWB 2007 and SDLX), four use the latest version of MemoQ (5.3), one uses SDL Worldserver, and two use proprietary tools that are only available within the company for which they work. Nine interviewees are female and four are male. Interviews with C, D, E, F, G, H, and I were via a Skype voice-over-IP phone call. The remaining interviews took place mostly in private offices in the participants' workplaces. Each participant was given a plain language statement explaining the research and signed an informed consent form as approved by the DCU Research Ethics Committee (see section 4.7.2 and appendices A and B).

6.2.2 Coding of Interview Data

The audio interviews were between 16 and 59 minutes in length and were transcribed, imported into NVivo qualitative analysis software (see section 4.7.5) and coded, initially per interviewee and per interview question (IQ). As per the themes identified in section 4.7.3, interview questions were:

IQ1. What is your job?
IQ2. What TM tool do you use, and what version of that tool?
IQ3. What do you consider the benefits and disadvantages of TM?
IQ4. You mentioned (or did not mention) consistency. What is the effect of TM on consistency?

Chapter 5 showed inconsistencies in our case study TMs. The intent of this question was to find out whether our participants felt that TMs were inconsistent, addressing RQ1, whether they felt that TMs encourage TT inconsistency, and whether they found inconsistency a problem, in an effort to answer RQ3. (See Chapter 4 and section 6.6)

IQ5. Have you come across source text (ST) inconsistency?

Scripted prompt: For example, letter case, punctuation, or script inconsistency?

IQ5a. If the answer to 5 is 'yes': Do you find ST inconsistencies frequently, and do you have processes to deal with ST inconsistency? What effect do these have on target texts?

IQs 5 and 6 are intended to investigate whether our interviewees believe that our findings are externally valid (see section 4.3.2). If interviewees required further clarification, the scripted prompt was used as per section 4.7.2, based on the findings in Chapter 5.

IQ6. Have you seen similar examples to any of the following (see examples 4.1 to 4.7)? What do you consider their effect on translation quality?

This question addresses sub-segment inconsistencies which are discussed in Chapter 4: target text (TT) noun or term inconsistencies (6a); whole phrases alternated in TT (6b); ST format retention (6c) and TT explicitation (6d).

IQ7. How could TM tools be improved?

This final question looks to the future of TM tools, and addresses RQ5.

While coding and re-reading the interview transcripts, recurrent themes emerged. In the open phase of coding, these were added to nodes (NVivo terminology for codes) in a folder called 'Open Codes', and then re-organised into a branching tree structure as specified in section 4.7.5. These themes did not emerge from responses to single questions in most cases. Figure 6.2 shows the themes that emerged and links them to the associated interview questions. The blue text boxes represent the interview questions arranged clockwise. The beige text boxes represent the emerging themes.

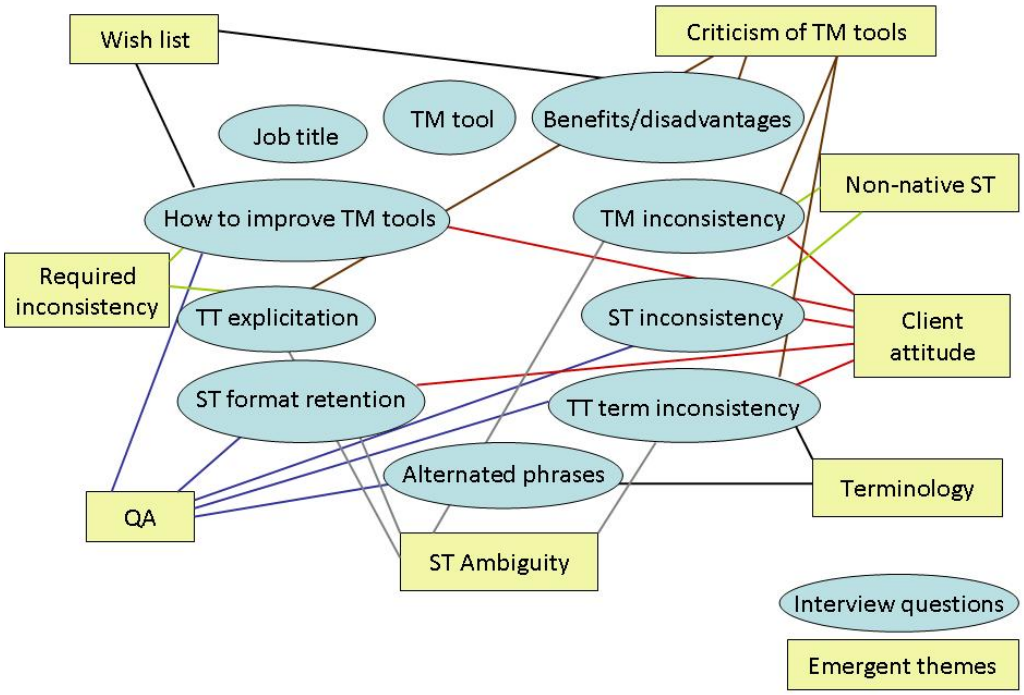
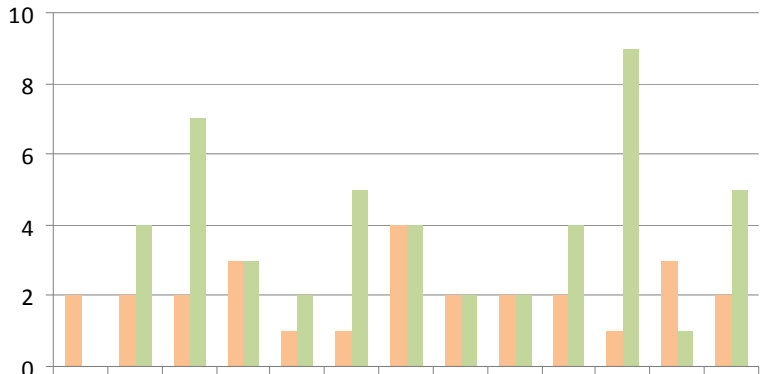


Figure 6.2 Relationship of Interview Questions and Emerging Themes

In general there was a high level of similarity in interviewees' responses. Small differences will be noted in the following sections. The results of coding of positive and negative comments about tools are shown in figure 6.3. Most of these comments were in response to IQ3, but criticisms arose in response to IQ4, IQ6a, and IQ6d, as may be seen in figure 6.2. Interviewee K, a COO with experience as a translator and a project manager, expressed the most negative opinions of TM tools.



	A Tra	B Tra	C Tra	D Tra	E Tra	F QA	G QA	H LTC	I PM	J PM	K CO	L WM	M SLE
Positive	2	2	2	3	1	1	4	2	2	2	1	3	2
Negative	0	4	7	3	2	5	4	2	2	4	9	1	5

Figure 6.3 Attitudes to TM per Interviewee

Several interviewees stated that clients' attitudes to translation had a great influence on source text quality and target text expectations. Interviewees who expressed frustration at limitations on time and money made available for translation, or complained of clients' lack of understanding of the translation process, tended to be those whose job was customer-facing. Interviewees who made the most references to clients or customers were M (Software Localisation Engineer; 36 references), K (COO; 28 references), H (Language Technology Consultant; 25 references), and L (Workflow Manager; 16 references). In contrast, only one translator (B) mentioned the client more than twice.

In the following sections we discuss interviewees' responses to questions and themes that emerged related to each stage of the translation workflow from figure 6.1, focussing on causes of and ways to minimise inconsistency, addressing RQ4 and RQ5. As translation clients are influential stakeholders at all stages of the process, we begin with source text and clients' attitudes to and influence on ST consistency.

6.3 Source Text

In Chapter 5 we found that TMs in our case study contain ST inconsistencies that are in turn associated with high rates of TT inconsistency (see table 4.17). In this section we examine interviewees' perceptions of ST inconsistency, and possible causes of ST inconsistency such as clients' attitude to ST creation, non-native written ST, and ambiguous ST, and find how interviewees deal with these causes of inconsistency.

6.3.1 Interviewees' Perceptions of ST Inconsistency

All but one of the interviewees (E) said, in response to IQ5, that they had seen many instances of ST inconsistencies as found in categories 1 and 2, detailed in Chapter 5. Interviewee G, a QA specialist, said that every time he checks a TM "I have the same translation for different source texts that should have been the same". He continued "the TM technology… the quality issue of technology is there, but the source text is the problem". Five interviewees said that inconsistent ST segments were, in their opinion, a cause of further TT inconsistency (again, in response to IQ5).

6.3.2 Clients' Attitude to ST

All twelve of the interviewees with experience of ST inconsistency said that they revert to clients with problems and one even tends to suggest a third party, a technical writing consultancy firm, to assist clients with their ST consistency. Eight interviewees suggested the attitude of some of their clients as a cause of ST inconsistency. They felt that the focus of clients is usually on time and cost savings, which means hurriedly written and inconsistent source text, and said that educating their customers about the effect of inconsistent ST on the translation process is one way in which they attempt to minimise inconsistency.

Several interviewees said that this education usually proves difficult. For example, K, a COO, commented that this is “one of the biggest fights that everyone in the localisation industry has to fight with clients” to make them understand that “if you don't control your English… you can't possibly expect to have savings and leverage over time with the TMs”. F, a QA specialist, agreed: “If they (clients) can't control their source text, then we can't be expected to control the target text for them”.

L (workflow manager) responded to IQ5, commenting that, while some clients pay his company to check or even rewrite the source, others want “the quickest turnaround at the lowest cost possible, so issues in the source text are completely ignored unless they cause a serious issue to the translators”. H and M also offer to “normalise” source material for their clients. All of the non-translator interviewees other than I and J spoke of a conflict, as far as clients are concerned, between cost and consistent ST. Those clients most concerned with making an initial cost saving choose not to control their ST, ignoring the extra time and cost later associated with this decision. Four interviewees (G, H, L, and M) have returned to customers with an assessment of the financial repercussions of ST inconsistency in order to “show just how much money they're wasting”, and to tell clients that “they could do things a hell of a lot cheaper and a lot better” (H; language technology consultant). G mentioned a test he had carried out comparing a company’s TM against an updated manual to demonstrate the financial impact of source text inconsistency, comparing 95-99% fuzzy matches and 100% matches. He found that “something like 80%” of the fuzzy matches should have been 100% matches, but the writers “had made little changes, a comma here or something there”. When he worked out how much these ST inconsistencies had added to the cost, his client was “amazed by the amount of money they would have saved” and the number of issues they would have resolved by controlling their ST.

L contended that what is necessary is “educating the customer, or at least educating the developers on the customer's side” to make them aware that the

content will be localised, and thus “the consequences of what they do in their source code before the project is translated”. Many interviewees said that, in addition to convincing their direct clients that controlled ST is beneficial, other parties on the client side need to understand the translation process, as the direct clients are usually not the creators of the ST. G said that it is “one of our main issues nowadays”, that “technical writers and editors are not aware of the impact they can have in translation”. In addition to the lack of controlled content from technical writers, K said that, in her experience, ST inconsistency can be caused by a lack of organisation, especially in big companies who do not maintain control over who creates their content. “There might be different technical writers working on different manuals for different products”, she says, “so even though there's a style guide and… guidelines for copywriting and all of that, it doesn't actually work”. This is not true of all customers. L sees more of his customers using content management systems to lessen this negative impact. These customers choose to “manage their source content, reuse their source content when they can, and be consistent with existing source content with their new material”, in order to maximise leverage.

6.3.3 Non-native ST

Aside from their clients’ attitudes, interviewees suggested several other causes of ST inconsistency. Six interviewees suggested that source text inconsistency is prevalent as a result of content written in English by non-native speakers. H estimated that “60% of the translations we do have English as source text but it's practically never written by a mother-tongue English [speaking] person”. She found that non-native ST tends to have inconsistent grammar and word choice. E (translator) regularly receives poorly-written English ST that has been written by Chinese authors. L receives a great deal of “Singlish” (Singaporean English) ST, adding that, depending on the customer, related source issues are either ignored, a “sanity check” on the ST is run, or the ST is completely rewritten. I and M both spoke of difficulties due to STs that alternate between British and US English.

6.3.4 ST Ambiguity

Six interviewees suggested that consistent but ambiguous ST may be a cause of TT inconsistency by being open to multiple interpretations. Reinke, cited in section 2.6.1, had previously identified an inability to cope with ambiguity as a limitation in current TMs. I (PM) and B (translator) have found problems due to ambiguity in ST that interviewee B said can be multiplied via auto-propagation proposing 100% matches in the wrong context (see also Rieche in section 2.7.1). K (COO) believes that ST ambiguity is a problem because “English, especially technical English, isn't always the clearest”. Her chosen solution is to search ST for ambiguities and to revert to the client for guidance. While she believes that some ST inconsistency is difficult to avoid, she said that it can be “minimised”.

G (QA specialist) also said that he regularly finds problems originating from imprecise ST. He has found that reverting to the technical writer does not always make things clearer. When asking writers which interpretation he should use, they often reply “it's this one, but actually it could also be the other one as well”. As a result translators “make different choices and sometimes those choices get out of hand and in time they are very difficult to fix, if you really want to achieve consistency”. In English, he has found it difficult to be “100% sure of what are the proper uses”, and has found himself “taking a long time to describe something [in Spanish] that in English is just two words”.

M (software localisation engineer) gives the examples of ‘setup’ and ‘settings’ as typical ambiguous ST segments. His solution is to provide translators with as much context as is possible such as “screenshots or the build in English so they can see exactly where it's coming from” and believes that if translators are “experienced enough in UI files” they will translate appropriately.

H finds the capability of MemoQ to use bilingual corpora as TMs useful in clearing up ambiguities in source texts and viewing context “so that we can really make an informed decision and decide whether that really was better in that

context or whether it should have been different". ST ambiguity was suggested as a cause for many of the TT inconsistencies in section 6.5, particularly in the case of target text explicitation.

6.4 Translation

Once the source text has been created, the next step in our simple workflow model in figure 6.1 is the translation process. In this section we discuss TMs and TM maintenance, followed by the interviewees' interpretation of the evolution of TM tools and some problems with matching, tags, and formatting. In section 6.4.3 we consider the effects of translators and their translation decisions, and the emergent theme of terminology management appears in section 6.4.4. Focus is initially on whether TMs are consistent, addressing RQ1, and on causes of inconsistency and ways to minimise inconsistency, addressing RQ4 and RQ5.

6.4.1 Translation Memories

In responses to IQs 3, 4, and 7, interviewees spoke about TMs. In answering IQ3, about benefits and disadvantages of TMs, all interviewees mentioned cost savings, time savings, and leverage as benefits. Five interviewees listed consistency as an advantage, following the often repeated axiom about TM, but two considered the introduction of inconsistency a disadvantage. Including answers to the other IQs above, nine interviewees said that TMs become less consistent over time, suggesting that their answer to RQ1 – are TMs consistent? – would be 'no'. G (QA specialist) said that "from a theoretical point of view the TM should improve consistency" but added that he has seen "over and over again" in projects using several translators and over several iterations that "consistency is not there". K (COO) believes it is "impossible to have a fully consistent TM", and thinks that nobody in the industry has "found a way of eliminating duplicates or inconsistencies in a TM".

This introduces the issue of TM maintenance, which was brought up by nine interviewees (B, C, F, G, I, J, K, L, M). B (translator) said that TMs get "inconsistent over time because they go through many hands", and that "not everybody is

good in maintaining the TMs so sometimes, over time, they get polluted". The result is that, as suggested by Bowker in section 2.7.1, TMs inevitably become 'dirty', i. e. filled with inconsistencies, and require maintenance to clean them up. F (QA specialist) said that cleaning TMs retrospectively is "much too expensive to do", aside from the difficulty of editing an active memory that is being "added to and leveraged from on an hourly basis". K said that to keep TMs in perfect shape is "a huge, huge effort".

Seven interviewees mentioned QA tools that they use for TM maintenance: three (G, H, and M) use ApSIC Xbench, two use QA Distiller (K and M), and two (F and I) use internal proprietary tools. Other than assiduous maintenance of TMs, some interviewees suggested that inconsistencies can be minimised by controlling the translation process. I (PM) said that while "ten glossaries, no penalties, 15 TMs" may result in a lot of leverage, she and interviewee K both stressed the benefit of small, targeted TMs that provide a low number of relevant suggestions, in effect prioritising precision over recall (see Bowker, section 2.7.1).

6.4.2 TM tools

Interviewees' answers to IQ3 and IQ7 revealed various frustrations with TM tools. They cited dissatisfaction with the evolution of TM tool functionality (8 interviewees), with matching and segmentation (5), and with how tools deal with tags in ST (7).

Evolution of TM tools

Interviewees felt that there has been a lack of useful development of TM tool functionality (see section 2.2 on the history of TM tools) and that a lot more could be done within the tools to minimise inconsistency. This can be seen by the number of interviewees who discussed their requirement for separate QA tools.

Instead, the focus from developers appears to be on workflow and technical improvement of the TM system (see discussion of Lagoudaki 2008 in section 2.5). Amongst the interviewees, translators and non-translators were dissatisfied by tool development in different ways. G (QA specialist), for example, felt that rather than changing how the tool works with data, developers are changing the workflow and he says “actually I haven't seen that much development” (see section 6.7). This was typical of the comments of non-translators. Most of the SDL Trados users cited similar reasons for not upgrading to the more recent Studio iteration of the software. K (COO) said that developers may have “put bells and whistles around it (the software)”, but compared to the older TWB functions are “pretty much the exact same”. She feels that the TM tool has become a translation management system, “whereas it should really, truly, focus back on consistency, language quality, and really incorporating and integrating all of those other functions”.

In section 2.5 we also discussed how TM technology has been said to serve industry goals rather than users’ needs. This theme recurred particularly when speaking to translators in the current interviews. In answering IQ7, C (translator) recalled when she used to work for a TM tool developer. She said she would “give the developers feedback and comments” but that “they can't really see it… they haven't translated *ever*, so they don't know the problems you encounter or the things you would like to see”. As a result, she finds that TM tools show too much focus on “superfluous” design aspects, such as colours and options, whereas the translation interface leaves little room for context. She says that “if translators had the knowledge and the capacity to develop the tools as they use them, no doubt they would be better”.

B spoke about the problem of interoperability as another reason for not upgrading or changing TM tools. She cited a loss of leverage when moving between tools, and the effect that this can have on consistency, explaining that “sometimes you migrate resources from one version to another within the same

tool, even then you can have a loss", and developers "wouldn't acknowledge [it], but it happens".

Matching and segmentation

In section 6.3 the interviewees agreed that ST segments with minor inconsistencies often evince further inconsistency on the TT side, as seen in Chapter 5. Two ways of preventing such TT inconsistencies would be by controlling the ST or by ignoring certain minor inconsistencies in the TM tool. K, in response to IQ4, said that if ST segments are "even one space different, one comma different, it (the TM tool) won't pick them up as being identical", rather "it will still show up as a different segment". However, eight interviewees (including K) said that many 100% match suggestions are erroneous. E (translator) said that, when a glossary has been updated, the TM may have the "old style and old terms, so you have to review the 100%. You can't trust 100%". K concurred, "You might have stuff in the TM, but really, you know, they're not 100% matches". The interviewees also find that, in the case of fuzzy matching, translators may "over-correct" (F, also mentioned by G, I, and J), or spend more time editing a fuzzy match than it would take to translate from scratch. This view concurs with research reviewed in section 2.6.2 that found that, below a certain fuzzy match threshold, TM matches cease to be beneficial.

Several interviewees also cited problems with segmentation. L (workflow manager) finds segmentation limiting for translators, removing flexibility and context. D (translator) and M (software localisation engineer) said that segmentation rules are designed for English and are not always appropriate for other languages. D has had problems with French language segmentation and wishes that the software would "recognise more signs or ways to segment the text grammatically depending on the language." She said that French text may have an "unbreakable space before exclamation marks, question marks", which often leads to a segment containing "the mark alone, without the rest of the

sentence". C (translator) has had difficulties with incorrect segmentation at URLs, tags, and at other types of non sentence-ending punctuation (see section 6.3.4).

Tags and format issues

In Chapter 5 we found a high rate of category 1 tag inconsistencies in each of the four TMs (particularly in TM D). Six interviewees said that they found dealing with tags in ST was a problem for TM tools. B (translator) said that "sometimes due to difference[s] in tags you get a different leverage" and that this is exacerbated by tags being handled differently by different tools. M feels that it is too easy to modify tags in TMs and C has had problems with tools disallowing the removal of unnecessary tags. She has sometimes been unable to convert the files back to the original format from her TM tool and as a result, "the only way you can sort that out is in the final document". This means that she has to leave unused tags at the end of segments and leave instructions for engineers to remove them, which she finds unsatisfactory.

L also said that tags had caused inconsistencies in his TT when STs were incorrectly segmented at the tag, separating a single segment into two segments that "you can't reconcile… unless you can merge the two segments". B finds that when moving between TM programs, "every tool handles tags differently" (see sections 6.4.2 and 6.6). M said that his solution to problems with tags is to remove "unnecessary formatting" from the text if the client is agreeable. For him, a line break or hard return in the ST provides a more difficult problem to resolve, and has a "great impact because you're not gonna get any leverage". H complained that, although the latest version of MemoQ allows translators to merge segments "even if there's a hard return, which is one of the biggest problems with Trados and all the other tools", translators are unused to being able to merge the segments, so they still tend not to.

6.4.3 Translators

The use of TM is intended to minimise inconsistency in translation. However, in Chapter 5 we saw the results of different translation decisions, comments left in segments by translators, and explicitation or context added by translators. Some of the interviewees believe that, when a human translator is involved, errors are inevitable (E, G, K, L). Several suggested that the level of experience of a translator may make them more or less likely to over-adjust a fuzzy match. I (PM) feels that, rather than relative experience, it is down to the "discipline of the translators to be aware of that (the possibility of over-editing) and not to just go and change everything they get".

B, D (translators), J (PM) and K (COO) agree with H (Language Technology Consultant) that TM "needs to be used correctly by the translators and this is another one of the big problems". B believes that many translators "aren't prepared to invest money and quite often aren't also even prepared to invest the time to learn how to use [TM tools] properly". In answering RQ5, most of the suggested methods of minimising inconsistency involve essentially removing any element of random choice from the computer-aided translator, such as in controlling ST (section 6.3), managing terminology (section 6.4.4), and the use of style guides (section 6.5.2). The increasing integration of MT proposals below a fuzzy match threshold may be seen to limit translators' decisions further. Despite this squeeze on the influence of the translator, only two interviewees, C (translator) and L, commented that, when using TM to translate, "the creativity of the translator is a bit reduced".

6.4.4 Terminology management

Nine of the thirteen interviewees highlighted the need for terminology management as a method of minimising inconsistencies in response to IQ6a. Terminology management is a function of most current TM tools, although we

still found TT term inconsistencies in the study in Chapter 5, which our interviewees all said are commonplace in their experience of TMs (see section 6.5.3). M warned that the use of TM without glossaries means "you're going to end up with inconsistencies, mistranslations, and some [other] issues". F (QA specialist) uses glossaries as his company "can't afford to care about what happens to the syntax and the grammar and everything else", rather they just "have to focus on terminology being consistent because that's what the clients really care about". H (language technology consultant) said that her clients, on the other hand, pay insufficient heed to terminology, and she considers it a "gift from heaven" if she can get "20, 50 words that … you guarantee in your documentation. According to her, "customers are not interested, don't appreciate, do not understand; they're not willing to pay for terminology work".

This discrepancy in clients' reported interest in terminological consistency is likely to be a result of the different domains in which these interviewees work. H works for a large LSP, while F works for a medical device translation company, a domain for which G (QA specialist) explained accuracy is a "regulatory compliance issue". J (PM) works in game localisation, where the focus, rather than on accuracy, is on keeping up to date with trends. In the case of an upgrade, it is mostly important "to adjust the user environment for new trends in the market". G believes that this reliance on terminology management demonstrates the inadequacy of TM tools.

> "That's why we still use glossaries. If TMs were really working as a means of achieving consistency, we wouldn't need glossaries 'cause once you have a term in the TM it should be used over and over, but really really that's not the case because of the way the industry works".

B (translator) said that it would improve a TM tool to "enrich it with terminology". This is largely already the case, but she feels that the relationship and interoperability need to be "more precise… closer together with terminology tools". D (translator) explained how, as designated terminologist on a project,

she would deal with multiple terms in a terminology list and prevent propagation of more than one TT term.

> “What I do is mark that in the termbase as a forbidden term and then the other person cannot use it anymore. (Laughs) It's the solution I found for this kind of thing... Having to fish that out of thousands and thousands of words? Yep. Forbidden term. End of the story”.

6.5 Target Text

The translations process culminates in the production of the target text (figure 6.1). In this section, we discuss clients' expectations of a delivered translation and whether they appear to find consistency important or desirable. Section 6.5.2 is focussed on ST format retention, a repeated cause of inconsistency in the quantitative data. Interviewees' views on further inconsistencies from Chapter 5 and their causes follow, such as TT term inconsistencies, phrase inconsistencies, and TT explicitation, focussing again on ways to minimise inconsistency, addressing RQ4 and RQ5. Interviewees struggle with inconsistency, yet also find occasions where inconsistency is required. We end this section with the emergent themes of required inconsistency and quality assurance.

6.5.1 Clients' Expectations

In section 6.3.2 we discussed translation clients' attitudes to ST and the related effects on consistency. Completing the workflow model in figure 6.1, we now come to clients' expectations for target text and the related effects on consistency. The consensus amongst customer-facing interviewees is that most clients are ambivalent about consistency. Five interviewees said that their customers are interested enough to use TM, but their main aim is to save time and money. M (software localisation engineer) said "it seems more and more that clients are not bothering much about consistency". In response to IQ4, he said that, for incorrectly suggested 100% matches, the clients say "don't worry, don't bother reviewing it… if it was wrong before we don't care if it's wrong now". In comparison, he says that roughly 30% of his clients consider each step of the process carefully and give specific instructions for the final translations. This extra care is reflected in the provision of better quality translations, he says, leaving both parties with greater satisfaction.

L (workflow manager) said that he thinks that some customers pretend to understand the translation process less than they really do. For them “it’s a game” to “try to get the highest quality at the lowest cost possible”. K (COO) said that some of her clients do not understand that 100% matches could be incorrect and “specifically request that anything that has already been translated before, they won't pay for it”. When she has tried to charge a little for 100% matches to minimise inconsistency, her customers have used machine translation as a threat, replying “No way! We'll get a machine to do it”. With this threat in mind, if she could “really achieve that maximum consistency through a single tool and we could pass on those savings to customers”, she said “I don't think they would even look at machine translation”. At present, she and the other interviewees feel that TM tools do not offer “maximum consistency” for segments or terminology.

6.5.2 ST Format Retention

IQ6c was about retention of ST punctuation or letter case in the TT, even when it may not be appropriate. In section 5.7, we noted that ST formatting was inconsistently retained in the TT in our case study. IQ6c was intended to ascertain whether this finding is externally valid. Interviewees were shown the following examples of ST format retention, wherein formatting or punctuation used in the ST was retained in one translation unit in the TT despite it being inappropriate for the target language, and not in another instance. They were asked about the effect of this inconsistency on translation quality. Example 6.2 is repeated from Chapter 5 (example 5.54 from TM D category 3).

6.1.1s *(Do Not Add to Workspaces or Add to Woprkspaces)*	6.1.1t *(Zu Arbeitsbereichen Nicht Hinzufügen oder Zu Arbeitsbereichen Hinzufügen)*
6.1.2s *(Do Not Add to Workspaces or Add to Workspaces)*	6.1.2t *(Zu Arbeitsbereichen nicht hinzufügen oder Zu Arbeitsbereichen hinzufügen)*

6.2s *Accesses Dimensioning mode* 6.2.1t 寸法記入モードにします

6.2s *Accesses Dimensioning mode* 6.2.2t 寸法記入モードにします。

Twelve of thirteen interviewees had seen this inconsistency frequently. I (PM) suggested that a formatting issue such as the one in segment 6.1.1t occurs when no match is suggested, in which case the translator might copy the ST to the target segment and overwrite it without considering whether the formatting is "actually compliant with German rules". She said that she would deal with this by compiling a style guide specifying what formatting to use, and that "that's something you have to clarify up front, even if you have the best TM ever".

Seven interviewees said that they use style guides. M (software localisation engineer) highlighted the importance of ensuring correct grammar in the TT, saying that along with "terminology and TM", TT grammar rules are top priorities for him. Despite this, inconsistencies continue to occur in the TT. E (translator) recalled a scenario where she was given different TT style instructions for different parts of a translation job. She was also told not to change any 100% matches even if they were inconsistent with the style guide, leaving inconsistency. A (translator) explained that style guides sometimes enforce rules that she finds counterintuitive.

> "There is a particular company that, at least for Brazilian Portuguese – actually for other languages as well, requires the translators to use the same capitalisation as is used in English even though this does not follow the grammar rules for Portuguese. It's for standardisation, for simplifying the process".

K (COO) believes that style guides are becoming less popular. She said that they were "very widely used up until five years ago", until "all of a sudden" a decision was made to "automate everything". Now, she said, "everything has been nearly left to the tool, but the tools don't actually catch things like that". She continued that QA tools would "probably catch something like that (inappropriate target

text formatting) but might not always", leaving it down to a "very labour intensive" manual quality control check to prevent inconsistencies in the final target text.

6.5.3 Target Text Term Inconsistency

IQ6a focussed on noun or term inconsistency in target texts. In Chapter 5 the largest proportion of TT sub-segment inconsistency in each of the four TMs was noun inconsistency. As a percentage of the total number of category 3 sub-segment inconsistencies, the lowest rate of noun inconsistency was in TM D (35%) and the highest was in TM C (49%).

Interviewees were shown the following examples of TT term or noun inconsistency. Example 6.3 is repeated from example 5.11 in the previous chapter and example 6.4 is repeated from example 5.53.

6.3s *All lines that have been converted using the {1}Create surface borders{2} function can be recognized easily since they are drawn with the {3}Border{4} pen.*	6.3.1t *Alle Linien, die mit der Funktion {1}Flächenränder anlegen{2} konvertiert wurden, k*önn*en Sie leicht erkennen, da sie mit dem Stift mit der Bezeichnung {3}Border{4} gezeichnet werden.*
6.3s *All lines that have been converted using the {1}Create surface borders{2} function can be recognized easily since they are drawn with the {3}Border{4} pen.*	6.3.2t *Alle Linien, die mit der Funktion {1} Flächenränder anlegen{2} konvertiert wurden,* könn*en Sie leicht erkennen, da sie mit dem Stift mit der Bezeichnung {3}Rand{4} gezeichnet werden.*

6.4s *A new layer group filter can be nested only under another group filter.*	6.4.1t 新しいレイヤグループ フィルタは、他のグループ フィルタに対してのみネストできます。
6.4s *A new layer group filter can be nested only under another group filter.*	6.4.2t 新しい画層グループ フィルタは、他のグループ フィルタに対してのみネストできます。

All 13 interviewees agreed that this was a phenomenon that they saw often. E (translator) suggested that these inconsistencies may be caused by having several translators working on the same material without a follow-up consistency check to “catch this before it is propagated to the TM”. C (translator) suggested that this problem may be caused by merging TMs from different sources. She often receives TMs that contain inconsistent terminology, so that “when you look for a term in a translation memory, you do a concordance search”, in which case she often finds “two, three, or four different translations” for the same term.

L (workflow manager) suggested that inconsistencies such as in example 6.3 could emerge even with terminology databases as a result of the conflict between approved terminology and search engine optimisation (SEO). In this situation, a German translation for ‘border’ may have been “approved and reviewed”. Despite this, the client may have realised that “in German, people don't actually go to google.de and look for the German translation, but they actually look for the English words”, so when they start optimising their German website, they might decide that “the approved term is actually not what's going to get them the hits.” This is one explanation for the prevalence of noun inconsistencies such as in example 6.3, featuring a native German word in one instance and a borrowed English word in another.

Non-translator interviewees felt that new or inexperienced translators on a job tend to add new translations to the memory (see also section 6.4.3). K (COO) said that translators may also choose to accept inappropriate matches. If inconsistencies exist in the TM, she continued, one cannot expect consistency to be increased in the TT via "human decisions". Three interviewees said that they had experienced problems with inconsistencies such as in example 6.4 in Japanese TMs, whereby in one case a kanji compound was used and the phonetic katakana loan word was used in another. J (PM), a native Japanese speaker, felt that "the trend is now phonetical translation, so that's inconsistency just depending on the translator preference, so we have to give them guidelines, style guidelines".

Nine interviewees believe that inconsistencies such as examples 6.3 and 6.4 are the result of insufficient terminology work (see section 6.4.4), leaving the translator with a lack of guidelines and glossaries. D (translator) said that "bigger projects" require a dedicated terminologist who "gets paid more… to correct these kind of entries because in a huge TM you cannot go through the whole TM, it takes too long".

M (software localisation engineer) returned to the theme of client attitude when answering IQ6a. "For the client, if they say "leave it as it is", for us it's a 100% match, there you go. As a translator I would say, of course, this is not right". Thus, two methods of minimising term inconsistency are terminology management and client education.

6.5.4 Target Text Phrase Inconsistency

IQ6b concerned inconsistent whole phrases that had been propagated throughout TMs in Chapter 5, such as in example 6.5 (repeated from example 5.31 in section 5.3.2) and example 6.6 (repeated from example 5.26 in section 5.2.3). The TT phrases in example 6.5.1 both appeared in the TT aligned with

consistent (or almost consistent) ST 219 times in TM C - 126 times in category 1 (as below, with inconsistent ST) and 93 times in category 3 (with consistent ST). These examples were shown to all interviewees.

6.5.1s *At the Command prompt, enter subtract.*	6.5.1t *Geben Sie in der Befehlszeile differenz ein.*
6.5.2s *At the command prompt, enter subtract.*	6.5.2t *Geben Sie an der Eingabeaufforderung DIFFERENZ ein.*

6.6s *Binding ISO Elements to XML Elements using Object Info*	6.6.1t オブジェクト情報を使用して ISO エレメントを XML エレメントに拘束する
6.6s *Binding ISO Elements to XML Elements using Object Info*	6.6.2t オブジェクト情報を使用して ISO エレメントを XML エレメントに関連付ける

11 of the 13 interviewees said that they often find similar inconsistent phrases propagated in TMs. K said that she sees these kinds of inconsistencies “day in, day out”, where there is “a new member of the translation team who will reckon he or she has a better solution”. If this new team member ignores the suggested match and rewrites the TT segment, “it's just so easy to bring in or upload a new version onto the TM and there's nothing that stops it”. F (QA specialist) also said that he sees this “all the time” and suggested that it may be the result of translators working independently without adequate terminological guidance. Also, with the software that his company uses, for fuzzy matches “it’s up to the translator to identify where that fuzziness is and correct it, and sometimes they just over-correct”. Despite the importance of consistency in his domain of medical device translation, he says “we wind up with this all the time”.

Two translators (C and D) suggested that the inconsistency in example 6.5 may have been caused by auto-propagation and not chosen by a human translator. G

(QA specialist) said that he often finds inconsistencies such as these caused by differences of interpretation between translators as discussed in section 6.4.4. J and K also remarked that a lack of clarity in the English ST may have been a cause of the inconsistency in example 6.6. Both considered that 'Binding ISO Elements' could mean 'binding of ISO elements' or 'ISO elements that bind'. K said: "In reading it at first it was clear to me what it meant, but having gone through the two options in Japanese, it's not clear any more what the English actually meant". As in section 6.5.3, interviewees' suggestions for minimising TT phrase inconsistency were tighter terminological control and a reduction in ST ambiguity.

6.5.5 Target Text Explicitation

IQ6d related to target text explicitation, a phenomenon seen particularly in Japanese TT in Chapter 5. Explicitation in translation is considered by some theorists to be a universal characteristic in translation, and can take the form of addition, adding new information, or specification, giving "more specific information" (Englund Dimitrova 2005, p34). Becher, however, does not believe that explicitation is "translation-inherent", but suggests that where there is explicitation there will always be a reason, not least to "minimise the risk of misunderstanding" (2011, p215/216).

Example 6.7, which was shown to all interviewees, specifies what is being selected in each TT segment and is repeated from example 5.24 in section 5.2.3 (see also translations of the word 'generation' in example 5.25).

6.7s *Selecting*	6.7.1t エレメントの選択
6.7s *Selecting*	6.7.2t コールアウトエレメントの選択
6.7s *Selecting*	6.7.3t アセンブリの選択
6.7s *Selecting*	6.7.4t 多角形の選択
6.7s *Selecting*	6.7.5t 線の選択
6.7s *Selecting*	6.7.6t 楕円の選択
6.7s *Selecting*	6.7.7t 選択
6.7s *Selecting*	6.7.8t 長方形の選択
6.7s *Selecting*	6.7.9t ベジエ曲線の選択
6.7s *Selecting*	6.7.10t めねじの選択
6.7s *Selecting*	6.7.11t おねじの選択

Nine interviewees said that they saw this phenomenon frequently, not just in Japanese, but in German (B and I), Brazilian Portuguese (A), Romanian (H), Spanish (K), and Malay (K). K again felt that the ST in segment 6.7 is unclear (see section 6.3.4). She suggested that the translator needs more details as “it's critical that they [ST segments] are taken in context”. She also returned to the topic of client expectation (see 6.5.1), saying in her answer to IQ6d that customers need to be told that “they might be 100% match, but they're not perfect matches all the time”. Her policy is not to lock 100% matches to prevent new translations, but she tells translators not to touch 100% matches unless absolutely necessary. In the case of incorrect 100% matches it’s left to the assiduousness of her translators as to whether they revert to her with problems or leave the 100% matches untouched. If the context has changed, then the introduction of some inconsistency in TT may be necessary (see section 6.5.6).

M also said that the source text 6.7s is unclear and remarked that, on receipt of a similar ST segment, he would revert to his client to ask “what actually do you mean with this?”. His policy is to get as many details as possible from the client and, if the client agrees, to penalise ST segments with more than one 100% match so that none will pass without review. This source text ambiguity is a continuous problem for him. “You know, all they do is come with one word,

which works fine in English, but maybe when you try to translate that is when you have issues". He has found that target text explicitation such as in example 6.7 occurs frequently in Japanese as the language is "very challenging… because it's very very context-specific. We get a lot of queries from translators trying to say 'what's the context, why is this like that?'". He would like to see the 'perfect' or ICE (In-Context Exact) match technology perfected, so that he could see a 100% match that "actually is a 100% match and you don't have to worry".

K has also found a tendency toward target text explicitation in Spanish due to a tendency towards ellipsis in English ST. She said that "in Spanish we [are explicit about] everything, you know, everything, if it's 'selecting', it's 'selecting what'". G said that target text explicitation can happen in Spanish "where we need articles or gender, [if] we have adjectives". Accordingly, there is a requirement for some context to be explicit if an article is to be used so that the correct gender may be chosen, or the addition of an article may be necessary in what Englund Dimitrova calls an obligatory explicitation (2005, p34). The requirement for obligatory explicitation may lead to a situation where inconsistency is required.

6.5.6 Required Inconsistency

In this research we have measured inconsistency in TM data without consideration of whether this inconsistency is required or not. As seen in section 6.4.2, the interviewees would like "maximum consistency" (K), yet have problems with clients' assumptions that all 100% matches can be automatically accepted. In section 6.5.1, eight interviewees said that 100% matches may be erroneous, a point previously made by Reinke (2004; see section 2.6.1). Particularly when answering IQ6d (regarding TT explicitation), several interviewees (particularly non-translators) felt that some TT inconsistency may be necessary. F is "guided by my translators" as to whether it might be better to introduce inconsistency. He said that, for him, it is more important for a translation to be "accurate and natural and fluent than it is for the resulting

translation unit to be recyclable infinitely in all other documents", adding that this loss of leverage is "a sacrifice we have to make".

This divergence between consistency and what F calls 'accuracy' and 'fluency' means that, according to M, clients can make one of two choices with regard to translation using a TM: They can either have one ST segment to one TT segment and leverage with little or no editing where possible (requiring consistent and clear ST), or they can choose to allow multiple TT translations of a single ST segment, which he would prefer to penalise so that 100% matches would be reviewed. H said that allowing multiple translations of one ST segment "can create chaos", and that her translators complain that "it means that when I try to correct something I've done myself, it's not overriding it, it just puts a new one in".

J (PM) said "that's the difficulty" with TM, that it cannot get what she calls "the proper leverage", a balance of maximum leverage while preventing the acceptance of inaccurate 100% matches. In answering RQ3, minimising inconsistency means minimising this required inconsistency by making ST clear, consistent, and unambiguous by educating writers and developers on the client side that their work is to be localised. L also suggested that inconsistency may be minimised by a "fairly thorough linguistic review process" as part of quality assurance (QA), which we address further in the following section. He adds that "inconsistency is not necessarily something to be eradicated", and may be necessary "depending on the type of project, the type of customer, the type of content".

6.5.7 Quality Assurance

One of the themes that emerged from the interview data (see figure 6.2) was quality assurance. Although in section 6.4.1 seven interviewees said that they use QA tools for TM maintenance, QA tools were also discussed with regard to

checking ST consistency prior to translation and, especially, checking whether a translator has introduced inconsistency in the TT after translation. This final check may also be done manually by an assigned language QA person. M said that the "biggest impact" automatic QA can have for him is "doing the final checks before delivering the files". He asks translators, as does H, to "Xbench their files [before] they come to us". I (PM) said that automatic QA tools are important for her, because if a translator adds an unnecessary new TT translation it "may not be caught" at the translation stage, but "when we would do a consistency check with the tool, it would definitely be caught, so it wouldn't be a big issue".

Seven interviewees (D, E, F, G, H, J, and K) said that they work with (or are) a QA person who manually checks translations before they are delivered to clients. D (translator) said that, unless a person is designated to make the final decision on delivering work to clients "you get a huge mess", especially for "online projects" and when people are "working together on the same server for big projects". She added that "otherwise achieving consistency is almost impossible because it's kind of a personal choice".

K said that, despite QA, often the final translation is "an imperfect product". She said she has accepted that there is no such thing as "a 300% consistent, 300% perfect delivery when it comes to language".

6.6 Wish List for TM Tools

This section highlights features that interviewees would like to see improved in or added to TM tools. The section begins with their dissatisfaction with progress to date and their preference for the integration of improved quality assurance. Several interviewees would like to see changes to segmentation and matching, as well as interoperability.

6.6.1 Misdirected Updates

The author Neal Stephenson, bemoaning a lack of recent general technological innovation, writes that "the best an audacious manager can do is to develop small improvements to existing systems" within what he says at present is "a system that celebrates short-term gains and tolerates stagnation, but condemns anything else as failure" (2011, p16). As discussed in section 6.4.2, this accurately describes how most interviewees feel about innovation in TM tools. G (QA specialist) said that in recent attempts to improve TM tools, "they [TM tool developers] just make them a lot more difficult", and as such is one of several interviewees who sees no benefit in keeping up with the latest iteration of tools. He also believes that "when you focus too much on the tool itself, you forget about the bigger picture". K (COO) was also unimpressed with recent updates to TM tools. She felt that "the progress and, I don't know, the "improvements" so to say, have been in the wrong areas, which is just the management systems and the file processing." L (workflow manager) said that the tools could fit in better with workflow technology, but that "the algorithm behind it (the TM tool), the context matching, the quality control, I think, are all very good".

6.6.2 Quality Assurance

Interviewees suggested several additions to tools for minimisation of inconsistency. Four interviewees (G, H, K, and M) would like improved integration of QA functions, with K favouring the integration of "a QA Distiller-type tool." M (software localisation engineer) would like to see a concordance check for each segment to "self-validate a TM". He currently does this using Xbench, but only after exporting his TMs to the TMX format and adjusting settings in Xbench. Using such a tool, he said, it would be easier to decide "OK, this TM is really inconsistent" and as a result to choose to spend more time working on the TM "rather than just finding the issues when you're doing the translation". He finds this type of functionality increasingly necessary as he has to deal with "more complex languages", citing examples of Hindi and Marathi. E would like to see a "mechanism of checking that there is only one source and one target". K believes that there is a real need to improve the performance of TM tools with regard to consistency, and says that whatever tools can incorporate "minimising inconsistency issues like [those in this study] will definitely be the winner… in the industry".

6.6.3 Segmentation and Matching

D (translator) and M would like to see segmentation improved. D wishes the software would "recognise more signs or ways to segment the text grammatically depending on the language" citing problems she has had with punctuation in French. M complained that "all the settings are based on English and [the assumption that] the target language is going to be very similar to English, which is not the case and it's not very easy to customise". He would also like to be able to expand or shrink segments outside of the TM tool's editing environment, adding that "it would be very helpful for some languages".

F (QA specialist) and M would like to see improvements in fuzzy match classification. F said there is a "need for fuzziness to be more accurate, the

calculations to be more explicit, but they are in some tools and in other tools they aren't". Three interviewees (A, B, and J) said that they would like to see improved sub-segment matching in order to "see more detailed matched text" (J). M said that, with more file formats and new tag types to deal with, the current fuzzy or 100% match classification has become outdated so that only context matches are trustworthy, and that analysis based on possibly inaccurate 100% or fuzzy matches, not to mention pricing based on that analysis, is often skewed. If the system was changed, he believes it would lead to "more realistic expectations" and be "more transparent for translators and also for us (LSPs) and for the client". He would also like to see the options for TM maintenance improved, as he finds the current options difficult and "primitive".

6.6.4 Other Suggestions

C (translator), K, and J (PM) think that consistency would be improved by simplifying tools and thus lowering the learning curve for translators when beginning to use TM tools. J said that current tools are inflexible and constrained by licensing methods that hamper accessibility. F said that updates are too slow, and that his company has resorted to building an in-house tool so that changes may be made as necessary. B and K would like to see improved interoperability (see section 6.4.2). K said that currently she finds that if she does an analysis on "a simple file", then someone else analyses the same file using the same version of the same TM tool, "you come up with two different analyses". This, she said, could be down to differences in settings, or problems of interoperability, but can make translation companies seem inept at "a time where we really need to be efficient and translation is seen as an unnecessary commodity".

6.7 Summary

In this chapter we presented the results of the second part of our sequential mixed methods study, explanatory qualitative interviews. We profiled the interviewees and discussed coding before beginning a discussion of the interviewees' contributions, organised to follow a simple computer-aided translation workflow. The interviewees added to the results of the quantitative study in Chapter 5 by sharing their experiences and opinions of inconsistency. We had found several examples of inconsistent source text in Chapter 5. In section 6.3 we discussed clients' attitudes to ST creation, ST written by non-native speakers, and ambiguous ST, and how interviewees deal with these causes of inconsistency. Section 6.4 was focussed on TMs and TM maintenance, particularly interviewees' problems with matching, tags, and formatting. They considered translators' translation decisions and terminology management as possible causes of target text inconsistency.

Following translation, we looked at clients' expectations of a delivered translation, clients' focus on time and money rather than consistency, and whether most translation clients consider consistency important or desirable. Interviewees then responded to inconsistencies found in the quantitative data such as source text format retention, target text term inconsistencies, phrase inconsistencies, and TT explicitation. They explained how they attempt to minimise such inconsistencies in their processes. We also introduced themes that emerged from the interview data such as required inconsistency and quality assurance.

The interviewees' familiarity with the inconsistencies found in Chapter 5 – and their identification of those inconsistencies as frequently-occurring issues for them – gives our findings in Chapter 5 a degree of external validity. In answering the interview questions, the interviewees also helped to address the research questions, particularly RQs 3, 4 and 5 (see section 7.2). Procedures to minimise

inconsistency included designation of QA specialists and terminologists, use of QA tools, and the prescribed use of glossaries and style guides for consistent TT formatting. Finally, having discussed interviewees' dissatisfaction with the development of TM tool functionality, we summarised their wishes for the future direction of TM tools, which include the integration of improved quality assurance, and improvements to matching and segmentation.

7 Conclusions

7.1 Introduction

In this chapter we present our conclusions, addressing the five research questions from Chapter 1 in section 7.2. In section 7.3, we begin to look to the consequences of the current research, providing a series of recommendations for users and developers of TM tools, and continuing with a cost model for a hypothetical localisation project to highlight the financial impact of inconsistency in TMs. Thereafter we discuss the contributions of the current research to the field of translation technology, and address some of the limitations of the research that may in turn suggest possible directions for future work leading from this research.

7.2 Addressing Research Questions

In this section we address each of our research questions from Chapter 4 in turn, referring to the results of our mixed-methods study.

7.2.1 Research Question 1

RQ1: *Are TMs consistent?*

Our hypothesis (section 4.6.1), based on the results of the pilot study in Chapter 3, is that they are not. In section 5.7 we saw that our four case study TMs are not completely consistent. Our interviewees confirmed that, in their opinions, our findings in Chapter 5 conformed to their experience. The first quantitative phase of the main study found that the four TMs were inconsistent on both source and target sides, and showed the types of inconsistencies common to these TMs. For repeated source text segments it was found that approximately 5% were translated inconsistently in each of the TMs. (The potential cost associated with this level of inconsistency is suggested in section 7.3.4). The most prevalent type of inconsistency was introduced noun or term inconsistency in category 3. Many noun inconsistencies also demonstrated influence from the source language (in this case, English). TT verb and punctuation changes were common throughout the corpora. In each of the TM corpora there were cases of inconsistent formatting, leaving a combination of ST-appropriate and TT-appropriate in the TM data. The English-to-Japanese TM data also showed evidence of explicitation in the TT.

RQ1 was based on the assumption in the localisation industry that TMs result in cost savings, time savings, and consistent translations, but according to our interviewees, translation clients tend to emphasise one or two of these benefits at a cost to the others. A focus on cost-savings will result in uncontrolled ST, lost

leverage, erroneously accepted 100% matches, and more time spent on editing the final translation. A focus on time-savings results again in uncontrolled ST and associated lost leverage, rushed translations, and insufficient QA, leaving a 'polluted' TM that is very expensive to clean. A focus on consistency involves putting time and effort into controlling ST, maintaining the TM, and concentration on QA. Our interviewees, depending on their clients, attempt to find a balance between these three competing emphases.

More consistency will result in higher leverage, but in section 6.5.6 we saw that there may be required inconsistency. In section 6.4.1, two interviewees suggested prioritisation of precision over recall, using small bespoke TMs that are job- and domain-specific as a way of minimising required inconsistency (see also Bowker 2005 in section 2.7.1). This would, of course, depend on other factors such as the consistency of the ST and behaviour of the translators. The interviewees tended to believe that the introduction of unnecessary inconsistency to TMs is inevitable.

7.2.2 Research Question 2

***RQ2**: How can consistency be identified and measured in TM data?*

Using the novel methodology outlined in Chapter 3, we have found, measured, and categorised inconsistency at the segment and sub-segment level in four sets of TM data. We applied this methodology to TM data in Chapter 5, and, when presented with our findings, our interviewees subsequently agreed that the type of inconsistency we observed in our TM data was familiar to them. This, we contend, serves to validate the methodology and also the findings from Chapter 5.

7.2.3 Research Question 3

RQ3: *In the opinion of translators and those working in the localisation industry, is inconsistency a problem?*

When asked generally about TM, only two interviewees cited inconsistency as a disadvantage, but all agreed that TT inconsistencies found in this study were commonplace and that inconsistency in TM is a problem. This confirms that the answer to this question is 'yes'.

7.2.4 Research Question 4

RQ4*: In the opinion of those same people, what causes inconsistency?*

Interviewees identified several causes of inconsistency, as discussed in Chapter 6. The client's attitude to translation was seen as an underlying cause of ST, TM, and TT inconsistencies. On the source text side, inconsistent or ambiguous ST, inconsistent ST formatting or tags, and non-native ST was said to cause TT inconsistency. TMs were found to become "polluted" and "inconsistent over time" due to factors such as: choices made by translators; permission for multiple translations of a single ST segment; inadequate management of terminology; and inadequate TM maintenance. Interviewees felt that TM tools may also cause inconsistency through incorrect segmentation, erroneous match suggestion, inconsistent handling of tags, and inadequate integration of terminology and QA tools.

These causes of inconsistency may be mitigated or at least influenced by actions on the translator's part (such as those suggested in section 7.2.5), but the interviews also revealed further factors that militate against consistency. Following trends means updated terminology is to be used, which interviewees highlighted as a constant source of inconsistency. Keeping language up to date is

difficult and particularly problematic in the domain of games localisation, where "the use of natural, fluent and idiomatic language is essential" (Mangiron Hevia 2006, p312). Interviewees spoke of having to adjust the user environment to cope with new trends in the market, such as the preference for currently popular colloquialisms, and, in the case of translation into Japanese, the preference for phonetic translation of some words. This could be one explanation for the influence of English ST evident in both German and Japanese TMs in this study.

Another possible reason for the use of English in the TT, and a factor that may cause inconsistency, is the need for search engine optimisation (SEO) of web-based content. Designers optimise web sites in order to appear at the "top of the results page when users search for particular keywords or phrases" (Cahill and Chalut 2009, p234). As popular keywords (those searched for by users) tend to be in English, English loan words in texts may lead to a higher page ranking, and thus appear higher in a page of Internet search results.

7.2.5 Research Question 5

***RQ5**: What procedures could be implemented to minimise inconsistency?*

Based on the findings in Chapter 5, standardisation of source text would minimise inconsistency, with particular attention to be paid to standardisation of terminology, punctuation, spacing, and tags. It could be seen from the metadata that some errors were propagated due to choices made by different translators on a project. Maintenance of the TM and of term lists would limit these choices and improve consistency in target texts, as would the use of style guides to retain consistent formatting of target text. Some suggestions to minimise inconsistency follow in sections 7.3.2 and 7.3.3.

Many of the interviewees have already put procedures in place to minimise inconsistency. In section 6.3.2, interviewees highlighted the importance of

educating clients and those on the client side who produce source content about the importance of controlling ST. Much of the ST that the interviewees currently receive appears to be written poorly (in their opinion) by non-native speakers of the ST language, lacks clarity, and contains unnecessary inconsistency.

Several interviewees said that they use QA tools (such as those in sections 2.7.3 and 5.6) to run checks on the source text, to maintain TMs, and to search for introduced inconsistency post-translation. They work with the same TM tool throughout a project to avoid problems of interoperability, and where possible use translators who are familiar with the project. Managers can constrain translator choice by either locking 100% matches or penalising multiple 100% matches, and by introducing style guides for consistent TT formatting. Most interviewees also said that they use glossaries or terminology tools in order to minimise inconsistency.

The level of post-translation editing, review, and quality assurance is dictated again by the client's instruction and willingness to spend time and money on minimising inconsistency. The client's focus on finance may increase consistency at a cost to fidelity to the ST if erroneous 100% matches are accepted when context would suggest that a new, albeit 'inconsistent', translation is required. The interviewees felt that more could be done within the TM tool environment to minimise inconsistency. This was discussed further in section 6.6.

7.3 Consequences of Research

In this section we look to the consequences of our research, in particular to how it might make a difference to the future use of TM tools or be applied in the development of a QA tools optimised to check for consistency. In section 7.3.2 we present some recommendations resulting from both phases of the current research, beginning with standardisation of source text and followed by ways to mitigate inconsistency when translating using TM. Section 7.3.3 contains recommendations relating to TM tool development. In section 7.3.4, we propose a hypothetical localisation scenario so that we might estimate the financial impact of inconsistency.

7.3.1 A Focus on Consequences

This current research adopts the pragmatic worldview, which insists not on precedents but "upon the possibilities of action" (Dewey 1931, cited in section 4.3). Mixed methods research based on pragmatism is said to look to consequences and the validity of research conducted within the pragmatic paradigm is judged on the effectiveness of the research and the actions it engenders once completed and disseminated (Kvale and Brinkmann 2009, cited in section 4.3). In this section, therefore, the focus is on consequences of this research and on dissemination of key findings. A series of recommendations for users and developers is thus contained in sections 7.3.2 and 7.3.3.

Thereafter, we use our findings to envisage realistic scenarios in which leverage is lost due to inconsistencies in the TM applied in a hypothetical job. This necessitates making a number of assumptions about the project, but the creation of such scenarios can help make the financial case for ensuring consistency (where possible) in TMs, despite the cost associated with doing so. If cost-benefit analyses based on scenarios like the one presented in section 7.3.4

can demonstrate that it is financially worthwhile to implement measures that promote consistency, then the major obstacle to doing so (clients' perception of cost) may be removed. Following on from this study, we have begun to collaborate with colleagues from the Centre for Next Generation Localisation to create a component for the Localisation Research Centre's SOLAS platform[36] that will measure and categorise inconsistencies based on the specifications developed in this research.

7.3.2 Recommendations for Users

Recommendation 1: Standardise Source Text

In Chapter 5 we saw ST inconsistency in each of the TMs and in Chapter 6 interviewees highlighted the importance of source text standardisation. All of the interviewees had experienced difficulties with inconsistent punctuation in source text segments, leading to fuzzy rather than 100% matches. Consistency of punctuation, formatting, spelling, and terminology was considered important. Chapter 5 showed the prevalence of ST space inconsistency, especially trailing end spaces and double spacing. These could be automatically removed if using a content management system or ignored if the settings on the TM tool can be set to discount end spaces when searching for matches. The ability to ignore not only trailing spaces, but further minor source text inconsistencies could have an impact on TM proposals, cost, and the possiblity of further inconsistencies being introduced.

[36] SOLAS (Service Oriented Localisation Architecture Solution) is a component-based localisation platform. More details are available from http://www.localisation.ie/solas/

Recommendation 2: Use Style Guides

The inconsistencies of formatting in Chapter 5 showed that style guides are important to retain target text consistency despite, as interviewee K claimed, an increasing tendency to leave format consistency issues to be solved by CAT tools. Aside from the use of style guides, there are several other choices to make before translating with TM that may affect consistency. One is whether or not to lock 100% match proposals so that a translator cannot edit them. This means retaining only one translation of a single source text segment, but according to interviewees "you can't trust" 100% matches. Interviewee M considers that ICE (in-context exact) match technology has not been perfected, but could be one way of verifying that a 100% match can be trusted.

Recommendation 3: Manage Terminology

The TMs in this study showed a high rate (40-50% of all TT inconsistencies) of introduced noun inconsistency, highlighting the need for a focus on terminology management. Interviewees were also consistent in their preference for terminology management. Maintenance of terminological resources and TMs is time consuming and difficult, but worthwhile if the resources are to be best exploited. This step, along with the update of TMs with final edited translations, appears to be often forgotten.

Recommendation 4: Maintain TMs

In the hypothetical localisation scenario in section 7.3, we suggested that the introduction of inconsistency has a financial impact, making it less worthwhile for recycling in TM and in other translation tools (such as SMT). For this reason, TM maintenance is important. TM tools alone were seen as insufficient for resource maintenance by interviewees and as a result, dedicated QA tools are an important part of the current translation workflow, providing a good overview of the consistency of a TM and providing a useable interface for maintenance.

Recommendation 5: Use Bespoke TMs

In Chapter 2, we mentioned that Bowker suggested using small, targeted TMs containing "well-chosen texts" to reduce inconsistency (2005, p19). This theme recurred in interviews, as interviewees felt that precision (i.e. few but accurate match suggestions) was more important than recall (many inaccurate match suggestions, or noise).

7.3.3 Recommendations for Developers

Recommendation 1: Improve Tool Interoperability

Interoperability has long been an issue for CAT tools users and in section 2.2.3 we highlighted the poor implementation of standards by tool developers. Interviewees were also dissatisfied with discrepancies between tools in segmentation rules and tag handling.

Recommendation 2: Enhance Integration of Terminology and QA

Interviewees in Chapter 6 were unhappy with the terminology and QA functionality in existing tools, as a result of which stand-alone QA tools were preferred despite problems of interoperability with TM tools. Several stand-alone QA tools also allow users to carry out an internal TM consistency QA check.

Recommendation 3: Involve Translators in Tool Development

Interviewees reiterated frustration with the development of functionality in TM tools evident in Lagoudaki's thesis from 2008 (see section 2.5.2). Interviewees felt that software upgrades added translation management and project management functions without improving noticeably how they deal with consistency and language quality. They also felt that the user interface in most

current tools is too cluttered and does not suit their needs. The application of user centred design, with improved affordance of commands (see section 2.5.2), may lead to an improved user interface, increasing usability and speed.

Recommendation 4: Permit Users to Add or Replace TM Content for Each Segment

At present the choice of whether to add to or replace matches in the TM is made at the beginning of a project, but if the choice could be made for each segment as it is saved to the TM, inaccurate matches could be removed and the TM maintained extemporaneously.

7.3.4 The Cost of Inconsistency

The purpose of using TM tools to translate texts is primarily to save time and cost. Webb estimated that an end client might save 40% on translation costs for a 40,000 word project based on two in-house translators and six freelance translators using a TM from a previous iteration of the project, while the translation agency would also make a vastly increased profit from translation using TM (1998, p29-32). During his interview, interviewee G mentioned that when he did a study for his clients, providing a breakdown of the consequences of inconsistent ST across a localisation project, they were “amazed” by the amount of money and issues they would have saved by controlling their ST. We now consider what the cost of inconsistencies might be for both source and target texts.

Our model translation project is a first translation and subsequent updates for software documentation, with 50,000 ST segments to be translated to 20 languages – a total of 1,000,000 segments to be translated. In the ST segments in TMs C and D investigated in the current research there are an average of 17 words per segment overall, but an average of only 14 words per repeated

segment. Based on those figures, we will assume there are on average 15 words per segment in our hypothetical project. Translatorscafe.com[37] lists the average per-word rate for translation from English to German or to Japanese as just under US11c, so for this exercise we will use a per-word rate of €0.08, meaning the average segment will cost €1.20 to translate from scratch. This means that the cost of initially translating from scratch for this project without an existing TM is €1,200,000.

For applying TM discounts, we will use the 30/60/100 model formerly on translationzone.com (2004), so that 100% matches are paid at 30% of the full per-word rate (in this case 36c per average segment), and 85-99% matches are paid 60% of the full rate (in this case 72c per average segment). TM discounts are a contentious issue, which may be why this model was removed from the Translation Zone website, run by SDL. (Anecdotal evidence exists for translators allowing higher or lower discounts than the 30/60/100 model suggests, but few present their discount rates up-front, and there appears to be little consensus as to how much discount to offer, or even whether to offer a TM discount at all (see McBride 2009 in section 2.5.1).)

After one year, an update of the software is to be released. When translating the updated documentation the TMs can provide 100% matches for 300,000 segments (30% of the total to be translated) and 85-99% matches for another 300,000 segments (again 30% of the total), the cost of translating the updated documentation using TM is now €804,000, saving the client €396,000 on year one as shown in table 7.1. This estimate is more conservative than that from Webb (1998), but still represents a saving for the end client of 33%.

[37] http://www.translatorscafe.com/cafe/communityrates.asp [last accessed 16th May 2012]

Match Type	No. of Words Translated in Year One	No. of Words Translated in Year Two using TM (with associated cost)
Translated from scratch	15,000,000	6,000,000 (€480,000)
85-99% matches	-	4,500,000 (€216,000)
100% matches	-	4,500,000 (€108,000)
Total cost	€1,200,000	€804,000

Table 7.1 Savings on Model Project Using TM

Although interviewee G cites rates of ST inconsistency far greater in section 6.3.2, the TMs in our study contained minor ST inconsistencies (such as those specified in section 4.4) at a rate of between 12.2% of category 1 to 4 segments (in TM D) and 17.9% of category 1 to 4 segments (in TM B). Assuming that 15% of ST segments show minor inconsistencies, that would suggest that, had the ST in our hypothetical job been standardised, 45,000 ST segments (15% of the total; containing 675,000 words) receiving 85-99% matches would have been 100% matches, saving the client a further €16,200 or (4.8c per word x 675,000)- (2.4c per word x 675,000).

Table 5.15 in Chapter 5 shows 4.3% to 6.3% of TUs with repeated ST segments contain introduced TT inconsistency. Assuming that 200,000 ST segments contain repetitions, we will suggest that 5% or 10,000 TUs (150,000 words) contain inconsistent TT in the year 2 update. After translation, the target texts are checked with an automatic QA tool which highlights these inconsistencies and they are edited by a reviewer earning the average hourly rate for the language pairs (again from Translatorscafe.com) of €25 per hour, editing 5000 words per 8 hour day. In this case, the cost of TT inconsistency is 30 reviewer days of editing at €200 per day: €6000. According to interviewees in Chapter 6, these edits are rarely saved back to the TM.

When it comes time for the year three documentation update, we will again assume that 1,000,000 segments (15,000,000 words) are to be translated with 300,000 ST segments receiving 100% match proposals and 300,000 ST segments receiving 85-99% match proposals. The cost of the translation stage is therefore the same, but a further 10,000 inconsistent TT segments are introduced, doubling the editing required at the QA stage. Table 7.2 shows the extra cost of inconsistent ST segments and introduced TT inconsistency for years two, three, and four. These calculations are based on a scenario where many TT translations are permitted for a ST segment.

	Year 2	Year 3	Year 4
Total Translation cost	€804,000	€804,000	€804,000
Cost of inconsistent ST	**€16,200**	**€16,200**	**€16,200**
Number of words for QA edit	150,000	300,000	450,000
Cost of QA Edit	**€6,000**	**€12,000**	**€18,000**
Total Cost of Inconsistency	**€22,200**	**€28,200**	**€34,200**

Table 7.2 The Rising Cost of Updates with Polluted TM

Allowing the TM to get more polluted as in this scenario would result in diminishing savings and a longer lead time for QA, but would still save money when compared to translating without TM. Cost could be saved by ensuring consistent ST, managing terminology to restrict TT inconsistency, and maintaining the TMs (see recommendations in section 7.4). Ultimately, the translation client needs to decide whether to consider the long term view or if they feel it more important to make short term savings, as “some suppliers will skip quality steps to offer lower rates, especially if the client is not measuring quality” (Kelly 2012, p2).

7.4 Contributions

It is hoped that this research has adequately answered the research questions and may contribute to the wider discussion of translation technology. Some possible contributions are considered in this section.

This research introduces a novel methodology that may be used to test for TM inconsistency. It also presents a full study, in which that methodology is applied to real-life TMs in different language pairs, and in doing so, challenges one of the axioms of TM use (that TM maximises consistency). As far as we are aware, this is the first published model of consistency in TMs. Where most previous research on TM quality (cited in section 2.7) has been in the form of pilot studies (using student translators) focussed primarily on target text, and has not contained an examination of real-world TM data, this research measures and categorises source and target text inconsistency using TMs that have been created by professional translators in the course of real localisation tasks. The follow-on qualitative interviews provided explanatory detail beyond that available from an analysis of TMs alone.

Current commercial tools, discussed in section 5.6, can identify inconsistency in TMs at a segment level. They do not, however, provide qualitative analyses of inconsistency and, as they are proprietary tools, their methods of searching for inconsistency are not transparent. As far as we are aware, there are no published accounts of how they work. As mentioned in section 7.3, a component for the SOLAS localisation platform is in development based on the methodology published in Chapter 4 of this research. This work also forms part of a larger discussion of consistency in translation technology, suggesting that consistency of bilingual resources will continue to be an issue into the future and that the "quality control of TMs needs to become much more sophisticated" (Zetzsche 2012, p51).

7.5 Limitations and Future Research

Throughout this research our best efforts have been made to maximise the validity of this study by creating a methodology that may be applied beyond the current research and enriching the quantitative findings with a follow-up qualitative study. There are, however, some limitations to this research which may lead to further avenues of research.

This project began in 2008 as part of the research agenda of a centre (Centre for Next Generation Localisation) that was created in 2007. Since that time there has been a great deal of change in TM tool development. As technology is in a state of constant change and development, the current research is limited by focussing on an area that dates quickly.

The study was carried out by one person; it was limited by the researcher's language competence, restricting the study to English, German, and Japanese. As a result, it may be useful to replicate the methodology herein with different language pairs to see whether the results are repeated. The people interviewed in the second phase work predominantly in translation of IT documentation, which may have impacted on their views. As noted in section 6.4.4, the interviewee working in medical device translation (F) had divergent views on terminological consistency, possibly due to regulation in that domain. The sample was limited to those who received and responded to our call for participants via email, social media, and at industry events. Interviewees from other domains may have brought a different perspective.

In connecting the first to the second phase, the findings from the quantitative study were necessarily interpreted by the researcher to formulate questions for the interviewees. This meant that not all findings from the first phase were represented in the interview schedule. The findings from the first phase remain accessible to other researchers, however, and could be used to inform

alternative studies and to generate further research hypotheses worthy of investigation.

As mentioned previously, a component is currently in development that will automate the measurement and categorisation of inconsistencies from this work. In an era of increasing integration of TM and MT systems, it would also be useful to see what change in consistency results from the addition of MT proposals at a segment or sub-segment level. This study could also be developed by focussing on the effect of controlling ST (possibly using author memory), by trying to identify where inconsistency is required in target texts, and by adjusting the source text to limit required inconsistency on the target side.

In section 6.3.2 and 6.5.1, interviewees discuss translation clients' influence on consistency, yet the contribution of clients themselves is missing. The views of other stakeholders in the translation process on consistency would add to the explanatory detail on causes of inconsistency, where inconsistency is required, and whether consistency is desirable other than for increasing leverage. Similarly, it would be valuable to know whether the quality of information is valuable to end consumers if the product performs as expected.

References

ALPAC. 1966. *Languages and machines: computers in translation and linguistics*. *A report by the Automatic Language Processing Advisory Committee*. Washington, DC.

Arthern, P.J. 1979. Machine translation and computerized terminology systems: a translator's Viewpoint. *IN: Proceedings of Translating and the computer.* 14th November 1978, London. p77-108.

ATA Language Technology, 2007. Accessed online at http://www.ata-divisions.org/LTD/?p=53

Austermühl. F. 2001. *Electronic Tools for Translators.* Manchester: St Jerome.

Bass, S. 2006. Quality in the Real World. *IN:* Dunne, K. (Ed.), *Perspectives on Localization.* Amsterdam/Philadelphia: John Benjamins.

Bazeley. P. 2007. *Qualitative Data Analysis with NVivo.* London /Thousand Oaks, CA: Sage Publications.

Becher, V. 2011. *Explicitation and Implicitation in Translation: A Corpus-Based Study of English-German and German-English Translations of Business Texts.* PhD Thesis. Universität Hamburg.

Bédard, C. 2000. Mémoire de traduction cherche traducteur de phrases. *Traduire,* 186, p41-49. Available from http://www.terminotix.com/t_fr/index.htm [Accessed on 15/12/2008]

Benis, M. 2003. Letting the CAT out of the bag. *The Linguist,* 42(3), p70-72.

Benis, M. 2008. Review of Quah, C. K. 2006. Translation and Technology. *The Interpreter and Translation Trainer,* 2(2), p261-267.

Bennet, S. and Gerber, L. 2003. Inside Commercial Machine Translation. *IN:* Somers (ed.) *Computers and Translation: A Translator's Guide*. Amsterdam / Philadelphia: John Benjamins. p175-190.

Bergmann, F. 2008. The Race for Open Source. *Multilingual Computing*, 100 or 9(8), p52-58

Bey, Y., Boitet, C, Kageura, K. 2006. The TransBEY Prototype: An Online Collaborative Wiki-based CAT Environment for Volunteer Translators. *IN: Proceedings Of 5th International Conference on Language Resources and Evaluation.* Genoa, Italy. p49-54.

Bowker, L. 2002. *Computer-aided translation technology: A practical introduction.* Ottawa: University of Ottawa Press.

Bowker, L. 2005. Productivity vs quality? A pilot study on the impact of translation memory systems. *Localisation Focus,* 4(1), p13-20.

Bowker, L. 2007. Translation Memory and ''Text''. *IN:* Bowker, L. and Meyer, I. (eds.) *Lexicography, terminology, and translation. Text-based studies in honour of Ingrid Meyer*. Ottawa: University of Ottawa Press. p175-187.

Bowker, L. and Barlow, M. 2008. A comparative evaluation of bilingual concordancers and translation memories. *IN:* Yuste, E. (ed.) *Topics in Language Resources for Translation and Localisation*. Amsterdam/Philadelphia: John Benjamins. p1-22.

Brace, C. 1992. Trados: Smarter Translation Software. *Language Industry Monitor,* 11, p5-6.

Cahill, K. and Chalut, R. 2009. Optimal Results: What Libraries Need to Know About Google and Search Engine Optimization. *The Reference Librarian,* 50(3), p234-247.

Carl, M. 2009. Grounding Translation Tools In Translator's Activity Data. *In Proceedings of MT Summit XII.* Ottawa, Canada.

Carl, M., Dragsted, B., Elming, J., Hardt, D., Jakobsen, E. L. 2011. The Process of Post-Editing: A Pilot Study. *IN: Proceedings of the 8th International Workshop on Natural Language Processing and Cognitive Science (NLPCS 2011)*. Copenhagen Business School, Denmark, 20th and 21st August. p131-142.

Cherryholmes, C. H., 1992. Notes on Pragmatism and Scientific Realism. *Educational Researcher*, 21(6), p13-17.

Colominas, C. 2008. Towards chunk-based translation memories. *Babel*, 54(4), p343-354.

Creswell, J. W. and Plano Clark, V., 2007. *Designing and Conducting Mixed Methods Research.* London /Thousand Oaks, CA: Sage Publications.

Debove, A., Furlan, S., Depraetere, I. 2011. A contrastive analysis of five automated QA tools (QA Distiller. 6.5.8, Xbench 2.8, ErrorSpy 5.0, SDLTrados 2007 QA Checker 2.0 and SDLX 2007 SP2 QA Check). *IN:* Depraetere, I. (ed.), *Perspectives on Translation Quality*. Berlin/Boston: Walter de Gruyter. p161-192.

DePalma, D. 2005. SDL-Trados: Language Service Provider Reaction to SDL's Purchase of Trados. *Globalization and Localization Association*. Andover, USA: GALA.

Désilets, A., Barrière, C., Quirion, J. 2007. Making WikiMedia resources more useful for translators. *IN: Proceedings of Wikimania 2007, The International Wikimedia Conference*. Taipei, Taiwan, August 3-5.

Désilets, A., Patenaude, G., Melançon, C., Brunette, L., 2009. How translators use tools and resources to resolve translation difficulties: an ethnographic study. *IN: Proceedings of MT Summit XII.* Ottawa, Canada.

Dewey, J. 1931. The Development of American Pragmatism. *IN:* H. S. Thayer (ed.), 1989. *Pragmatism: The classic writings.* Indianapolis, USA: Hackett, pp. 23-40.

Dillon, S. and Fraser, J. 2006. Translators and TM: an investigation of translators' perceptions of translation memory adoption. *Machine Translation*, 20(2), p67–79.

Dragsted, B. 2005. Segmentation in translation: Differences across levels of expertise and difficulty. *Target,* 17(1), p49-70.

Drugan, J., Babych, B. 2010. "Shared resources, shared values? Ethical implications of sharing translation resources". *IN: Proceedings of JEC 2010: Second joint EM+/CNGL Workshop "Bringing MT to the user: research on integrating MT in the translation industry", AMTA 2010*. Denver, Colorado, November 4, 2010; p.3-9.

Englund Dimitrova, B. 2005. *Expertise and Explicitation in the Translation Process*. Amsterdam / Philadelphia: John Benjamin.

Frey, L. R., Botan, C. H., Freidman, P. G., Kreps, G. L., 1991. *Investigating communication: an introduction to research methods.* Cornell University: Prentice Hall.

Fulford, H. and Granell-Zafra, J. 2005. Translation and technology: a study of UK freelance translators. *The Journal of Specialised Translation*, 4. Available from www.jostrans.org/issue04/art_fulford_zafra.php [Accessed 12/12/09].

García, I. 2003. Standard Bearers: TM brand profiles at Lantra-L. *Translation Journal,* 7(4). Available from http://www.proz.com/translation-articles/articles/224/ [Accessed on 11/05/2012]

García, I. 2005. Long-Term Memories: Trados and TM Turn 20. *Journal of Specialised Translation,* 4, p18-31

García, I. 2006. Translation Memories: A Blessing or a Curse? *IN:* Pym, A., Perekrestenko, Starink (eds.) *Translation Technology and its Teaching*. Tarragona: Universitat Rovira i Virgili.

García, I. 2006. Power Shifts in Web-Based Translation Memory. *Machine Translation,* 21, p55–68.

García, I. 2008. Translating and Revising for Localisation: What Do We Know? What Do We Need to Know? *Perspectives: Studies in Translatology,* 16(1, 2), p49–60.

Global Watchtower, 2007. Idiom Releases Worldserver 9. Available from http://www.globalwatchtower.com/2007/09/11/idiom-releases-worldserver-9/ [Accessed on 11/11/2009]

Global Watchtower, 2008. SDL Buys Idiom. Available from http://www.globalwatchtower.com/2008/02/11/sdl-buys-idiom/ [Accessed on 11/11/2009]

Global Watchtower, 2009. Globalsight Adds Crowdsourcing To Its Vision. Available from: http://www.globalwatchtower.com/2009/08/05/crowdsight/ [Accessed on 8/8/2009]

Google Blog, 2009. Google Translator Toolkit and Minority Languages. Available from http://googleblog.blogspot.com/2009/10/google-translator-toolkit-and-minority.html [Accessed 10/11/2009]

Gow, F. 2003. *Metrics for Evaluating Translation Memory Software*. MA thesis. University of Ottawa.

Gow, F. 2007. You Must Remember This: The Copyright Conundrum of "Translation Memory" Databases. *Canadian Journal of Law and Technology,* 6(3), p175-192.

Gubrium, J. F. and Holstein, J. A., 2003. Active Interviewing. *IN:* Gubrium, J. F. and Holstein, J. A. (Eds.) *Postmodern Interviewing*. London /Thousand Oaks, CA: Sage Publications, p67-80.

Guerberof, A. 2008. *Productivity and Quality in the Post-editing of Outputs from Translation Memories and Machine Translation*. Minor Dissertation. Universitat Rovira i Virgili, Tarragona.

Guerberof, A. 2012. *Productivity and Quality in the Post-editing of Outputs from Translation Memories and Machine Translation*. PhD Dissertation. Universitat Rovira i Virgili, Tarragona.

Guillemin, P. and Trillaud, S. 2012. What Has Become of LISA's OSCAR Standards? *Multilingual Computing,* 127 or 23(3), p38-41.

Heyn, M. 1996. Present and Future Needs in the CAT World, *LISA Newsletter* V3, September 1996. p15-33.

Hilmarsson-Dunn, A.M. 2001. Protectionist Language Policies in the Face of the Forces of English: The Case of Iceland. *Language Policy*, 5, p293-312.

Hodász, G. 2006. Evaluation Methods of a Linguistically Enriched Translation Memory System. *IN: Proceeedings of LREC 2006. Genoa, Italy*. Budapest, Pázmány Péter Catholic University.

Hutchins, J. 1998. The Origin of the Translator's Workstation. *Machine Translation,* 13(4), p287-307.

Johnson, B. and Turner, L. A. 2003. Data Collection Strategies in Mixed Methods Research. *IN:* Tashakkori and Teddlie (eds.) *Handbook of Mixed Methods in Social and Behavioral Research*. London /Thousand Oaks, CA: Sage Publications. p297-320.

Karamanis, N., Luz, S., Doherty, G. 2010. "Translation Practice in the workplace and Machine Translation". *IN:* Hansen, V. and Yvon, F. (eds.) *EAMT (2010): Proceedings of the 14th Annual conference of the European Association for Machine Translation*. 27-28 May 2010, Saint-Raphaël, France. pp8. [Accessed from http://www.mt-archive.info/EAMT-2010-Karamanis.pdf on 10th November 2010]

Kay, M. 1997 (Originally 1980). The Proper Place of Men and Machines in Language Translation. *Machine Translation,* 12, p3-23.

Kelly, N. 2012. What to Do When Translation Vendors Raise Prices. *Common Sense Advisory Article.* MA, USA: Common Sense Advisory.

Kemper, E. A., Stringfield, S., Teddlie, C. 2003. Mixed Methods Sampling Strategies in Social Science Research. *IN:* Tashakkori and Teddlie (Eds.) *Handbook of Mixed Methods in Social and Behavioral Research.* London /Thousand Oaks, CA: Sage Publications, p273-296.

Kenny, D. 2007. Review of Quah, C. K. 2006. Translation and Technology. *Machine Translation,* 20, p291–295.

Kenny, D. 2012. Corpus Analysis in Translation Studies. *IN:* Chapelle, C. (ed.) *The Encyclopedia of Applied Linguistics*. Oxford: Wiley-Blackwell.

Kvale, S. and Brinkmann, S. 2009. *InterViews: Learning the Craft of Qualitative Research Interviewing.* London /Thousand Oaks, CA: Sage Publications.

Lagoudaki, E. 2006. Translation Memories Survey 2006: Users' Perceptions Around TM Use. *IN: Proceedings of ASLIB Translating and the Computer 28*. London, UK. 15-16 November 2006.

Lagoudaki, E. 2008. *Expanding the Possibilities of Translation Memory Systems: From the Translator's Wishlist to the Developer's Design*. PhD thesis. Imperial College, London.

Lauriston, A., 1997. Terminology and the Computer. *IN:* Dubuc, R. (Ed.) *Terminology: A Practical Approach.* Brossard: Linguatech. p179-192.

Levitt, G. 2003. Internet-Based Sharing of Translation Memory. *Multilingual Computing,* 57 or 14(5), p38-41.

LISA. 2004. Translation Memory eXchange (TMX). *LISA Website.* [Accessed from http://www.lisa.org/Translation-Memory-e.34.0.html 11/11/2009]

Lindlof, T. R. and Taylor, B. C. 2002. *Qualitative Communication Research Methods*. London /Thousand Oaks, CA: Sage Publications.

LISA. 2005. TMX 1.4b Specification. Accessed from [http://www.lisa.org/fileadmin/standards/tmx1.4/tmx14-20050426.htm on 19/4/2010]

McBride, C. 2009. *Translation Memory Systems: An Analysis of Translators' Attitudes and Opinions*. MA thesis. University of Ottawa.

Mangiron Hevia, C. 2006. Video Games Localisation: Posing New Challenges to the Translator. *Perspectives: Studies in Translatology,* 14(4) p306–323.

Maldonado Guerra, A. and Itagaki, M. 2008. Checking Terminology Consistency with Statistical Methods. *IN: Proceedings of LRC XIII: the Localisation Research Centre conference 2008*. Dublin, Ireland.

Marshman, E. 2012 Powering translation and empowering translators: A dual role for language technologies? 4th International Association for Translation and Intercultural Studies Conference. Queens University, Belfast. July 24-27.

Melby, A.K. 1982. Multi-level translation aids in a distributed system. *IN: Coling 82. Proceedings of the Ninth International Conference on Computational Linguistics.* Prague, July 5-10.

Merkel, M. 1996. Checking translations for inconsistency – A tool for the editor. *IN: Proceedings from AMTA '96.* Montreal, October 2-5.

Merkel, M. 1998. A Study of Consistency and Variation in Technical Translations: A Study of Translators' Attitudes, *IN:* Bowker, L., Cronin, M., Kenny D., and Pearson, J. (Eds.) *Unity in Diversity? Current Trends in Translation Studies*. Manchester: St. Jerome Publishing, p137-149.

Mitkov, R. and Corpas, G. 2008. Improving Third Generation Translation Memory systems through identification of rhetorical predicates. *IN: Proceedings of LangTech'2008.* Rome.

Mogensen, E. 2000. Orwellian Linguistics: How Using Computer-Aided Translation Tools Impacts The Target Language. *Language International,* 12(5), p28-31.

Moorkens, J. 2009. Total Recall? A Case Study of Consistency in Translation Memory. *IN: Proceedings of LRC XIV.* 24th and 25th September. Limerick, Ireland. p101-110.

Moran, J. and Lewis, D. 2011. Unobtrusive Methods for Low-Cost Manual Evaluation of Machine Translation. *IN: Proceedings of Tralogy 2011, Translation Careers and Technologies: Convergence Points for the Future.* Available from http://lodel.irevues.inist.fr/tralogy/index.php?id=141 [Accessed 24/5/2012]

Morse, J. M. and Niehaus, L. 2009. *Mixed Method Design: Principles and Procedures*. Walnut Creek, CA: Left Coast Press.

Nakada, G. 1999. TRADOS: Translating Success into 55 Languages. *Computing Japan,* 6(11). [Accessed from http://www.japaninc.com/cpj/magazine/issues/1999/nov99/docs/nov99_trados.html 10/11/2009]

Nyberg, E., Mitamura, T. and Huijsen, W. 2003. Controlled Language for Authoring and Translation. *IN:* Somers, H (Ed.) *Computers and Translation: A Translator's Guide.* Amsterdam / Philadelphia: John Benjamin. p245-282.

O'Brien, S. and Barany, R. 1996. S-Tagger Converts Between FrameMaker and Workbench. *Multilingual Computing,* 8(1). Available from http://www.multilingual.com/articleDetail.php?id=1115 [Accessed 10/11/2009]

O'Brien, S. 2006. Eye-tracking and translation memory matches. *Perspectives: Studies in Translatology,* 14(3), p185–205.

O'Brien, S., O'Hagan, M. and Flanagan, M. 2010. Keeping an Eye on the UI Design of Translation Memory: How Do Translators Use the "Concordance" Feature? *IN: European Conference on Cognitive Ergonomics*, Delft, The Netherlands. 25-28 Aug 2010.

Olohan, M. 2011. Translators and Translation Technology: The Dance of Agency. *Translation Studies* 4(3), p342-357.

Onwuegbuzie, A. J. and Teddlie, C. 2003. A Framework for Analysing Data in Mised Methods Research. *IN:* Tashakkori and Teddlie (eds.) *Handbook of Mixed Methods in Social and Behavioral Research*. London /Thousand Oaks, CA: Sage Publications, p351-384.

Pelino, S. 2011. An Empirical Model for MT Quality Assessment and Implications for Business Models. *ASLIB Translating and the Computer 33*. London, UK. 17-18 November 2011.

Pym, A. 2004. *The Moving Text*. Amsterdam / Philadelphia: John Benjamins.

Pym, A. 2008. "Professional corpora": teaching strategies for work with online documentation, translation memories and content management. *Chinese Translator's Journal,* 29(2), p41-45.

Pym, A. 2011. Democratizing translation technologies - the role of humanistic research. *IN: Proceedings of Luspio Translation Automation Conference*. Rome, Italy. 5th and 6th April.

Quah, C. K. 2006. *Translation and Technology.* New York: Palgrave McMillan.

Quinn Patton, M. 2002. *Qualitative Research and Evaluation Methods*. London / Thousand Oaks, CA: Sage Publications.

Rasmussen, K. W. and Schjoldager, A. 2011. Revising Translations: A Survey of Revision Policies in Danish Translation Companies. *The Journal of Specialised Translation,* 15. [Accessed from http://www.jostrans.org/issue15/art_rasmussen.php on 11th March 2011]

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

Ribas López, C. R. 2007. *Translation Memories As Vehicles For Error Propagation: A Pilot Study*. Minor dissertation. Universitat Rovira i Virgili, Tarragona.

Richards, L. 2005. *Handling Qualitative Data: A practical guide*. London / Thousand Oaks, CA: Sage Publications.

Rieche, A. C. 2004. *Memória de tradução: auxílio ou empecilho?* MA Thesis. Pontifícia Universidade Católica do Rio de Janeiro.

Saffer, D. 2006. *Designing for Interaction: Creating Smart Applications and Clever Devices.* Berkeley: Peachpit.

Saraireh, M. 2001. Inconsistency in Technical Terminology: A Problem for Standardization in Arabic. *Babel,* 47(1), p10-21.

Schofield, J. W. 2000. Increasing the Generalizability of Qualitative Research. *IN*: Gomm, R., Hammersley, M. and Foster, P. (eds.) *Case Study Method: Key Issues.* London: Sage. p69-97

Shaughnessy, J. J., 2009. *Research Methods in Psychology.* New York, USA: McGraw-Hill Higher Education.

Smith, R. 2008. Your Own Memory. *The Linguist,* 47(1), p22-23.

Smith, R. 2009. Copyright Issues in Translation Memory Ownership. *IN: Proceedings of ASLIB Translating and the Computer 31*. London, UK. 19-20 November 2006.

Somers, H. (Ed.) 2003. *Computers and Translation: A Translator's Guide.* Amsterdam / Philadelphia: John Benjamin.

Stake, R. E., 1995. *The Art of Case Study Research*. London / Thousand Oaks, CA: Sage Publications.

Stephenson, N. 2011. Innovation starvation. *World Policy Journal,* 28(3), p11-16. [Accessed on 20th March 2012 from http://www.worldpolicy.org/journal/fall2011/innovation-starvation]

Tashakkori, A. and Teddlie, C. B., 2003. *Handbook of Mixed Methods in Social and Behavioral Research*. London /Thousand Oaks, CA: Sage Publications.

TAUS, 2009. Profit of sharing: Up to 50%. Accessed from: http://www.tausdata.org/index.php/news/news/70-increasing-leveraging-from-shared-industry-data

Teddlie, C. B. and Tashakkori, A., 2009. *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences.* London /Thousand Oaks, CA: Sage Publications.

Translationzone.com, 2004. *Marketing Tip: Using Pricing and Trados to Your Advantage* [Online]. Available from: http://web.archive.org/web/20040606125633/http://www.translationzone.com/tip.asp?ID=3 [Archived from 6/12/2004 at web.archive.org; accessed 1/5/2012].

Vashee, K. 2009. Clean Data Improved SMT Engine Results. *Multilingual Computing,* 108 or 20(8), p46-48.

Wallis, J. 2006. *Interactive Translation Vs Pre-Translation in the Context of Translation Memory Systems: Investigating the Effects of Translation Method on Productivity, Quality and Translator Satisfaction*. MA thesis. University of Ottawa.

Waßmer, T. 2003. SDLX™ Translation Suite 2003. *Translation Journal,* 7(3). Available from http://www.accurapid.com/Journal/25sdlx.htm on 11/11/09. [Accessed 11/11/09]

Webb, L. E. 1998. *Advantages and Disadvantages of Translation Memory: A Cost/Benefit Analysis*. MA Thesis. Monterey Institute of International Studies.

Whyman, E., and Somers, H. 1999. Evaluation Metrics for a Translation Memory System. *Software – Practice and Experience,* 29(14), p1265-1284.

Wordfast, 2011. Wordfast Anywhere. Available from: http://www.wordfast.net/?whichpage=anywhere

Williams, J. and Chesterman, A. 2002. *The Map. A Beginner's Guide to Doing Research in Translation Studies*. Manchester: St. Jerome Publishing.

Wu, H., Wang, H., Liu, Z., Tang, K., 2005. Improving Translation Memory With Word Alignment Information. Beijing: Toshiba Research Centre.

Yamada, M. 2011a. *Revising Text: An Empirical Investigation of Revision and the Effects of Integrating A TM and MT System into the Translation Process*. PhD Thesis*.* Rikkyo University.

Yamada, M. 2011b. The Effect of Translation Memory Databases on Productivity. *IN*: Pym, A. (ed). *Translation Research Projects 3*. Tarragona: Intercultural Studies Group. p63-73.

Zetzsche, J. 2012. Translation Technology Comes Full Circle. *Multilingual Computing* 127 or 23(3), p50-51.

Appendices

A: Informed Consent Form

This research by Joss Moorkens of the School of Applied Languages and Intercultural Studies at Dublin City University is intended to contribute to a doctoral thesis with the provisional title 'A Study of Consistency in Translation Memories', funded by Science Foundation Ireland via the Centre for Next Generation Localisation. When completed in 2012, it is hoped to contribute to the field of localisation.

Participation in the research is entirely anonymous. Throughout the research, the participants will remain anonymous and every effort will be made to conceal their identity. Therefore, the researcher will change the names of participants and institutions, geographical names or any other particulars which can be traced back to any of the participants. In order to maintain confidentiality, all audio recordings and transcripts will be stored on the researcher's password-protected computer based at DCU.

All material related to the research will be destroyed by the researcher five years following the publication of the doctoral thesis. It is very important that all participants understand that their participation is entirely voluntary. They can withdraw from the research any time up to the publication of the information. There will be no penalty for withdrawing at any stage and the contents of the interviews will have no bearing on any assessments or grades. By signing the informed consent form, all participants indicate that they understand what the research is about and their participation in the research.

Participant – please complete the following (Circle Yes or No for each question)	
I have read the Plain Language Statement (or had it read to me)	*Yes/No*
I understand the information provided	*Yes/No*
I have had an opportunity to ask questions and discuss this study	*Yes/No*
I have received satisfactory answers to all my questions	*Yes/No*
I am aware that an audio recording of my interview will be made	*Yes/No*

I may withdraw from the Research Study at any point.

I have read and understood the information in this form. My questions and concerns have been answered by the researchers, and I have a copy of this consent form. Therefore, I consent to take part in this research project

Participants Signature: _____

Name in Block Capitals: _____

Witness: _____

Date: _____

B: Plain Language Statement

The current research, provisionally titled 'A Study of Consistency in Translation Memories', is conducted by Joss Moorkens of the School of Applied Languages and Intercultural Studies at Dublin City University. He is supervised by Dr Dorothy Kenny, Dr Sharon O'Brien and Mr Reinhard Schäler and funded by Science Foundation Ireland via the Centre for Next Generation Localisation. When completed in 2012, it is hoped to contribute to the field of localisation and to improve our understanding of the process of translation using translation memories.

Participation in this research is in the form of interviews to be recorded digitally. These audio recordings will then be transcribed and analysed. The transcribed interviews are intended to add detail to quantitative analyses of TM data. Participation in the research is entirely anonymous. Throughout the research, the participants will remain anonymous and every effort will be made to conceal their identity. Therefore, the researcher will change the names of participants and institutions, geographical names or any other particulars which can be traced back to any of the participants. In order to maintain confidentiality, all audio recordings and transcripts will be stored on the researcher's password-protected computer based at DCU.

All material related to the research will be destroyed by the researcher five years following the publication of the doctoral thesis. It is very important that all participants understand that their participation is entirely voluntary. They can withdraw from the research any time up to the publication of the information. There will be no penalty for withdrawing at any stage and the contents of the interviews will have no bearing on any assessments or grades. By signing the informed consent form, all participants indicate that they understand what the research is about and their participation in the research.

If participants have concerns about this study and wish to contact an independent person, please contact:

The Secretary,
Dublin City University Research Ethics Committee,
C/o Office of the Vice-President for Research,
Dublin City University,
Dublin 9.
Tel 01-7008000

Supervisors contact:

Dorothy Kenny dorothy.kenny@dcu.ie
Sharon O'Brien sharon.obrien@dcu.ie

C: Python Script

```
import sys
import codecs
from xml.etree import cElementTree as ET

def parse_tmx(tmx_file):
   """Parse a TMX file in a non-robust manner and write seg text to 'out' file'.
   The parse makes use of the itertext()method that is new in Python 2.7.
   Doing it with Python < 2.7 involves playing with .text, .tail, which is much more
cumbersome.

   @tmx_file -- path to a tmx file.

   """

   try:
      data = open(tmx_file)
      parsed_file = ET.parse(data)
   except:
      print '%s cannot be parsed' % tmx_file
   else:

      root_element = parsed_file.getroot() #Get the root element
      all_segs = [] #Initialize a list to store the final results (that will be written to
the out file)

      for trans_unit in root_element.getiterator('tu'): #Create an iterator on a
specific element: tu
         #FIXME: This should wrapped in a try/except
         segs = trans_unit.findall("seg") #For each tu value, we try and locate the
seg(s) element(s)
         cur_segs = [] #Initialize a temp list to store the text of the current segs
         for seg in segs: #Loop over the segs
            cur_text = "" #Initialize a string that will be incremented with text found
            text = seg.itertext() #This is the crux of the function: an iterator is
created
            while text:
               try:
                  cur_text += text.next() # All text is extracted
               except:
                  break #Until there is no text left
            cur_segs.append(cur_text) #The string is appended to the temp list
```

```
        #Once we're done, we join the results of each temp list using a tab and we
append a new line character.
        all_segs.append("\t".join(cur_segs) + "\n") #The resulting string are
appended to the list

    out_file = 'outJap.txt'
    codecs.open(out_file, 'wb', 'utf-8').writelines(all_segs) #We assume UTF-8 to
write the final strings
    print 'Text has been extracted from %s and written to %s' % (tmx_file,
out_file)


if __name__ == "__main__":

  usage = "python tmx.py path_to_tmx_file"
  if len(sys.argv) == 2:
    try:
      parse_tmx(sys.argv[1])
    except:
      print usage
  else:
    print usage
```