

**Interactive Translation vs Pre-translation in the Context of  
Translation Memory Systems:  
Investigating the effects of translation method on productivity, quality  
and translator satisfaction**

Julian Wallis

Thesis submitted to the  
Faculty of Graduate and Postdoctoral Studies  
In partial fulfillment of the requirements  
For the MA in Translation Studies program

School of Translation and Interpretation  
Faculty of Arts  
University of Ottawa

© Julian Wallis, Ottawa, Ontario, Canada 2006

## **Acknowledgements**

There are many people I would like to thank for helping make this thesis what it is. Primarily, I would like to thank my supervisor, Dr. Lynne Bowker who has been nothing short of amazing since the day we met. Lynne, without your confidence in me from the start, I would never have written a thesis to begin with! Your guidance and motivation have been an inspiration to me, and many others. Thanks for being there for everything, thesis related or not!

I would like to thank Sylvie Lambert and Malcolm Williams for volunteering their time, knowledge and resources, and Fernand Aubin for permitting me to use the Writing Centre for my experiment and research. I would also like to thank Sinclair Robinson for interesting me in translation studies to begin with and for being the inspiration for the career path I have chosen.

Exceptional thanks go to Matt Fifer, Christina Pachis (you are my saving grace!), Marta Gómez Palou, Marie Alwick and Linda Ham, for helping me develop and test my methodology and to Renée Desjardins for editing and helping translate my Abstract. To Matt, Renée, Marta and Marie thanks for being there to talk, listen, laugh and sometimes scream! The fact that we were all in this together has made this process that much easier, and much more entertaining!

Throughout the writing of this thesis and always, I am lucky to have the support of great friends and family. To Gramma Smale and Grandma and Grandpa Wallis, thanks for all the love and support (moral and financial!) and for helping get me to where I am today. To Katie Maw, thanks for keeping me sane, keeping me in the loop, and making me feel like I can do anything!

Very important thanks go to Stephen Webb, who has been there for me every single day. Thank you for listening, thank you for holding my hand, for always reminding me to stay positive, and for believing in me no matter what. You are an amazing person and I am lucky to have you in my life.

Finally, I would like to thank Mom, Dad and Dain. Brother? Thanks for being you. Your carefree attitude reminds me that no matter what happens, everything will always turn out alright. You enrich the life of every person you meet, and you are a great inspiration to me. Thanks for always being there and for making me laugh! Mom, thanks for remembering every date and deadline, and for always being right about everything! Thank you for pushing me to be the best and for always having encouraging words. Dad, thank you for making me feel like I am the smartest person alive! Thanks for listening, for being so supportive, and for always assuring me that I am making the right decisions. You are the most amazing people in the world. Thanks for laughing with me, always supporting me and loving me unconditionally. I wouldn't be here if it wasn't for you...

## **Abstract**

Translation Memory (TM) systems are currently among the most popular translation tools available on the market. Clients today are looking for productivity and quality, but at the same time are looking to reduce costs. A TM stores portions of previously translated texts in a database so that they can be reused in new texts.

There are two ways of working with a TM: interactive mode and pre-translation mode. In interactive mode, the TM system offers proposals to the translator, who can choose to accept and adapt them, or refuse them and translate from scratch. In pre-translation mode, the TM system's proposals are automatically inserted into the new source text, producing a hybrid text containing a mixture of source and target language elements. The translator's job is then to verify the translation of the automatically replaced portions of text, as well as to translate any text that was left in the source language.

As the demand for translation continues to rise, more and more translators are looking to TMs to help increase productivity; however, for a variety of reasons, such as cost and incompatible file formats, they do not always have access to a useful TM. Therefore, translators need access to the TM database of the client they are working for in order to complete their translations. However, clients are increasingly hesitant to give out their databases due to proprietary issues. For this reason, many clients are turning to the pre-translation option so that they can get their translations done without having to give out their TM databases. To date, however, no one has yet studied whether the choice of pre-

translation vs interactive translation has an effect on productivity, quality and translator satisfaction. This thesis aims to go some way towards filling this gap by designing and testing a methodology to compare these two methods of working with a TM system.

The thesis is divided into three main parts. Part I provides background information on TMs through a literature survey, the findings of which form the basis for three hypotheses concerning the ways in which interactive and pre-translation will impact translation productivity and quality, as well as translator satisfaction. Part II describes an experiment that was designed and conducted to investigate these hypotheses. Finally, Part III evaluates the research carried out in this thesis and suggests possible ways to expand it in the future.

## Résumé

Les mémoires de traduction (MT) sont présentement l'un des outils de traduction les plus convoités sur le marché. Aujourd'hui, les clients cherchent la productivité et la qualité, sans que cela affecte de façon négative leurs revenus. Une MT stocke des extraits de textes déjà traduits dans une base de données afin que ceux-ci puissent être réutilisés dans de nouveaux textes.

Deux méthodes s'offrent à l'utilisateur lors de la traduction avec une MT : la traduction interactive et la pré-traduction. Lorsqu'on travaille en mode interactif, la MT propose des traductions, et le traducteur peut soit accepter et adapter les traductions proposées, soit les refuser et traduire à partir de zéro. En utilisant l'option pré-traduction, les traductions proposées par le système sont automatiquement insérées dans le nouveau texte de départ, ce qui produit un texte hybride qui contient à la fois des éléments de la langue de départ et des éléments de la langue d'arrivée. Le travail du traducteur consiste donc à assurer une traduction adéquate des éléments insérés de manière automatique dans le texte de départ, en plus de traduire les parties du texte source non traduites.

Puisque la demande de traduction augmente, de plus en plus de traducteurs se tournent vers les MT; cependant, pour diverses raisons, notamment le coût des MT et l'incompatibilité des formats de fichiers, les traducteurs n'ont pas toujours accès à une MT utile. Pour cette raison, les traducteurs doivent avoir accès à la base de données de la MT de leur client de façon à pouvoir effectuer leur travail. Or, les clients hésitent de plus

en plus à partager leurs bases de données en raison de questions relatives à la propriété intellectuelle. Par conséquent, beaucoup de clients se tournent vers la pré-traduction, ce qui leur permet de limiter l'accès à leurs bases de données. Jusqu'à maintenant, on ignore si le choix entre la traduction interactive et la pré-traduction produit un effet sur la productivité, la qualité et la satisfaction du traducteur. Cette thèse se propose d'analyser les deux méthodes de traduction avec MT avec une méthodologie comparative.

Cette thèse est divisée en trois parties. La première partie offre un aperçu général des recherches déjà effectuées au sujet des MT. Sur ces recherches sont basées les trois hypothèses relatives à la traduction et la pré-traduction (productivité, qualité, et satisfaction du traducteur). La deuxième partie comprend une étude qui a tenté de répondre aux questions relatives aux hypothèses de la première partie. Finalement, la troisième partie évalue la recherche effectuée et offre des suggestions quant aux développements éventuels de certaines idées proposées dans la présente thèse.

## Table of Contents

|   |          |
|---|----------|
| ACKNOWLEDGEMENTS .....                                    | II       |
| ABSTRACT.....   | IV       |
| RESUME .....  | VI       |
| TABLE OF CONTENTS.....                                    | VIII     |
| LIST OF FIGURES .....                                     | XI       |
| LIST OF TABLES .....                                      | XII      |
| <b>CHAPTER 0 – INTRODUCTION.....</b>                      | <b>1</b> |
| 0.1 OBJECTIVES.....                                       | 1        |
| 0.2 BACKGROUND AND CONTEXT.....                           | 1        |
| 0.2.1 Translation Market: Challenges .....                | 2        |
| 0.2.1.1 Increased demand for translation .....            | 2        |
| 0.2.1.2 Shorter deadlines.....                            | 4        |
| 0.2.1.3 Shortage of translators.....                      | 4        |
| 0.2.2 Translation Market: Some Impacts of Technology..... | 5        |
| 0.2.2.1 Tools available to translators .....              | 5        |
| 0.2.2.2 TMs and the question of ownership .....           | 6        |
| 0.2.2.3 TMs and the question of payment.....              | 7        |
| 0.2.3 Current Situation.....                              | 9        |
| 0.3 JUSTIFICATION AND MOTIVATION FOR RESEARCH.....        | 12       |
| 0.4 GENERAL METHODOLOGY .....                             | 13       |
| 0.5 SCOPE AND LIMITATIONS .....                           | 14       |
| 0.5.1 TM System.....                                      | 14       |
| 0.5.2 Languages .....                                     | 15       |
| 0.5.3 Participants.....                                   | 15       |
| 0.5.4 Texts.....  | 16       |
| 0.6 OUTLINE.....  | 16       |

## **PART I: BACKGROUND INFORMATION AND THEORETICAL FRAMEWORK ..... 18**

|  |           |
|--|-----------|
| <b>CHAPTER 1 – INTRODUCTION TO TRANSLATION MEMORIES, INTERACTIVE TRANSLATION, AND PRE-TRANSLATION.....</b> | <b>19</b> |
| 1.1 FROM INITIAL CONCEPTION TO PRESENT-DAY POPULARITY.....   | 19        |
| 1.2 ORGANIZATION AND STORAGE OF INFORMATION IN A TM.....   | 23        |
| 1.2.1 Segmentation.....  | 24        |
| 1.2.2 Alignment.....   | 25        |
| 1.3 RETRIEVAL OF INFORMATION FROM A TM .....   | 26        |
| 1.3.1 Exact Match .....  | 27        |
| 1.3.2 Fuzzy Match.....   | 27        |
| 1.3.3 Sub-segment Match.....   | 31        |
| 1.3.4 Term Match.....  | 32        |
| 1.4 WORKING WITH A TM.....   | 33        |
| 1.4.1 Interactive Mode .....   | 33        |
| 1.4.2 Pre-translation Mode.....  | 34        |
| 1.4.2.1 History of pre-translation .....   | 36        |
| 1.4.2.2 Pre-Translation in the current context of TM.....  | 37        |

|  |   |           |
|--|---|-----------|
| 1.4.2.3                                      | Pre-translation and post-editing .....      | 40        |
| 1.5  | CONCLUDING REMARKS .....                    | 43        |
| <b>CHAPTER 2 – RESEARCH HYPOTHESES .....</b> |   | <b>44</b> |
| 2.1  | HYPOTHESIS 1: PRODUCTIVITY .....            | 44        |
| 2.2  | HYPOTHESIS 2: QUALITY .....                 | 45        |
| 2.3  | HYPOTHESIS 3: TRANSLATOR SATISFACTION ..... | 47        |
| 2.4  | CONCLUDING REMARKS .....                    | 49        |

**PART II: EXPERIMENTAL DESIGN AND TESTING .....** 50

**CHAPTER 3 – EXPERIMENT: PREPARATION AND EXECUTION .....** 51

|         |   |    |
|---------|---|----|
| 3.1     | PREPARATORY WORK .....  | 51 |
| 3.1.1   | Identifying Participants .....                                | 51 |
| 3.1.1.1 | Translators.....  | 51 |
| 3.1.1.2 | Evaluators.....   | 53 |
| 3.1.2   | Identifying a Domain .....                                    | 53 |
| 3.1.3   | Selecting a TM Tool .....                                     | 54 |
| 3.1.4   | Building a TM Database and Termbase.....                      | 54 |
| 3.1.4.1 | Verifying alignment of texts .....                            | 55 |
| 3.1.4.2 | Adding to the termbase .....                                  | 56 |
| 3.1.5   | Choosing Source Texts .....                                   | 56 |
| 3.1.6   | Producing Pre-translations .....                              | 57 |
| 3.1.7   | Preparing Instructions and Questionnaire for Translators..... | 58 |
| 3.1.8   | Preparing Evaluation Criteria for Evaluators .....            | 58 |
| 3.1.9   | Training Translators on the Use of Fusion Translate .....     | 59 |
| 3.1.10  | Pre-Testing and Refining the Pilot Study.....                 | 60 |
| 3.2     | EXPERIMENT .....  | 61 |
| 3.2.1   | Translation .....   | 61 |
| 3.2.2   | Evaluation of Translations .....                              | 63 |
| 3.3     | CONCLUDING REMARKS .....                                      | 63 |

**CHAPTER 4 - DATA ANALYSIS .....** 64

|         |  |    |
|---------|--|----|
| 4.1     | PRODUCTIVITY .....   | 64 |
| 4.1.1   | Data .....   | 64 |
| 4.1.2   | Discussion .....   | 65 |
| 4.1.2.1 | Texts.....   | 65 |
| 4.1.2.2 | Technical difficulties.....  | 67 |
| 4.1.2.3 | Familiarity with software .....  | 68 |
| 4.1.2.4 | More results to evaluate in interactive mode .....                           | 69 |
| 4.1.2.5 | Quality of the finished translations .....                                   | 69 |
| 4.1.3   | General Conclusions Regarding Productivity .....                             | 70 |
| 4.2     | QUALITY.....   | 71 |
| 4.2.1   | Data .....   | 71 |
| 4.2.2   | Discussion .....   | 73 |
| 4.2.2.1 | Perceived difficulty of texts and quality of the finished translations ..... | 74 |
| 4.2.2.2 | Amenability of one text for interactive translation .....                    | 75 |
| 4.2.3   | General Conclusions Regarding Quality.....                                   | 75 |
| 4.3     | TRANSLATOR SATISFACTION .....  | 76 |
| 4.3.1   | Discussion .....   | 76 |
| 4.3.1.1 | Pre-translation mode .....   | 76 |
| 4.3.1.2 | Interactive mode.....  | 78 |

|       |   |    |
|-------|---|----|
| 4.3.3 | General Conclusions Regarding Translator Satisfaction ..... | 78 |
| 4.4   | CONCLUDING REMARKS .....                                    | 79 |

**PART III: EVALUATION AND CONCLUSION ..... 81**

|   |           |
|---|-----------|
| <b>CHAPTER 5 – CONCLUDING REMARKS.....</b>            | <b>82</b> |
| 5.1 GENERAL EVALUATION .....                          | 82        |
| 5.1.1 Database Size .....                             | 83        |
| 5.1.2 Nature of Participants.....                     | 84        |
| 5.1.3 Instructions for Translators .....              | 85        |
| 5.1.4 Evaluation Phase .....                          | 85        |
| 5.1.5 Analysis of Source Texts .....                  | 87        |
| 5.1.6 Order of Operation .....                        | 87        |
| 5.2 SUGGESTIONS FOR FUTURE RESEARCH.....              | 88        |
| 5.2.1 Increase Number of Participants .....           | 89        |
| 5.2.2 Add New Languages and Language Directions ..... | 89        |
| 5.2.3 Investigate Different Text Types .....          | 90        |
| 5.2.4 Experiment with Different TM Tools .....        | 90        |
| 5.3 CONCLUDING REMARKS .....                          | 90        |
| <b>REFERENCES .....</b>                               | <b>92</b> |

**APPENDICES ..... 98**

|  |            |
|--|------------|
| <b>APPENDIX A: SOURCE TEXTS.....</b>   | <b>99</b>  |
| <b>APPENDIX B: PRE-TRANSLATED SOURCE TEXTS.....</b>  | <b>101</b> |
| <b>APPENDIX C: TUTORIAL ON FUSION TRANSLATE .....</b>  | <b>103</b> |
| <b>APPENDIX D: INSTRUCTIONS FOR TRANSLATORS, TIME SHEET, RESOURCES SHEET AND<br/>TRANSLATOR QUESTIONNAIRE.....</b> | <b>106</b> |
| <b>APPENDIX E: COMPLETED TIME SHEETS .....</b>   | <b>113</b> |
| <b>APPENDIX F: COMPLETED RESOURCES SHEETS .....</b>  | <b>117</b> |
| <b>APPENDIX G: COMPLETED QUESTIONNAIRES .....</b>  | <b>121</b> |
| <b>APPENDIX H: EVALUATOR GUIDELINES .....</b>  | <b>131</b> |
| <b>APPENDIX I: COMPLETED FINAL SCORE SHEETS.....</b>   | <b>134</b> |
| <b>APPENDIX J: TARGET TEXTS .....</b>  | <b>136</b> |

## List of Figures

|             |  |    |
|-------------|--|----|
| Figure 0.1  | <i>Example of a pre-translated text</i> .....                                | 11 |
| Figure 1.1  | <i>Example of how a text is divided into segments</i> .....                  | 24 |
| Figure 1.2  | <i>Example of how translation units are stored in a TM</i> .....             | 25 |
| Figure 1.3  | <i>Example of a misaligned segment</i> .....                                 | 26 |
| Figure 1.4  | <i>Example of an exact match</i> .....                                       | 27 |
| Figure 1.5  | <i>Example of a fuzzy match</i> .....  | 28 |
| Figure 1.6  | <i>Example of a sub-segment match</i> .....                                  | 31 |
| Figure 1.7  | <i>Example of a term match</i> .....   | 32 |
| Figure 1.8  | <i>Example of how a text is segmented in interactive translation</i> .....   | 34 |
| Figure 1.9  | <i>Example of a pre-translated text</i> .....                                | 35 |
| Figure 1.10 | <i>Example of a pre-translated text containing a sentence fragment</i> ..... | 39 |

## List of Tables

|           |  |    |
|-----------|--|----|
| Table 1.1 | <i>Example of multiple fuzzy matches for the same ST segment</i> .....                             | 29 |
| Table 1.2 | <i>Example of fuzzy matches ranked according to superficial character string resemblance</i> ..... | 30 |
| Table 1.3 | <i>Example of a pre-translated text as conceived by Bédard</i> .....                               | 37 |
| Table 1.4 | <i>Example of post-editing machine translation output</i> .....                                    | 41 |
| Table 1.5 | <i>Example of editing a pre-translated text produced by a TM system</i> .....                      | 41 |
| Table 3.1 | <i>Method of translation for each individual translator for each ST</i> .....                      | 61 |
| Table 4.1 | <i>Time required for each method of translation for each translator</i> .....                      | 64 |
| Table 4.2 | <i>Results of translation evaluations</i> .....  | 72 |

## **Chapter 0 – Introduction**

Translation technology has been increasing in popularity in recent years, and one of the most popular tools on the market today is the Translation Memory (TM) system. TMs are now being discussed in translation courses, at conferences, in professional magazines and academic journals. However, since it is only relatively recently that translators have begun using these tools, there has not yet been a substantial amount of research into the impact that these tools have on translators or on the texts that they produce. With this thesis, I hope to take some small steps towards filling this gap.

### **0.1 Objectives**

The main aim of this thesis is to investigate two different methods of working with a translation memory (TM) system – interactive mode and pre-translation mode – in order to determine if one method is more beneficial than the other with regard to increasing translation productivity, producing high quality texts, and generating a high level of “job satisfaction” among translators using TMs.

### **0.2 Background and Context**

Before delving into a detailed description of TM systems and how they operate, it is worth taking some time to explore the situation of the translation market, which has been

largely responsible for bringing about an increased interest in the use of this type of technology. This section will explore both some challenges and potential solutions present on the translation market.

### *0.2.1 Translation Market: Challenges*

Translators working in today's market face a number of challenges, including a growing demand for translation coupled with shorter deadlines and a shrinking workforce.

#### *0.2.1.1 Increased demand for translation*

Today, the volume of translation is rapidly increasing. This is due, in large part, to Globalization since, when selling products to foreign markets, it is necessary to provide information on those products. For example, Allen (2003:299) observes that “for companies to be successful on an ongoing basis, they must present themselves in a multilingual way”. Meanwhile, Thibodeau (2000:127) notes that translating a product to be sold internationally “may increase a company’s sales by at least 25 percent”. Understandably, then, there is a great business incentive for translating products as well as any accompanying documentation such as marketing material, user guides, warranty information and product websites.

Increased demand for translation is also being driven by the fact that new products (e.g., Blackberries, iPods, digital cameras) are constantly being introduced, and if these

products are to be sold on foreign markets, it is necessary to translate the documentation so that consumers are able to use and understand the products that are becoming essential in today's fast-paced society. This is especially true nowadays since today's products are much more sophisticated than before, which in turn means that longer documentation may be required, thus adding to the increased volume of material to be translated.

An additional factor in the increased demand for translation is the fact that we find ourselves living in an "information society". That is to say that we want information, and we want it in our own language. The internet has greatly facilitated access to information, and some organizations are beginning to make information on their websites available in multiple languages. For instance, consider the case of the Canadian federal government. As a bilingual country, Canada has a mandate to provide information in its two official languages (i.e., English and French); however, some federal departments, such as Agri-Food and Agriculture Canada, are going beyond these domestic needs to also provide information in languages such as Spanish, Greek and Japanese, particularly on parts of their website that deal specifically with foreign trade.<sup>1</sup> Similarly, in addition to providing information in English and French, the Canadian rail service Via Rail has made much of the material on its website available in Spanish, German, Dutch, Portuguese, Japanese and Korean, no doubt with the hope of attracting business from tourists who speak these languages.<sup>2</sup>

---

<sup>1</sup> For example, see the Agri-Food Trade Service web page at <http://atn-riac.agr.ca/supply/factsheet-e.htm> [Consulted April 29, 2006].

<sup>2</sup> See the Via Rail website at [http://www.viarail.ca/en\\_index.html](http://www.viarail.ca/en_index.html) [Consulted April 29, 2006].

### *0.2.1.2 Shorter deadlines*

Not only is the volume of translation increasing, but “pressure on translators and translation companies to produce high-quality results in shorter times is likewise increasing” (Andrés Lange and Bennett, 2000:203). For instance, since web pages can be updated quickly and easily, it is necessary for the translated versions of the web pages to be updated just as quickly. This is especially important in the case of dynamic web pages, where the information is generated at the time of viewing (Cheng, 2000:29).

Similarly, many companies are looking to release their products and supporting documentation in multiple languages at one time in a process known as simultaneous shipment or “simship” (Bowker, 2002a:13). According to Anobile (2000:vii), some companies are releasing products in as many as 30 languages at once. Since companies cannot release a product unless the product and its accompanying documentation have been translated, translators are on tight deadlines so as not to hold up release of the product. The faster the translations are completed, the faster the product can be put on the shelves in the global market.

### *0.2.1.3 Shortage of translators*

Despite the fact that the volume of translation is increasing, there is a current shortage of trained translators (Anobile, 2000:vii), which is being exacerbated by the fact that current workers in the field of translation are aging and many of those belonging to the “babyboomer” generation will soon be retiring (CTISC, 1999:18). Although there is a

growing demand for translators, there are relatively few graduates of translation programs entering the industry (CTISC, 1999:19), which, as pointed out by Andrés Lange and Bennett (2000:203), means that there will continue to be a significant lack of translators to meet the demand of a growing volume of translation projects. This situation puts additional pressure on those few translators who currently make up the translation workforce.

### *0.2.2 Translation Market: Some Impacts of Technology*

Due to the fact that there are a great number of texts to be translated in a shorter amount of time by fewer individuals, more and more translators are turning to technology to see if it can help them to increase their productivity, while still maintaining a high level of quality in their work. The following sections will briefly describe some of the ways in which this technology is having an impact on the translation market and on translators themselves.

#### *0.2.2.1 Tools available to translators*

As discussed by a number of researchers (e.g. L'Homme 1999; Esselink 2000; Austermühl 2001; Bowker 2002a), there are many technological tools available to translators, such as corpus-analysis tools, terminology management systems, translation memory (TM) systems, and machine translation systems, as well as online resources and useful features in word-processing software (e.g., compare documents, track changes).

Out of these tools available to translators, TM systems are currently among the most popular and they will be the focus of this thesis.

A detailed description of a TM will be provided in Chapter 1; however, for the moment, it will suffice to say that a TM is a tool that stores previously translated texts in such a way that sections of these texts can be reused in new translations. A TM can therefore help translators to ‘recycle’ portions of a previous translation in a new translation. In fact, Yunker (2003:221) suggests thinking of a translation memory as a “recycle bin for your translated text”.

#### *0.2.2.2 TMs and the question of ownership*

Because of their potential to increase translator productivity, TM databases are proving to be a very valuable translation commodity, and both translators and their clients would like to claim ownership over the TM database. However, as pointed out by Topping (2000:59), from a legal perspective, neither party appears to have clear-cut ownership rights, so the matter is being hotly debated. For instance, a perusal of the postings on the *Translator’s Café* online forum – an electronic discussion board where professional translators can go to ask questions, exchange ideas, and discuss topics of relevance to the translation profession – reveals that many translators believe that since they are the ones putting their time and effort into building up the TM, they should have ownership of it.<sup>3</sup> One translator participating in the forum claims that it is his “intellectual property” and

---

<sup>3</sup> To read a thread on the Translator’s Café online forum that deals with ownership issues, see <http://www.translatorscafe.com/cafe/MegaBBS/thread-view.asp?threadid=4881&start=61> [Consulted April 29, 2006].

many others agree. Seldom are there posts where translators indicate that they would willingly give up a TM that they have been working on, unless of course the TM in question was given to them by their client to begin with and it therefore consisted of an already established database to which they simply added a few segments. It is interesting to note, however, that while translators may be reluctant to give TMs to their clients, some of them are exchanging these databases with other translators. This raises the question as to whether or not this type of exchange is appropriate because of the current confusion over ownership of the databases.

Meanwhile, Yunker (2003:226) presents the ownership issue from the client's perspective, noting:

Translation memory, as it increases in size, increases in value to your company. But suppose you decided to switch translation vendors [i.e., translators], and your vendor [translator] refused to return your memory. Some vendors [translators] claim ownership of the translation memory because it's a product they helped develop. As a rule, make sure your contract guarantees that you retain full ownership of the translation memory.

Until the slow-moving legal system catches up with the fast-paced world of translation technology, it is likely that ownership of TM databases will remain a contentious issue.

### *0.2.2.3 TMs and the question of payment*

Yunker (2003:221), again speaking from the client's perspective, also states that "once you translate a sentence, you shouldn't have to pay for it to be translated again". That is

to say that if a client is providing a TM to any translator he or she contracts to do a job, the client should only have to pay for newly translated segments since the translator is reusing segments of texts that the client has already paid to have translated. Many TM systems have a built-in repetition analysis module which is used to compare new source texts against the TM database to determine approximately how much of the text can be recycled (Bowker 2002a:121). This is presenting a challenge to translators because clients are beginning to demand discounts for any text translated with the help of a TM. According to Bowker (2002a:121), some clients will pay full price for a segment that has no match in the TM, but will only pay part of the standard rate for a segment which retrieves a fuzzy match.<sup>4</sup> As noted by Esselink (2000:364), generally speaking, a client will pay 60% of the full word rate for a fuzzy match and only 30% for an exact match. However, a few clients refuse to pay anything at all for a sentence that retrieves an exact match from the TM database.

Translators are understandably concerned about such attitudes. Again, examination of the *Translator's Café* online forum reveals that very few of the translators participating in the forum agree that clients should pay less for fuzzy matches.<sup>5</sup> Although many grudgingly tolerate not getting paid for 95%-100% matches, they still maintain that they should get *some* compensation since verifying the matches still takes time. For instance, one translator brings up the fact that when using a TM with languages such as Latin and

---

<sup>4</sup> A more detailed discussion of different types of matching will be presented in section 1.3. However, for the time being, it should be noted that a fuzzy match is a partial match that has some elements in common with the new text to be translated, but which will need to be edited – sometimes substantially – in order to produce a useful target text.

<sup>5</sup> For a thread discussing payment issues and TMs on the Translator's Café online forum, see <http://www.translatorscafe.com/cafe/MegaBBS/thread-view.asp?threadid=3499&messageid=46003#46003> [Consulted April 29, 2006].

Finnish, the fuzzy matches do not correctly render the grammar and syntax of the target language, so not getting paid for fuzzy matches is unreasonable. Meanwhile, another translator has a different way of dealing with the situation: when asked for a discount for fuzzy matches, he simply accepts the fuzzy match as it is and lets the client do the modification after the fact.

### *0.2.3 Current Situation*

The challenges described in section 0.2.1 and the consequent move towards the use of TMs as outlined in section 0.2.2 have resulted in a situation that can essentially be described as follows. Firstly, due to the increasing volume of texts to be translated and the shorter deadlines in which to translate them, many translators would like to use TMs to accelerate the translation process. Furthermore, clients want translators to use TMs because it can potentially save them time and money: they can get their translated material faster in order to get their products on the market, and they might not have to pay a translator who uses a TM as much as they would have to pay a translator who does not use a TM. In theory then, the use of a TM is a win-win situation. However, things may not always go as smoothly in practice.

For one thing, TMs are a relatively expensive technology, costing roughly between \$1000 and \$2500 per copy (Yunker, 2003:225), and this is one reason that not all translators own a TM. In addition, there are many competing TM products on the market, which means that even if a translator does purchase a TM, it may not be the same one that the

client uses and therefore the translator may not be an experienced user of that system.<sup>6</sup> Moreover, the TM databases created by different systems may not be compatible because each different TM system stores information using a proprietary format.<sup>7</sup> This means that in a market which is already experiencing a shortage of translation professionals, the pool of translators available to clients who insist on the use of a TM – not to mention a particular TM – is even smaller.

One solution to this problem could be for the client to provide a translator with access to the client's TM system and databases for the duration of the project. Some clients do choose this option; however, many other clients are frequently hesitant to give translators direct access to their TMs because of proprietary issues. The client does not want the translator using their TM to do work for another client. Even the signing of a non-disclosure agreement is sometimes not enough to allay the fears of some clients in this regard.

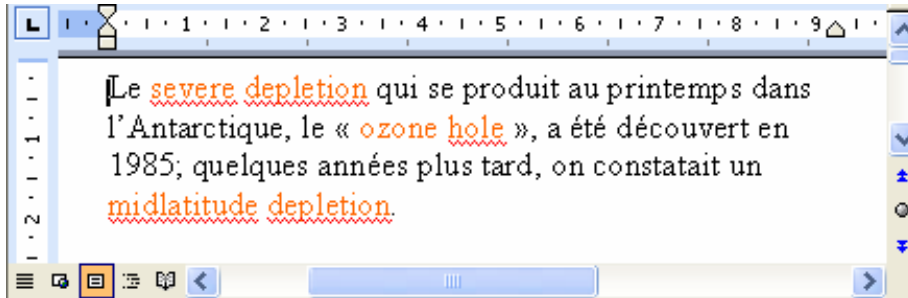
Therefore, another solution that is beginning to surface is one that makes use of a strategy called pre-translation. In a TM system, pre-translation is a function that compares the new source text against the contents of the TM database and automatically inserts directly into the source text any target-language matches that are found. As illustrated in Figure

---

<sup>6</sup> For example, Trados Translator's Workbench, Déjà Vu, SDLX, STAR Transit, and Fusion Translate are just some of the TM products available on the market today.

<sup>7</sup> There is a movement being spearheaded by a special interest group of the Localization Industry Standards Association (LISA) known as OSCAR (Open Standards for Container/Content Allowing Re-use). OSCAR is working on the development of a standard format for TM databases known as the Translation Memory eXchange (TMX) format. If all systems adopt this standard, the problem of file format compatibility will no longer be relevant; however, we have not yet arrived at that situation since many of the available TM systems on the market do not yet support TMX. See the LISA website for more information on OSCAR and TMX: <http://www.lisa.org/standards/tmx/> [Consulted May 18, 2006].

0.1, the result is a partially translated or hybrid text containing some fragments in the target language (i.e., sections of the text for which a match was recycled from the TM) and others that still remain in the source language (i.e., those sections for which no match could be found in the TM).



**Figure 0.1**

This figure shows an example of a “pre-translated” text where the untranslated portions of the original source text are shown in French, and the pre-translated chunks of English text have been recycled from the TM.

The client can then give the hybrid text to the translator whose job is to translate the remaining parts of the text for which no matches could be found. In principle then, the client can provide the translator with material from the TM database without actually having to give the translator direct access to the TM. However, one important question that remains to be answered is how pre-translation compares to the interactive approach of working with a TM. Are the two approaches more or less equal, or is one approach preferable to the other when it comes to issues of productivity, quality, and translator satisfaction? This is the question that I aim to answer based on the research reported in this thesis.

### 0.3 Justification and Motivation for Research

Clients today are looking for productivity and quality, but at the same time are looking to reduce costs. There have been many documented cases where clients have tried using machine translation - which is certainly faster and cheaper than human translation - but the resulting texts are almost always of an unacceptable quality unless they are subsequently post-edited by human translators.<sup>8</sup> The problem with this strategy is that, as pointed out by Allen (2003:303), it can frequently be more time consuming to post-edit machine translation output than to translate a text from scratch, and the quality of post-edited MT is often lower than that of human translation. Therefore, clients have high hopes for TM systems since they recycle texts that have been translated by humans and since they are intended to be used as an *aid* for human translators rather than as a *replacement* for them.

Nevertheless, as will be discussed in more detail in Chapter 1, although working with a pre-translated text is not exactly the same as post-editing machine translation output, there are some parallels between the two processes. Therefore, it is possible that the solution of providing translators with a pre-translated text instead of giving them access to a TM with which they can work interactively may actually be hindering productivity and reducing the quality of the translations that are being produced as is the case with post-editing MT output. Furthermore, since many translators do not like to work as post-editors, as observed by Allen (2003:303), it is possible that translators may not be happy

---

<sup>8</sup> There are a few exceptional cases where little or no post-editing is required, such as that of the Météo machine translation system, which has been used for many years to translate weather forecasts from English into French for Environment Canada (Somers, 2003b:289-294). However, this type of extremely restricted subject field is very rare.

working with pre-translated texts either, given the resemblance between the processes of post-editing and the revision of pre-translations. By conducting an experiment to compare the impact of interactive use of a TM against pre-translation with regard to productivity, quality of texts produced, and translator satisfaction, I intend to try to provide answers to these questions. The information resulting from this research could be useful for helping clients to decide whether they should provide their translators with pre-translated texts, or whether they might be better off giving them access to the TM databases in order to work interactively with the TM system.

#### **0.4 General Methodology**

The overall methodological approach employed in the context of this thesis can be divided into a number of steps as follows:

1. Conduct a literature survey to determine what researchers have already learned about TMs and their use in interactive translation and pre-translation modes.
2. Design an experiment to compare the use of interactive translation and pre-translation and the impact of these two approaches on productivity, quality and translator satisfaction.
3. Conduct preparatory work, including building a TM database, stocking the associated termbase, and generating pre-translations.
4. Carry out a pilot test whereby translators will translate texts using both approaches, and the resulting translations will be assessed by external evaluators.
5. Analyze the data generated from the pilot test.

6. Draw preliminary conclusions, suggest possible modifications to the experimental design, assess the overall project, and suggest possible directions for future research.

## **0.5 Scope and Limitations**

For practical reasons of time and economic constraints, this study was limited in a number of ways. Certain restrictions had to be placed with regard to the TM system, the languages, the participants, and the texts.

### *0.5.1 TM System*

In the experiment conducted for the purpose of this thesis, the translation memory system used was Fusion Translate™ (version 1.8.0.56), developed by Orca Development Corporation.<sup>9</sup> While there are many TM systems currently available, all of which can operate in interactive mode, not all systems have a comprehensive pre-translation function. Fusion Translate is one of the newest tools on the market and it offers a state-of-the-art pre-translation option. In addition, from a purely practical point of view, Fusion Translate is available for use at the Writing Centre computer laboratory of the University of Ottawa, and it was taught as part of the graduate course *TRA 5903: Computers and Translation*. This meant that I could recruit students from that course to

---

<sup>9</sup> For more information of Orca Development Corporation, developers of Fusion Translate, see: <http://www.orcadev.com/> [Consulted May 1, 2006].

participate in the experiment, and I could be sure that they already had a working knowledge of this tool.

### *0.5.2 Languages*

As a native speaker of English, I translate primarily from French into English, so I limited myself to working with that specific language direction for the research reported here.

### *0.5.3 Participants*

Two types of volunteer participants were recruited for the experiment conducted as part of this thesis: translators and evaluators. The precise criteria that were used to select participants will be described in section 3.1. For the time being, I simply want to note that, for practical reasons, I was limited to using four translators and two evaluators. Clearly this number is much too small to enable me to generate significant statistics or to make valid generalizations; however, I do hope that the pilot study presented in this thesis might nevertheless provide interesting preliminary results that could indicate whether or not it is worth conducting a future scaled-up study in this area. In addition, this pilot study will provide an opportunity to develop a sound methodology for comparing the effects of interactive translation against those of pre-translation, and this methodology could later be used as part of a larger study.

#### 0.5.4 *Texts*

A detailed description of the criteria used to select texts for this experiment will be presented in section 3.1; however, it is worth noting that for practical reasons, including limits on time and resources, only two source texts were used for the experiment. Both texts were approximately 350 words in length and both dealt with the same subject: ultraviolet radiation and the ozone layer.

### **0.6 Outline**

This thesis has been divided into three main parts. Following on from this introductory chapter, Part I consists of two chapters that provide background information and a theoretical framework that helps to contextualize this investigation. Chapter 1 summarizes the history of TM technology, provides a detailed description of how a TM system works, and presents an in-depth look at two different approaches to working with a TM: interactive mode and pre-translation mode. Based on the literature survey contained in Chapter 1, as well as on my own practical experience working with TM systems, three hypotheses are developed and presented in Chapter 2. Part II contains two chapters that focus specifically on the pilot study. Chapter 3 describes the methodology used to design and carry out the pilot study experiment, while Chapter 4 contains an analysis and discussion of the data gathered from this experiment. Part III consists of a single chapter – Chapter 5 – which contains an evaluation of the research carried out, some general closing remarks, and some suggestions for future research. To complete

this volume, a list of bibliographic references is provided, along with a number of appendices containing material used for the pilot study.

***PART I:  
BACKGROUND INFORMATION AND THEORETICAL  
FRAMEWORK***

## Chapter 1 – Introduction to Translation Memories, Interactive Translation, and Pre-translation

### 1.1 From Initial Conception to Present-day Popularity

A Translation Memory (TM) is a type of computer-aided translation tool which stores previously translated texts alongside their corresponding source texts (ST) and allows translators to ‘recycle’ or re-use these texts, or parts of them, in new translations. The original idea for a TM is usually attributed to a paper written by Martin Kay in 1980 called “The Proper Place of Men and Machines in Language Translation”.<sup>10</sup> In that paper, Kay ([1980]1997:19) proposed the development of a *Translator’s Amanuensis*, which, among other things, would enable a translator to “examine past and future fragments of texts that contain similar material”. As we will see in section 1.3, this is indeed one of the important principles upon which a TM operates.

Alan Melby and his colleagues at Brigham Young University (BYU) were also instrumental in the development of early TM systems. By the early 1980s, two different systems had been developed at BYU (Melby 1995:225). One, called a “Suggestion Box”, included three different levels. The first level was meant for projects where the source text was not available in machine-readable form, and the second level was meant for projects where it was, while the third level was meant to work in conjunction with a

---

<sup>10</sup> This article first appeared as a Xerox PARC Working Paper in 1980, but it has since been reprinted in the journal *Machine Translation*, vol. 12 (1/2), pp. 3-23, 1997. It is to this re-printed edition that I am referring with regard to page numbering.

machine translation system (Melby, 1995:189).<sup>11</sup> The other system that was implemented in the 1980s was actually an interactive machine translation system, rather than a TM system, but it incorporated some of the principles of a TM. Based on research carried out at BYU, the system called “Repetitions Processing” was commercially developed by Alps Inc. (Somers, 2003a:32).

Other translation companies also developed in-house systems for searching through their archives of past translations (Bowker, 2002a:92), but it was not until the mid 1990s that true TM systems – as we know them today – became commercially available, and they have been growing in popularity ever since (Somers, 2003a:32).

As an increasing number of translators have begun using these tools, TMs have correspondingly become the subject of a growing number of research investigations aimed at helping the translation community at large learn more about the effects that these tools can have on the translation process and product, as well as on the translator. For instance, Macklovitch and Russell (2000) and Bowker (2002b) study TMs from a development point of view, considering the shortcomings of existing TM systems and suggesting ways to make them more useful, such as modifying search techniques to take into account syntactic and semantic information (e.g., inflection, derivation and synonymy). Bowker (2002b), for example, suggests that superficial character-string matching presents a problem in that matches are proposed based on the degree of similarity between the physical appearances of the sentences and does not take semantic

---

<sup>11</sup> Some present-day TM systems work in a way similar to that of the third level concept and provide the option of linking to a machine translation system.

similarity into account. This means that, as discussed in more detail in section 1.3.2, the matches suggested by a TM might not be the best semantic matches. Bowker suggests employing a controlled language<sup>12</sup> to ameliorate the character-string matching, or using part of speech tagging to identify the syntactic similarity of the sentences. Other researchers, such as Schäler (2001) and Gow (2003) look specifically at some of the different types of matches proposed by TM systems (e.g., fuzzy vs sub-segment matching) and analyze which of these are more helpful to translators in various circumstances. Schäler (2001) suggests using what he calls a “Phrasal Lexicon” meaning that sentences are matched at the phrasal level instead of sentence level so that no fuzzy matching is required. According to Schäler, and as will be discussed in section 1.3.3, matching at the phrasal level can have more advantages than matching at the sentence level because the TM will be able to offer more exact matches and “...will be able to translate a larger amounts [*sic.*] of source text automatically without the need to adapt fuzzy matches manually” (Schäler, 2001:3), thus increasing productivity. Still other researchers compare the features of TMs against those of other tools, such as conventional corpus analysis tools (Bowker and Barlow 2004) or machine translation systems (Circé 2005), analyzing the strengths and weaknesses of each.

Productivity and quality are, of course, major points of interest for TM researchers. In terms of productivity, Webb (1998) and Circé (2005) investigate the advantages of using

---

<sup>12</sup> A controlled language (CL) is typically understood to be a subset of natural language that has a restricted vocabulary and a restricted syntax. The goal is to promote comprehension by reducing the ambiguity that is inherent in language. A CL might therefore dissuade the use of synonyms or of -ing verb forms or passive constructions, which have the potential to cause confusion. CLs are often used in combination with machine translation on the understanding that reducing ambiguity in the input text can help to improve the quality of the output text. For more information on CL, see Nyberg et al. (2003).

a TM system as compared to not using one, confirming that TMs do indeed have the potential to help translators increase productivity when used under the right circumstances. With regard to quality, Bédard (2000) and Mogensen (2000) focus specifically on the negative impact that TM use can have on the quality of a translation, such as making the resulting text less readable because of the relative absence of cohesive devices between sentences. Mogensen (2000) explains that using a TM could result in the creation of oversimplified translations because the sentences are often short and choppy, resulting in a monotonous rhythm. She indicates that instead of translation tools being modified to accommodate language, language is being modified to accommodate translation tools (e.g. controlled language) in order to make texts easier for these tools to process. This, however, can result in a target text whose syntax closely models that of the source text and is therefore less readable in the target language. Meanwhile, Bowker (2005) conducted a study that focused on productivity and quality simultaneously, demonstrating how it can be difficult for translators to achieve a balance because one often comes at the expense of the other when using a TM.

In another study, Bowker (2003a) examined the advantages and disadvantages of using a single TM database in both language directions, showing that when searching from L1 to L2, translators are likely to retrieve a greater number of matches, but these are of a lower quality, while translators searching from L2 to L1 are likely to find fewer matches, but the ones they do find are of a higher quality.

Finally, with regard to translator satisfaction, Merkel (1998) carries out a study where he examines the attitude of a group of technical translators toward terminological consistency and variation when working with TM systems, finding that translators can become frustrated by the way that consistency is sometimes inappropriately imposed on the texts by these systems.

What is striking is that in all of these investigations, the use of TMs has been studied in interactive mode only. This is probably because interactive mode has been, and continues to be, the most prevalent way of using these systems. However, as noted in section 0.2.3, the use of TMs in pre-translation mode is now beginning to gain some ground. Therefore, it seems appropriate at this point in time to undertake a comparative evaluation of these two methods of employing a TM. If we want to use TM tools to their highest potential, we need to find out more about how the different ways of using them can affect factors such as productivity, quality and translator satisfaction. The research presented in this thesis aims to go some way towards filling that gap; however, before we focus on this particular goal, it is necessary to first take some time to gain a better understanding of how TMs tools actually work in general, and in interactive and pre-translation modes in particular.

## **1.2 Organization and Storage of Information in a TM**

In order to understand how TM systems actually operate, it is necessary to be familiar with the way these tools organize and store information. This in turn requires an understanding of the basic concepts of segmentation and alignment.

### 1.2.1 Segmentation

TM tools process texts in small chunks known as segments. A segment usually corresponds to a sentence or some other sentence-like structure, such as a title, a heading, an item on a list, or the contents of a cell in a table. Usually punctuation marks (e.g., periods, question marks) or line breaks indicate the end of a segment, but it is also possible for the user to specify what should or should not represent the end of a segment (Bowker, 2002a:94-95). When texts are added to a TM system, the tool automatically divides each text into separate segments, as shown in Figure 1.1. Text segmentation in turn allows for alignment of corresponding segments of the source and target texts.

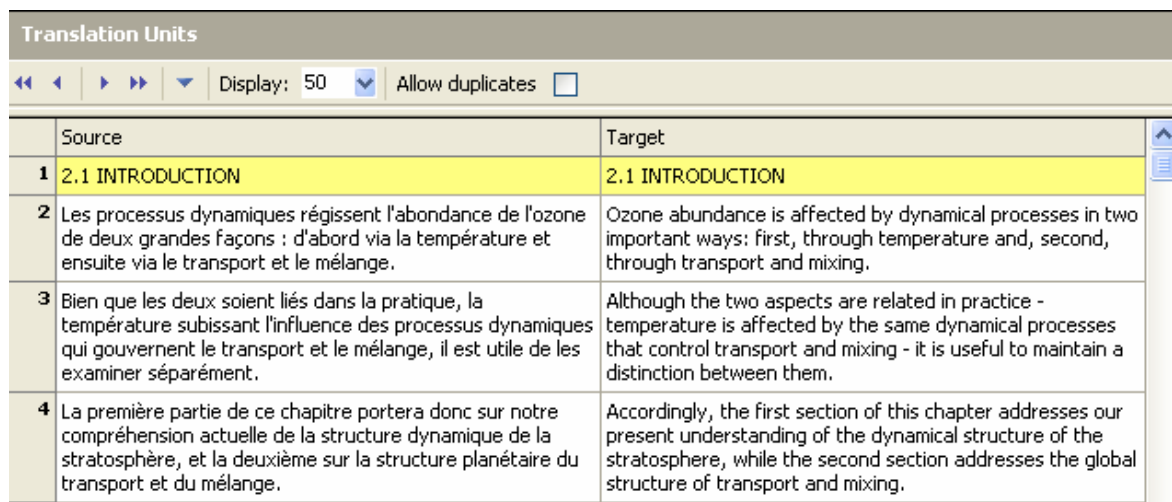
|   |   |
|---|---|
| <p><b>The Montreal Protocol: the Achievement and the Challenge</b></p> <p>The signing of the Montreal Protocol on September 16, 1987, was a remarkable and significant event in modern diplomatic history, one of those rare occasions when individual nations subordinated economic self-interest to the achievement of a common planetary goal. The event was even more remarkable when one considers that it was accomplished in spite of scientific uncertainties about detailed aspects of the depletion process and without immediate evidence of impacts on ecosystems and human health. That an agreement was eventually reached was due not only to an extraordinarily successful collaboration between scientists and policymakers but also to the enormous strides made by the international scientific community in expanding the boundaries of ozone science. The solidity of their achievement can be seen in the very real progress that has been made since 1987 in reducing emissions of ozone-depleting substances.</p> | <p><b>The Montreal Protocol:</b></p>  |
|   | <p><b>the Achievement and the Challenge</b></p>   |
|   | <p>The signing of the Montreal Protocol on September 16, 1987, was a remarkable and significant event in modern diplomatic history, one of those rare occasions when individual nations subordinated economic self-interest to the achievement of a common planetary goal.</p>    |
|   | <p>The event was even more remarkable when one considers that it was accomplished in spite of scientific uncertainties about detailed aspects of the depletion process and without immediate evidence of impacts on ecosystems and human health.</p>                              |
|   | <p>That an agreement was eventually reached was due not only to an extraordinarily successful collaboration between scientists and policymakers but also to the enormous strides made by the international scientific community in expanding the boundaries of ozone science.</p> |
|   | <p>The solidity of their achievement can be seen in the very real progress that has been made since 1987 in reducing emissions of ozone-depleting substances</p>  |

**Figure 1.1**

The text shown on the left-hand side of the figure has been divided into segments as shown on the right.

## 1.2.2 Alignment

In order for a TM to retrieve matches, each segment of a source text must be aligned with its target text equivalent. Aligning texts means matching up source and target text segments and linking them together to form translation units (TUs). Figure 1.2 shows how TUs are stored in the TM. An alignment tool is integrated into most TMs which allows for alignment of existing translations and their corresponding source texts (Bowker, 2002a:109). It is, however, advisable to view the automatically aligned segments to ensure that the alignment is accurate, and in the case of misalignment, to manually correct the alignment.



|   | Source   | Target  |
|---|--|---|
| 1 | 2.1 INTRODUCTION   | 2.1 INTRODUCTION  |
| 2 | Les processus dynamiques régissent l'abondance de l'ozone de deux grandes façons : d'abord via la température et ensuite via le transport et le mélange.   | Ozone abundance is affected by dynamical processes in two important ways: first, through temperature and, second, through transport and mixing.   |
| 3 | Bien que les deux soient liés dans la pratique, la température subissant l'influence des processus dynamiques qui gouvernent le transport et le mélange, il est utile de les examiner séparément.    | Although the two aspects are related in practice - temperature is affected by the same dynamical processes that control transport and mixing - it is useful to maintain a distinction between them.                 |
| 4 | La première partie de ce chapitre portera donc sur notre compréhension actuelle de la structure dynamique de la stratosphère, et la deuxième sur la structure planétaire du transport et du mélange. | Accordingly, the first section of this chapter addresses our present understanding of the dynamical structure of the stratosphere, while the second section addresses the global structure of transport and mixing. |

**Figure 1.2**

This figure shows how translation units are stored in a translation memory system.

If the alignment is incorrect, as shown in Figure 1.3, then the proposed equivalent that will be shown to the translator will not be accurate and will therefore be of no help whatsoever.

| Translation Units                                     |   |  |
|---|---|--|
| Display: 50 <input type="checkbox"/> Allow duplicates |   |  |
|   | Source  | Target   |
| 196   | Cette différence dans la localisation des vitesses de descente maximales peut s'expliquer par le fait que la bordure du vortex est une zone de mélange latéral significatif.  | Air parcels initially located at the vortex edge do not remain there, but perform large excursions into the midlatitude surf zone where they experience much less radiative cooling.   |
| 197   | Les particules d'air situées au départ en bordure du vortex n'y restent pas, mais entrent considérablement dans la zone de déferlement des latitudes moyennes, où elles subissent beaucoup moins de refroidissement radiatif. | Hence, the descent rate of the (Eulerian) mean mass circulation at the vortex edge is a significant overestimate of the actual Lagrangian descent of air parcels.  |
| 198   | Donc, la vitesse de descente de la circulation de masse moyenne (eulérienne) sur la bordure du vortex constitue une forte surestimation de la vitesse lagrangienne réelle des particules.                                     | As noted in Section 2.2, dynamical variability of the wintertime polar vortex is much greater in the Arctic than in the Antarctic, and temperatures sufficiently cold to allow PSC formation therefore occur far less often. |

**Figure 1.3**

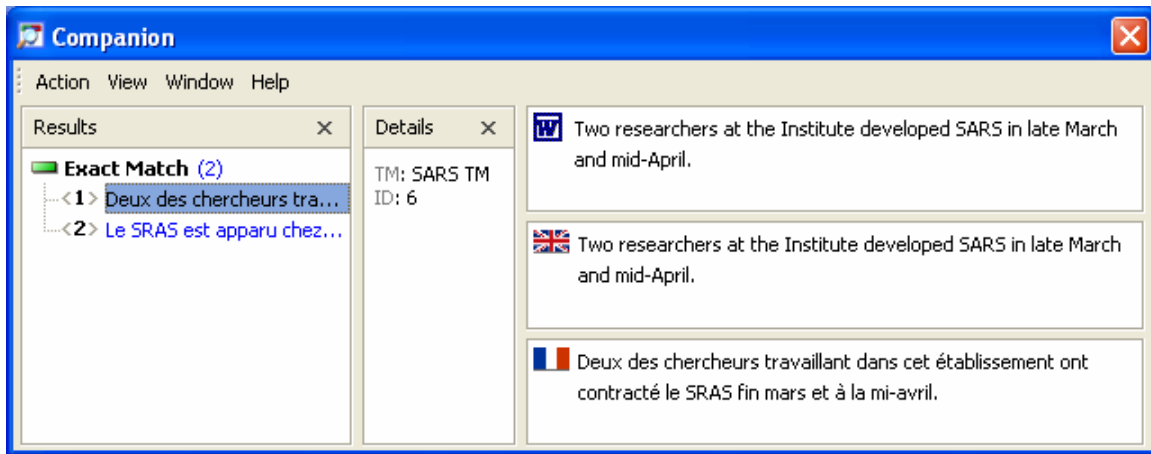
This figure shows a misaligned pair. In this case, the translator would need to align the segments manually.

### 1.3 Retrieval of Information from a TM

Now that we understand how information is stored in a TM, we can discuss how the retrieval process works. When a translator is given a new source text to translate, he or she opens that text in a text editor (e.g., a word processor), which interacts with the TM system. The first thing that the TM system does is divide this new source text into segments, as was shown previously in Figure 1.1. The TM then takes the first segment of this new ST and compares it against all the segments stored in the TM database. This comparison is done strictly on the basis of “superficial character string matching” (Bowker, 2002b:198) or “character-string similarity” (Somers, 2003a:38), meaning that the spelling, punctuation and all other parts of the segments are taken into account when the TM system tries to find matches in the database (Bowker, 2002b:200). When the system does find a match, the TM retrieves the matching TU from the database and presents it to the translator, who can accept this proposal, modify it, or reject it. There are a number of different types of matches that a TM system can retrieve, and these will be examined below.

### 1.3.1 Exact Match

The most straightforward type of match is an exact match, which means that the new source text segment is identical in every way to a segment that is already stored in the TM database. Figure 1.4 contains an example of an exact match. The segment shown in the upper pane (beside the Word icon) is the one taken from the new source text. The segment shown in the middle (beside the British flag) is the one that has been retrieved from the TM database, along with its translation which is shown in the bottom pane (beside the French flag).



**Figure 1.4**  
This figure shows an exact match

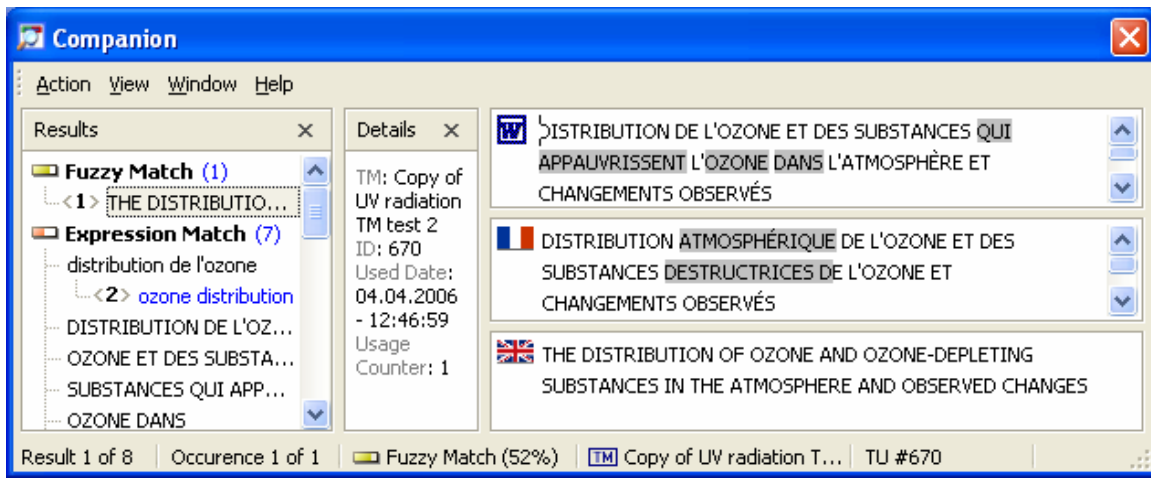
### 1.3.2 Fuzzy Match

With the exception of a few types of translation situations, it is rare for a significant number of segments from two different texts to be exactly the same.<sup>13</sup> For this reason, the

<sup>13</sup> Exact matches are commonly found, for example, in cases where a translator is translating a revision of a previous document. For instance, as noted by Trujillo (1999:59), a user manual for a product is a type of text that is frequently updated. Another example is a collective agreement. In such a case, the author does

powerful concept of fuzzy matching was developed.

As shown in Figure 1.5, a fuzzy match algorithm allows the TM to search in the database for segments that are similar to, but not exactly the same as, segments contained in the text to be translated. A translator can edit the fuzzy match as necessary in order to produce a valid translation for the source text segment. The parts of the retrieved segment that differ from the new source text will be highlighted to indicate what needs to be adjusted in the current translation in order to match the corresponding original. Some TM systems use different colors to indicate different types of changes that need to be made to the target text such as additions, omissions and modifications.



**Figure 1.5**

This figure shows a fuzzy match. Parts of the sentences have been highlighted to indicate differences between the new ST sentence and the sentence that is already stored in the TM.

---

not usually begin from scratch, but rather, s/he takes the previous version (e.g. version 1) and makes additions or modifications as necessary before sending the updated text (version 2) to the translator. If the translator has stored the translation of version 1 in the TM, then it will be possible to recycle relevant passages from version 1 into the translation of version 2. However, there are a limited number of text types that get updated in this manner. In most cases, the repetition of material from one text to the next is less extensive.

When working with a TM system, the translator can adjust the minimum match value according to the degree of fuzziness desired. For example, the translator can specify whether the two segments should be 90% similar in order to be retrieved, or 80%, or 70%, and so on. As indicated by Bowker (2002b:200), this is actually a challenging task: the translator must be cautious when selecting the degree of fuzziness so as to avoid problems of silence and noise. If the match value is set too high (e.g., 95%), the system may not retrieve matches that could be potentially helpful to a translator. In information retrieval, this type of problem is known as “silence”. In contrast, if the match value is set too low (e.g., 10%), then too many segments, or segments that are not useful, could appear leading to a problem known as “noise”, which may actually hinder the efficiency of the translation process.

It should be noted that it is possible for a TM system to retrieve more than one fuzzy match for any given segment, as illustrated in an example borrowed from Brockmann (1999:9) that has been reproduced in Table 1.1.

|   |   |
|---|---|
| <i>Original source text segment</i>               | What exactly is a <i>translation memory</i> ?   |
| <i>86% fuzzy match retrieved from TM database</i> | What precisely is a <i>translation memory</i> ? |
| <i>72% fuzzy match retrieved from TM database</i> | What is meant by <i>translation memory</i> ?    |

**Table 1.1**

Multiple fuzzy matches for the same ST segment retrieved from a TM database.

All of the fuzzy matches that are retrieved are added to the list of matches, thus allowing the translator to choose among several possibilities. As pointed out by researchers such as

Macklovitch and Russell (2000:138-139), Bowker (2002b:200) and Somers (2003a:39-40), the second (or third, etc.) match may in fact be closer to what the translator would actually like to put into the new translation. This is because, as mentioned in section 1.3, a TM ranks matches on the basis of the superficial similarity of the character strings in the segments rather than on the semantic similarity of the segments. For instance, consider the following example, borrowed from Macklovitch and Russell (2000:138-139) and reproduced in Table 1.2, which demonstrates how semantic similarity may not be well represented by superficial similarity.

|   |   |
|---|---|
| <b>New source text segment</b>                      | The wild child is destroying his new toy.       |
| <b>TU1: “Best” fuzzy match identified by TM</b>     | The wild chief is destroying his new tool.      |
| <b>TU2: “2nd best” fuzzy match identified by TM</b> | The wild children are destroying their new toy. |

**Table 1.2**

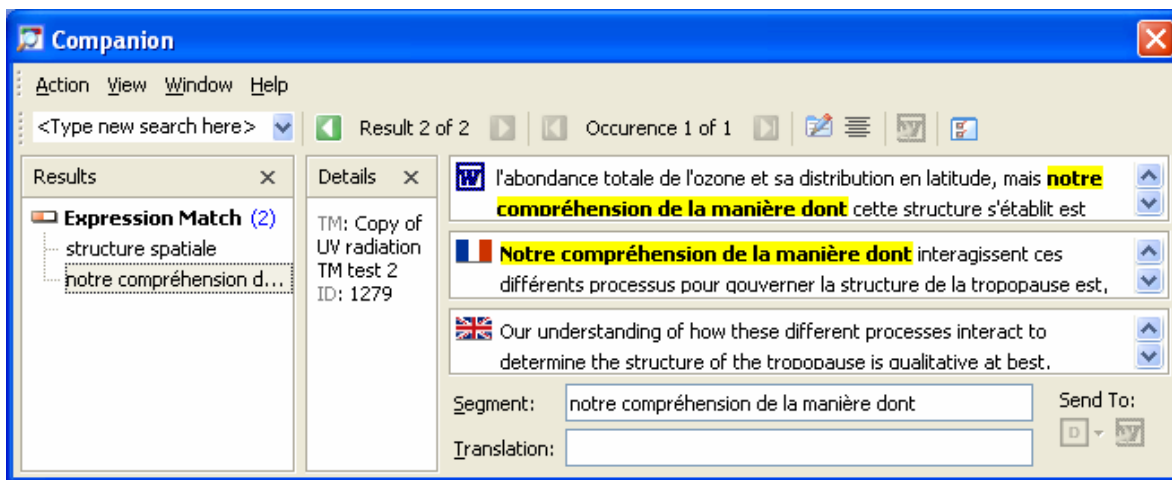
Fuzzy matches ranked according to superficial character string resemblance.

In this example, TU1 is determined by the TM system to be the “best” fuzzy match because it differs from the original segment by only four characters, while TU2 differs from the original segment by nine characters. Thus, superficially, TU1 is the most similar, even though TU2 is more similar in terms of semantic meaning, which is actually more helpful to the translator. By presenting all possibilities, the TM system leaves the choice to the user as to what match is most helpful for providing a solution in the current translation context.

### 1.3.3 Sub-segment Match

According to Schäler (2001), “matching segments at sentence level unnecessarily restricts the potential and the usefulness of translation memories as extremely valuable linguistic resources” because “the probability of finding exact matches at a lower phrasal level (e.g., at NP, VP or PP level) is significantly higher than the probability of finding exact matches at the current sentence level (i.e., the current TM segment level).”<sup>14</sup>

For this reason, as pointed out by Bowker (2002a:103), some newer TM systems now allow for sub-segment matches, where the system compares chunks of segments instead of whole segments.<sup>15</sup> This means that if two complete segments do not have many elements in common, smaller chunks of the segments might, and thus a match could be found. Figure 1.6 shows an example of a sub-segment match.



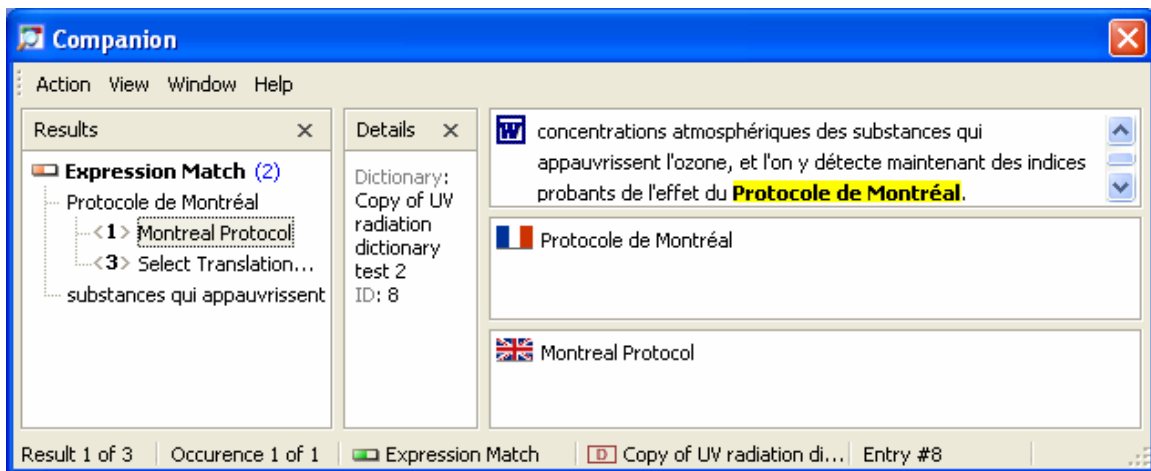
**Figure 1.6**  
This figure shows an example of a sub-segment match.

<sup>14</sup> NP = noun phrase, VP = verb phrase, PP = prepositional phrase

<sup>15</sup> A sub-segment match is called an Expression Match in Fusion Translate™.

### 1.3.4 Term Match

It is also possible for the TM to suggest term matches when no exact or fuzzy matches are available.<sup>16</sup> Since most TM systems are associated with termbases, the TM can compare terms found in the text with terms that have already been entered into the termbase by a translator (Bowker, 2002a:101). If a term is found in the associated termbase, the TM will display the term and its translation. This is known as a term match, and an example is provided in Figure 1.7.



**Figure 1.7**

This figure shows a term match. The term *Protocole de Montréal* was found in the TM dictionary and the equivalent, *Montreal Protocol*, is suggested to the translator.

The translator can then automatically insert the term into the new target text and continue translating. As we will see in section 1.4.2, term matches play an especially important role in pre-translation.

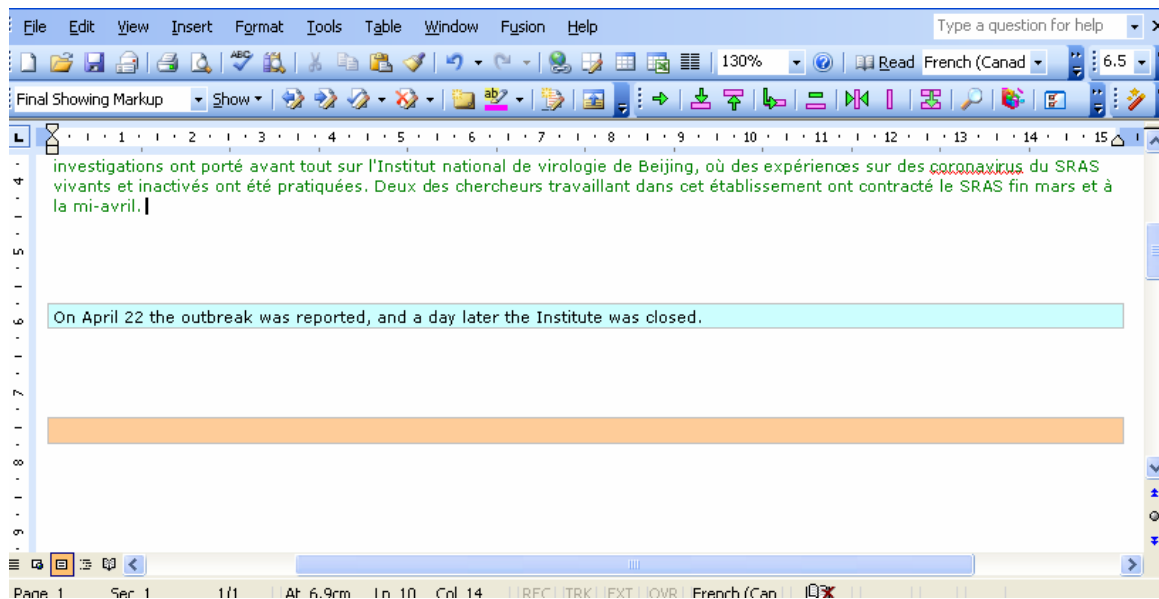
<sup>16</sup> A term match is also called an Expression Match in Fusion Translate™. The middle panes in Figures 1.6 and 1.7 show the source of the information. A term match comes from the dictionary and a sub-segment match comes from the TM database. Each TM system uses its own proprietary terminology, but the descriptions of different types of matches in this thesis are generic.

## 1.4 Working with a TM

There are two ways of working with a TM: *interactive mode* and *pre-translation mode*. These are the two methods that will be compared in the experiment described in Chapter 3. The following section gives an explanation of these two approaches to using a TM system.

### 1.4.1 Interactive Mode

At the present time, the most common method of working with a TM is in interactive mode where a translator opens the new text to be translated within the TM environment and proceeds to translate in a linear fashion (i.e., segment by segment). Figure 1.8 shows how the segment that is being worked on is presented in a highlighted box and the translation can be entered in a second box. To do the translation, the translator can consult the TM database or associated termbase, and accept, modify or refuse proposed matches. As soon as a segment has been translated, it will be linked with its corresponding source text segment and together, the pair will be stored as a TU in the TM database. This TU is now itself available for recycling.

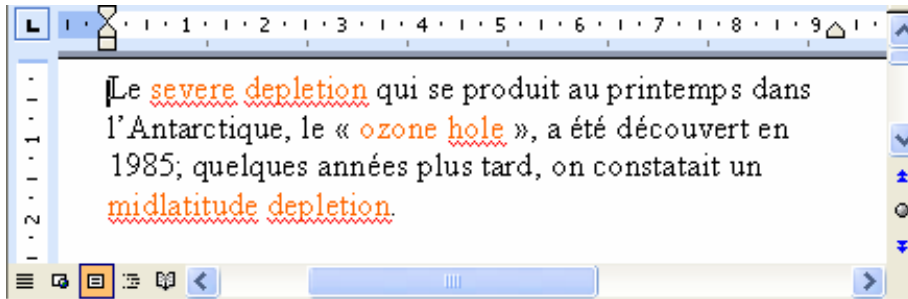


**Figure 1.8**

This figure shows how a text is segmented in interactive translation. The source text segment is shown in the upper box, and the translation of this segment will be entered by the translator into the lower box.

#### 1.4.2 Pre-translation Mode

Up to this point, we have been focusing on the use of TMs in interactive mode, which seems to be the most common way of using a TM. However, there is another way of using a TM – an option called pre-translation, which Heyn (1998:129) explains is the process of partially translating a text by having a TM system automatically replace those elements in the source text with target language equivalents taken from the TM database or termbase, as illustrated in Figure 1.9.



**Figure 1.9**

This figure shows an example of a pre-translated text. The source language term matches found in the termbase were automatically replaced in the target language creating a hybrid French and English text.

Similarly, L'Homme (1999:23) describes systems that perform pre-translation as “traducteurs de vocabulaire” and explains that their function is to produce partially translated texts. Meanwhile, in the words of Schäler (2001), when a TM system is operating in pre-translation mode, it is “little more than a sophisticated search and replace engine.”

Most of the works that have been written about TMs to date do little more than mention that pre-translating a text is a possibility. For example, O'Brien (1998), Bowker (2003b), and even Heyn (1998) and L'Homme (1999), give little more than a passing reference to this option – they do not explore it in any great detail, and they do not consider how it might affect the translation process or product, or the translator. The lack of attention that has been paid to pre-translation in the literature up to this point can most likely be explained by the fact that, until very recently, the vast majority of translators used TMs solely in interactive mode. However, as noted in section 0.2.3, pre-translation is now beginning to gain favour among those clients who do not wish to give translators direct access to their TM database, so it is now imperative that we learn more about the impact that pre-translation can have on issues such as productivity, quality and translator

satisfaction, particularly as compared to the way that these factors are affected by TMs that are operating in interactive mode. To summarize this situation, translators who own TM systems use them to translate different documents for different clients. This means that the TMs that translators are building up and using contain texts from all of the clients they work for. Therefore, texts translated for one client could potentially be recycled and used in a translation for another client. This makes some clients uncomfortable as they do not like to think that work for which they are paying could be used to benefit their competitors. In addition, some clients have very specific terminology that they want to appear in their texts, and they may even consider their terminology to be proprietary. According to Heyn (1998:129), the pre-translation approach is especially popular with translation agencies who contract work out to freelance translators, and for the reasons noted above, an increasing number of other clients are beginning to explore pre-translation as an option.

#### 1.4.2.1 *History of pre-translation*

Although the idea of pre-translation has been popularized in the context of TM, the idea pre-dates the use of TM systems. For instance, Claude Bédard (1990; 1992) proposed the idea of pre-translation in the early 1990s as an alternative to machine translation. Bédard (1992:745) described pre-translation in the following way:

Une prétraduction est une *traduction partielle* du texte, par simple remplacement global d'une partie des mots du texte au moyen d'un dictionnaire d'équivalences.

He suggested that machine translation potentially created more work for the translator in terms of post-editing the output, and he suggested that pre-translation “laisse au traducteur l’entière initiative de la traduction” (Bédard, 1992:738).

As conceived by Bédard (1990:440-441; 1992: 745-746) and illustrated in Table 1.3, pre-translation had a strong emphasis on replacing specialized terms in particular, rather than replacing all general language words or sentence fragments. In addition, there were certain restrictions in place. For instance, lexical items that could potentially create confusion or which had ambiguous part-of-speech categories (i.e., homographs) would be left in the source language for the translator to deal with, as would any lexical items that had multiple possible translations.

|                             |  |
|-----------------------------|--|
| <i>Original source text</i> | This <b>subsection</b> describes how <b>unbalance</b> is detected in the <b>vibration spectrum</b> , explains the different types of <b>unbalance</b> , and outlines the method for correcting <b>unbalance</b> . Methods of calculating the <b>correction masses</b> are described in <b>subsections 5.2, 5.3 and 5.4</b> , among which the <b>vector diagram</b> and the use of <b>programmable calculator</b> and a <b>balancing program</b> .                  |
| <i>Pre-translated text</i>  | This <b>sous-section</b> describes how <b>balourd</b> is detected in the <b>spectre des vibrations</b> , explains the different types of <b>balourd</b> , and outlines the method for correcting <b>balourd</b> . Methods of calculating the <b>masses d'équilibrage</b> are described in <b>sous-sections 5.2, 5.3 and 5.4</b> , among which the <b>diagramme vectoriel</b> and the use of <b>calculatrice à programme</b> and a <b>programme d'équilibrage</b> . |

**Table 1.3**

An example of a pre-translated text as conceived by Bédard (Source: Bédard, 1992:745).

#### 1.4.2.2 *Pre-Translation in the current context of TM*

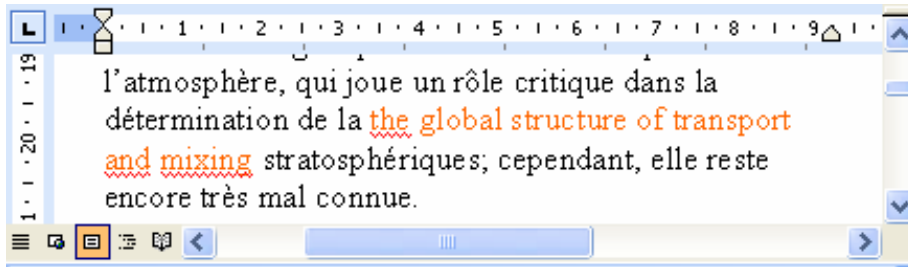
Today, in the context of TM, the notion of pre-translation has evolved somewhat from the

way it was first conceived by Bédard (1990; 1992). As discussed in section 1.3 above, a TM system looks for matches not only in a termbase containing specialized terminology (e.g., *northern hemisphere* or *greenhouse gas*), but also in a TM database, where matches can be exact or fuzzy, or may even consist of sub-segment chunks (e.g., *increased concentrations of greenhouse gases* or *in the Arctic*). L’Homme (1999:222) explains that in pre-translation mode, the system looks for matching segments from the TM database first, and then proceeds to match terms against the entries in the associated termbase or dictionary.<sup>17</sup> Most TM systems that support pre-translation automatically replace any of these types of matches.<sup>18</sup> Therefore, instead of being presented with a text that consists primarily of English-language sentences containing specialized French terms, as suggested by Bédard (1990; 1992) and as shown previously in Figure 1.9, translators may actually be faced with a text that contains partially translated sentences and translations of chunks of text that do not necessarily constitute semantic units, as illustrated in Figure 1.10.

---

<sup>17</sup> L’Homme (1999:225,226) also explains that some pre-translation programs require user interaction where the program stops on a word and asks the user to verify if the equivalent is correct. However, the Fusion Translate TM system, which has been used in the experiment described in chapter 3, does not require this type of user interaction.

<sup>18</sup> Note that fuzzy matches would only be replaced up to a certain threshold value specified by the user. For example, the user might specify that only matches that have a 75% or greater similarity should be replaced (Mossop, 2001:89). In addition, in cases where multiple fuzzy matches are retrieved for a given sentence (see section 1.3.2), the fuzzy match with the highest percentage value will be the one inserted into the new target text.



**Figure 1.10**

This figure shows a pre-translated text containing a sentence fragment.

By providing translators with this type of “pre-translated” text, clients are in a way giving the translator access to the content in the company’s TM system without actually having them use it directly. However, writing more recently, Bédard (2000) has been critical of this new approach, noting that TM databases are made up of many different texts that could have been produced by many different authors and translators. Each text and each translator will have a different style, and when bits and pieces from a variety of texts are forcibly brought together, the resulting text could be a stylistic hodgepodge that Bédard describes as a “salade de phrases” or a “sentence salad”.

Bédard’s “sentence salad” is similar to the notion of “disparity” described by Delisle (2006, in press). According to Delisle, disparities are breaks in style: “*disparity* describes stylistic incoherencies and discordances affecting the translated work. When compared to the original, the translation demonstrates a lack of linguistic, stylistic and tonal unity, among others.” Delisle uses a fashion analogy to compare a disparate translation to someone wearing a mixture of sartorial styles – a top hat, a ruff, a toga and patent-leather shoes. Although Delisle is discussing literary translation in particular, I feel that this concept could be applied to translation in general, and particularly to the case of pre-

translation, where the new target text is generated by automatically pasting bits and pieces from other texts together.

#### *1.4.2.3 Pre-translation and post-editing*

As mentioned briefly in section 0.3, working with a pre-translation bears some resemblance to a process known as post-editing of machine translation. Mossop (2001:89) defines post-editing as “human revision of MT [machine translation] output”. Loffler-Laurian (1996:83) describes it in the following way:

Si l'on donne une traduction brute à « revoir » à quelqu'un en le prévenant qu'il s'agit d'une sortie machine et en lui indiquant quelques consignes de travail, alors on est dans le domaine de la post-édition.

Post-editing of machine translation output is similar to revising pre-translations generated by a TM system in that the translator is given a text in which changes or corrections need to be made in the wake of some type of automatic processing step. In the case of post-editing machine translation output, the translator is given a “raw” machine translation that is considered to be unfinished and he or she must identify problems and make any necessary corrections (Allen, 2003:298). Revising a pre-translation done by a TM system follows the same basic principles, but differs from post-editing machine translation output in that both revision and translation work need to be done. First, the translator must translate anything that was left in the source language, and then he or she must revise what has been pasted into the source text from the TM in order to transform it into an acceptable translation (Mossop, 2001:89). The following examples, taken from

Mossop (2001:89-90) will serve to illustrate the difference between the job of a post-editor (Table 1.4) and that of a translator working with a pre-translation (Table 1.5).

|   |  |
|---|--|
| <i>Source text</i>  | Habilité à écouter et à comprendre afin de recevoir et répondre aux demandes des traducteurs et des clients.   |
| <i>Raw machine translation output</i>   | Skill to listen and include/understand in order to receive and answer at the requests of the translators and the clients.  |
| <i>Changes that a post-editor must make in order to make the translation intelligible</i> | Deletion of the words “at” and “include” would make the text intelligible.<br><br>An even more acceptable text would read “respond to requests from translators and clients”, but these additional changes are not needed to make the text intelligible. |

**Table 1.4**

An example of post-editing machine translation output (Source: Mossop, 2001:90)

|   |   |
|---|---|
| <i>Source text segment</i>  | Habilité à écouter et à comprendre afin de recevoir et répondre aux demandes des traducteurs et des clients.  |
| <i>72% fuzzy match inserted by TM</i>   | Ability to listen and understand in order to receive and respond to requests for information from visitors and telephone callers.   |
| <i>Changes that a translator must implement to make the proposed match translation acceptable</i> | The translator has to treat the expression “for information from visitors and telephone callers” as if it were a mistranslation and replace it with “from translators and clients”. |

**Table 1.5**

An example of editing a pre-translated text produced by a TM system (Source: Mossop, 2001:89-90)

As noted in section 1.4.2.1, when Bédard (1992) first conceived the notion of pre-translation, his idea was that the elements of the ST that would automatically be replaced by TL equivalents would primarily be specialized terms (see Table 1.3). Bédard (1992:745-746) suggests that a translator can open a pre-translated text on his or her

computer screen and use it as a starting point for creating his or her own translation. He is quick to point out that this is not a case of post-editing as such, since for him, a pre-translation is simply a source language text containing specialized target language terms – the actual syntactic structure of the source text is unchanged (Bédard 1990:442). In fact, Bédard (1992:746) coined the term *ptaduction* to refer to the process of editing a pre-translation:

Devant le besoin d'un terme pour désigner le travail de traduction effectué à partir d'un texte pré-traduit, j'ai forgé le mot « ptaduction » (téléscopage entre PTA [pré-traduction automatique] et traduction).

He continues by noting that “Il semble bien que la ptaduction soit plus alléchante pour les traducteurs que la post-édition” because among other things, when dealing with post-editing of machine translation, the translator is faced with a different set of circumstances from one sentence to the next. For example, a successfully translated sentence may be followed by a non-idiomatic sentence, followed by an erroneous translation, etc. In Bédard’s opinion, analyzing and correcting machine translation output is more mentally taxing than pre-translation. It was noted in section 1.4.2.2 that today, the notion of pre-translation has changed from Bédard’s initial conception to include the substitution of sentence fragments which may not even correspond to semantic units. Therefore, translators working with today’s version of pre-translated texts, as illustrated previously in Figure 1.10, are faced with an editing task which, although it is not exactly the same as post-editing machine translation, is certainly more complex and taxing than the *ptaduction* envisaged by Bédard.

## **1.5 Concluding remarks**

This chapter has introduced the concepts of translation memory, interactive translation and pre-translation. This literature review will form the basis for a number of research hypotheses, presented in Chapter 2, which will be investigated by means of an experiment described in Chapter 3.

## Chapter 2 – Research Hypotheses

Keeping in mind the objectives outlined in section 0.1, and based on the knowledge that I have acquired from both my own practical experience of TM systems and the review of the literature on TMs, interactive translation and pre-translation that was presented in Chapter 1, I will now formulate some hypotheses for investigation.

### 2.1 Hypothesis 1: Productivity

It has been noted, for example by Allen (2003:303), that some translators who have spent time post-editing MT output “have abandoned working in such an environment because this [post-editing] can often be as or even more time-consuming than translating a target text from scratch”. In section 1.4.2.3, it was pointed out that, in the initial conception of pre-translation presented by Bédard (1990; 1992), the pre-translated text consisted mainly of a source language text whose clear-cut specialized terms had been replaced by target language equivalents. However, the idea of pre-translation has now evolved so that any parts of the source text that have a match in the TM database – whether this match be at the level of an entire segment (exact or fuzzy), the sub-segment level, or the term level – can be automatically replaced. The replacements do not necessarily correspond to semantic units, and the resulting text is a hybrid document that switches between source and target languages with no underlying logic. Given these circumstances, the task of editing a pre-translated text no longer resembles the task of *ptaduction*, as described by Bédard (1992:746) but is in fact closer to the task of post-editing machine translation,

which, in Bédard's opinion, can be more complex and challenging than simply translating a text.

In pre-translated texts, translators are "bound" to a certain extent by the partially translated fragments contained in the text. They need to work around these fragments and spend time coming up with formulations that fit in with these existing translations. In contrast, translators working with a TM in interactive mode have the benefit of being able to consult the proposed translations that are retrieved from the database, but they can simply use these as inspiration and are still free to formulate their own constructions without feeling trapped by the structure of the proposal presented by the TM system. Therefore, the first hypothesis is as follows:

*Translators working with pre-translated texts will have a lower productivity than translators working with TMs in interactive mode because the former will need to spend more time trying to decipher the hybrid text and finding target language formulations that fit in with those parts of the text that have already been translated.*

## **2.2 Hypothesis 2: Quality**

It was noted above that translators working with pre-translated texts are somewhat bound by those sections of text that have already been translated. This occurs in part because, as noted in section 0.2.2.3, some clients will not pay translators to rework those sections of a

text for which a match has been retrieved from the TM database. In such a case, translators have no incentive to spend a lot of time reworking or retranslating that piece of the text since they would essentially be doing this work for free (or at least for a lower rate). Instead, they spend more time focusing on translating those portions of the text for which no matches were retrieved.

However, as observed by Bédard (2000), the resulting text may suffer from the “sentence salad” effect because the matches that have been automatically inserted into the text during pre-translation may potentially have originated from a range of different texts, all produced by different translators who all have their own individual style. Add to this the style of the translator working on this new text, who is trying to finish the sentences that have been partially translated by someone else, and the result could be a text that is of lower quality than one produced in its entirety by a single translator.

In contrast, a translator working with a TM in interactive mode has more freedom. The same problem exists in that the matches presented could potentially come from different sources and thus contain different styles, but the translator can consult them for inspiration rather than feeling obliged to insert them directly into the new target text.

Another point worth considering is that, as noted in section 1.3.2, TMs can actually retrieve multiple matches for a given segment, and these matches are ranked by the system on the basis of their superficial similarity to the source text segment. However, as noted by Macklovitch and Russell (2000), Bowker (2002b) and Somers (2003a), the

match that receives the highest ranking by the system may not necessarily be the one that is most useful to the translator. In the case of pre-translation, the highest match (i.e., the one with the greatest superficial similarity) is automatically inserted into the target text, and the translator has no opportunity to view the lower-ranking matches, even though these may actually be more helpful (e.g., in terms of semantic content). When working with a TM in interactive mode, translators have the opportunity to consult all the matches that have been retrieved and to decide for themselves which information best meets their needs.

Bearing all this in mind, the second hypothesis is as follows:

*Translators working with pre-translated texts will produce lower-quality translations than translators working with TMs in interactive mode because the former are obliged to work with texts that have been partially translated using a variety of different styles, and they do not have access to all the information contained in the TM database.*

### **2.3 Hypothesis 3: Translator satisfaction**

Some of the points raised in sections 2.1 and 2.2 above may also affect translator satisfaction. For instance, in the case of pre-translation, the fact that translators are constrained by being forced to adapt their work around pre-existing partial translations

may leave them feeling dissatisfied with the resulting quality of the text, or frustrated by the length of time required to analyze and adapt these partial translations.

Moreover, since the use of a TM promotes repetition, and the same terms, phrases and sentences are recycled, translators may feel that their creativity is being stifled when working with a pre-translated text because they have to work within the constraints of the material that has been pre-translated. Nyberg et al. (2003:272) report a similar finding among translators who are required to translate texts written in a controlled language, noting that “Translators ... tend to think of their work in holistic terms, and prefer to produce texts which flow from beginning to end with appropriate stylistic variation.”

Translators using a TM in interactive mode will still be presented with the same repetitive suggestions, but they may feel less obliged to adopt them wholesale into the new translation. Rather, they can consult all the types of matches that have been found, and then accept, modify or reject this information depending on their needs. In other words, they are more able to produce a “text” as compared to the translators working in pre-translation mode, who are reduced to stringing together a series of segments.

Therefore, the final hypothesis is as follows:

*Translators working with pre-translated texts will have a lower level of job satisfaction than translators working with TMs in interactive mode because the former are obliged to adapt their style to that which is already contained in the*

*pre-translated text, which means they will have less control over the creation of a holistic text.*

## **2.4 Concluding remarks**

The three research hypotheses presented in this chapter will now be investigated by means of a pilot study experiment as described in the following chapter.

***PART II:  
EXPERIMENTAL DESIGN AND TESTING***

## **Chapter 3 – Experiment: Preparation and Execution**

In this chapter, I will describe the steps that were taken to both design and carry out a pilot study experiment to test the research hypotheses discussed in Chapter 2.

### **3.1 Preparatory Work**

Preparatory work that needed to be done before the experiment could be carried out included identifying participants, a subject field, texts, and a Translation Memory tool, as well as building a TM database and termbase and generating pre-translations. These will be discussed in detail below.

#### *3.1.1 Identifying Participants*

Two types of participants were required for this study: translators and evaluators. The criteria used to identify each group are outlined in the following sections. It is important to note that because I am working within the scope of an MA thesis, I do not have funds to pay participants. I was therefore forced to rely on volunteers, which limited both the type and number of participants available.

##### *3.1.1.1 Translators*

My first concern was to identify a relatively homogeneous group of translators. The

following criteria were used to identify suitable participants:

- Language direction: French to English
- Translation experience: must have at least a BA in translation (or equivalent)
- Technology experience: must be familiar with TM tools

Using these criteria, I was able to identify four translators to participate in the experiment. Since I am a native speaker of English, the translations were to be done from French into English, thus, all four translators were Anglophones and were able to translate in that language direction. All four had completed the coursework for the MA in Translation at the University of Ottawa and were just about to commence work on their theses or major projects. Finally, three of the four translators had taken the graduate course *TRA 5903: Computers and Translation* in which they were introduced to TM tools and had used the Fusion Translate tool to complete an assignment.<sup>19</sup> The fourth participant had not followed the same course, but had taken a similar course called *TRA 4956: Translation Technologies* at the fourth-year undergraduate level during the qualifying year for the MA in Translation program.<sup>20</sup> As part of this course, the student had been introduced to TM tools, and had completed a practical assignment using the TM tool called Translator's Workbench developed by Trados, which operates according to the same principles as Fusion Translate.

---

<sup>19</sup> The course *TRA 5903 Computers and Translation* is offered as an optional course for the MA in Translation.

<sup>20</sup> The course *TRA 4956 Translation Technologies* is offered as a compulsory course for the BA in Translation and as part of the qualifying year for the MA in Translation program.

### 3.1.1.2 *Evaluators*

To test the quality-related hypothesis outlined in section 2.2, I needed to find evaluators who could competently judge the quality of the translations produced. I was able to find two experienced professors at the University of Ottawa School of Translation and Interpretation, both of whom work in the French-to-English language direction.<sup>21</sup> In addition, both professors have extensive experience grading professional certification exams for the Canadian Translators, Terminologists and Interpreters Council (CTTIC).

### 3.1.2 *Identifying a Domain*

As observed by researchers such as Webb (1998) and Bowker (2002a:112), texts dealing with specialized subjects are best suited for working with a TM because the vocabulary and structures tend to be repeated in texts within a specialized domain. However, since the translators participating in the experiment were translation students, and thus were not specialists in any given area (see section 3.1.1.1), I decided to opt for a domain that has been popularized to a certain extent and which would therefore be at least somewhat familiar to non-specialists. Therefore, the domain that I selected was *ultraviolet (UV) radiation and the ozone layer*. Over the past several years, there have been many reports in the general media about the depletion of the ozone layer, which consequently lets in harmful UV radiation and thus results in a need to limit sun exposure, to wear sunscreen and hats, and to monitor the UV index, etc. I felt that this choice of domain was a good

---

<sup>21</sup> Neither of these professors was involved in the supervision of this thesis and, in order to minimize bias, they were not informed of the specific details of the objectives.

one because it was accessible to non-specialists, but it still contained a relatively restricted vocabulary and should therefore be repetitive enough to get good results from a TM.

In addition, from a purely practical point of view, the popularity of this subject in the general media also meant that it would be relatively easy to find texts to include in the TM database, which will be discussed in more detail in section 3.1.4 below.

### *3.1.3 Selecting a TM Tool*

As noted in section 0.5.1, numerous TM tools are available on the market; however, I chose to work with the Fusion Translate tool because it is a state-of-the-art TM system that offers a pre-translation option, and it is available at the University of Ottawa Writing Centre computer laboratory. Moreover, as discussed in section 3.1.1.1, it was already familiar to most of the participants recruited for the experiment.

### *3.1.4 Building a TM Database and Termbase*

After having selected the domain UV radiation and the ozone layer, I needed to find some texts to incorporate into the TM database. Because a TM database consists of aligned pairs of texts, as explained in section 1.2.2, I needed to find previously translated pairs of texts in English and French. The Environment Canada website proved to be an excellent resource, and I was able to find fifteen text pairs to develop the TM database, which

provided approximately 2700 translation units (TUs) and 65,000 words.<sup>22</sup> Each of the text pairs were imported into the TM database using the *Import* option in Fusion Translate, which performed a preliminary alignment.

#### 3.1.4.1 *Verifying alignment of texts*

Since the text pairs were pre-existing translations, they were automatically aligned using Fusion Translate's automatic alignment feature; however, automatic alignment tools occasionally make mistakes. As pointed out by Bowker (2002a: 95), TM systems sometimes run into difficulty when trying to align cases where a single SL segment has been translated by multiple TL segments, or where multiple SL segments have been combined by a translator into a single TL segment. Additionally, problems can be caused when information is added to or omitted from the target text (e.g., to explain a culture-specific reference, or to delete information not pertinent to the target audience).

Having a well-aligned TM is a pre-requisite for getting useful and reliable information from the database. Therefore, to ensure that the database had been properly aligned, I verified each TU. Any un-aligned segments were aligned manually, and any segments that were misaligned were re-aligned properly. This is a painstaking and labour-intensive process that took approximately 25 hours; however, it is a vital step. If the contents of the TM database are not properly aligned, then the proposals retrieved by the system may be inaccurate and would therefore be of no help to the translator.

---

<sup>22</sup> The Environment Canada website is located at <http://www.ec.gc.ca/>. The texts used for this project came from the "Ozone" section of the site (English: <http://www.ec.gc.ca/ozone/en/index.cfm> and French: <http://www.ec.gc.ca/ozone/FR/index.cfm?>)

#### 3.1.4.2 *Adding to the termbase*

The Fusion Translate tool has a feature that will analyze the contents of the TM database and will automatically extract sub-segment chunks (see section 1.3.3) that appear more than once in the SL texts that have been stored in the database. The tool also tries to identify potential translations for these chunks from the corresponding TL texts.<sup>23</sup> However, as was the case with automatic alignment, the automatic identification of translation equivalents is not always perfect. Therefore, I manually verified or corrected all of the proposals before transferring them to Fusion Translate’s “dictionary” file (i.e., termbase) – a process that took approximately 3 hours. Once stored in the termbase, this information is available to be consulted during interactive translation or to be automatically inserted into a source text in pre-translation mode.

#### 3.1.5 *Choosing Source Texts*

After building the TM database, the next step was to choose two source texts to be used in the pilot study experiment. To get the best results from the TM, the texts clearly had to be on the same subject and of the same text type as those found in the database. In addition, since the translators and evaluators participating in the experiment were acting in a volunteer capacity, it was necessary to find texts that were relatively short and would not require an unreasonable investment of time on the part of the volunteers. Therefore, I selected another French-language text on UV radiation and the ozone layer from the Environment Canada website that can be categorized as a vulgarized scientific report.

---

<sup>23</sup> To identify potential equivalents, Fusion Translate uses a proprietary statistics-based algorithm.

From this text I took two self-contained extracts of approximately 350 words each to use as source texts. Since both source texts were taken from one larger text, I felt reasonably comfortable that the two extracts were of a comparable level of difficulty.

These texts were “pre-tested” in order to make sure that they generated not only a reasonable number of matches from the TM database and dictionaries, but also that they both generated approximately the same number of matches. This is important because of the comparative nature of the experiment. The source texts used in the pilot study experiment can be viewed in Appendix A.

### *3.1.6 Producing Pre-translations*

Because the aim of the experiment is to compare the results of translators working in interactive translation mode against those of translators working in pre-translation mode, it was necessary to generate some pre-translated texts. Using the *Translate* option in Fusion Translate, the source texts to be translated were uploaded and any matching segments from the TM database and matching terms from the TM dictionary were inserted into the source text by means of the *copy source to target* command. The resulting pre-translated texts were so-called hybrid texts, containing some words in the SL and others in the TL, as described in section 1.4.2. These pre-translated texts can be seen in Appendix B.

### *3.1.7 Preparing Instructions and Questionnaire for Translators*

Instructions for translators participating in the experiment were drawn up to ensure that each translator was following the same procedure. Similarly, time sheets (where participants would keep track of the length of time required to produce their translations), and resources sheets (where participants would keep track of which resources they consulted when producing their translations) were prepared.<sup>24</sup> Lastly, a questionnaire was developed in order to get feedback from the translators about their experience translating texts in both interactive and pre-translation mode. All of this material (i.e., instructions, time sheets, resources sheets, and questionnaire) can be found in Appendix D.

### *3.1.8 Preparing Evaluation Criteria for Evaluators*

Just as it was important to ensure that all the translators followed the same procedure, it was equally important to make sure that the two evaluators followed the same procedure. This meant that it was necessary to draw up evaluation criteria for them to follow. This proved to be quite a challenge. As part of my translation theory course in the MA program, I had the opportunity to research and write a paper on translation evaluation, and one of the main things that I learned is that there is no consensus on what makes a “good” translation, or what criteria are the most important. Much of the reading that I did

---

<sup>24</sup> Note that in order to control the variables in the experiment, it was important that each translator had access to the same resources. Therefore, I provided a list of resources – in addition to the TM database or pre-translation – that the translators were allowed to consult if desired, and I asked them to indicate which of these, if any, they consulted for each translation. This list of resources is part of the instruction sheet contained in Appendix D.

for the translation theory course was theoretical in nature (e.g., House, 1977, 1997; Williams, 2004) and not really able to be directly applied to the development of practical guidelines for the purpose of this pilot study. Instead, I took as my basis the guidelines issued by the Canadian Translators, Terminologists and Interpreters Council (CTTIC) to the evaluators who assess the professional certification examinations.<sup>25</sup> I modified these guidelines very slightly by reducing the penalties attributed to various types of errors. I did this because the translators participating in this study are not yet experienced enough to sit the CTTIC exam, so I felt that they should not necessarily be judged by those very exacting standards. In any case, since this experiment is comparative in nature (i.e., translations produced in interactive mode will be compared against those produced in pre-translation mode), the most important thing is that the same criteria be applied by both evaluators to all the translations. The guidelines that were given to the evaluators participating in the pilot study experiment can be seen in Appendix H.

### *3.1.9 Training Translators on the Use of Fusion Translate*

As mentioned in section 3.1.1.1, the translators participating in this experiment all had previous experience working with TM software, and most had even used Fusion Translate before. Nevertheless, I decided that it would be a good idea to give them all a quick “refresher”. I therefore prepared a tutorial, along with a source text and a sample TM database and termbase in the field of obesity.<sup>26</sup> The tutorial, which is provided in

---

<sup>25</sup> For more information on CTTIC, consult their web site: <http://www.cttic.org/>

<sup>26</sup> As was the case with the domain of UV radiation and the ozone layer which was to be used for the pilot study experiment (see section 3.1.2), the field of obesity was chosen as the subject of the tutorial because it

Appendix C, took the translators through all the major steps required to work with a TM in interactive mode. The translators were given the tutorial approximately ten days prior to the experiment, and they were encouraged to go through the tutorial as many times as necessary in order to feel comfortable with the tool. All of the translators assured me that they had completed the tutorial at least once before doing the experiment.

### *3.1.10 Pre-Testing and Refining the Pilot Study*

As a final step before proceeding to the actual pilot study experiment, I wanted to be sure that all the instructions, guidelines, etc. that I had prepared would be clear for the participants. Therefore, I asked a classmate who was not to be involved in the pilot study to carry out a mock session of the experiment using a copy of the TM and all documentation discussed above. Based on her feedback, I clarified some of the instructions for the participants and refined some of the questions on the questionnaire, but it did not prove necessary to make any major changes.

With regard to the evaluation criteria, these were examined by my thesis supervisor, who is also a certified French-to-English translator and an experienced translator trainer. Based on her feedback, I made some small refinements, but again, no major changes were made at this point.

---

was accessible to non-experts but still contained specialized terms and structures conducive to working with a TM tool.

## 3.2 Experiment

The pilot study experiment had two main components. Firstly, the translators produced their translations and answered their questionnaires, and secondly, the evaluators assessed the quality of the translations produced.

### 3.2.1 Translation

The experiment took place on Friday April 7, 2006 at 3pm in the Writing Centre computer laboratory at the University of Ottawa. I had already loaded the previously prepared TM database and termbase (see section 3.1.4) onto the computers.

The four translators were each assigned a code name – Blue, Red, Green and Yellow – so that the results of the experiment would remain anonymous. The translators were asked to translate two texts each – ST1 and ST2 – in the following order:

| <b>Translator</b> | <b>Method and text to translate first</b> | <b>Method and text to translate second</b> |
|-------------------|---|--|
| <i>Blue</i>       | ST2 using interactive translation         | ST1 using pre-translation                  |
| <i>Red</i>        | ST2 using interactive translation         | ST1 using pre-translation                  |
| <i>Green</i>      | ST1 using interactive translation         | ST2 using pre-translation                  |
| <i>Yellow</i>     | ST1 using interactive translation         | ST2 using pre-translation                  |

**Table 3.1**

This table shows the method of translation for each individual translator for ST1 and ST2

Although I tried to ensure that the two STs were highly similar (see section 3.1.5), it is always possible that one might be considered “easier” than the other, or that one might be more amenable to interactive translation than the other, etc. By alternating which ST was used with each method, I hoped to reduce any bias in this regard. While it is clear that the small number of participants will seriously limit the scientific validity of the results, I at least wanted to design a methodologically sound experiment which could perhaps be scaled up at a later stage.

As the translators were working on their translations, they were asked to record which additional resources they consulted in order to complete their translations, and they were also instructed to keep an accurate record of the amount of time required to complete each translation. The instructions sheet they were given at the beginning of the experiment reminded the translators to record their start and finish times as well as the resources used for each translation.

Once the translators had completed both translations, they were asked to fill in a questionnaire based on their experience working in interactive mode and pre-translation mode.<sup>27</sup>

The resulting translations, time and resources sheets, and questionnaires were then gathered for analysis.

---

<sup>27</sup> See Appendix D for instructions, time sheets, resources sheets and questionnaire.

### *3.2.2 Evaluation of Translations*

As noted in section 3.1.1.2, I located two experienced evaluators who were willing to assess the translations produced during the pilot study. In order to minimize bias, the evaluators were not told which texts resulted from translations done in interactive mode and which resulted from translations done in pre-translation mode. Instead, each evaluator was simply given eight translations (four translations of ST1 and four translations of ST2) and asked to evaluate each text by applying the evaluation guidelines that were provided (see section 3.1.7 and Appendix H).

### **3.3 Concluding remarks**

In this chapter, I have described in detail the steps taken to design and execute a pilot study experiment intended to determine whether interactive translation or pre-translation is the better choice for translators (and their clients) who want to have higher productivity, better quality translations, and greater job satisfaction. The results obtained from this experiment will be analyzed and discussed in Chapter 4.

## Chapter 4 - Data Analysis

In this chapter, I will present, analyze and discuss the data gathered from the experiment outlined in Chapter 3. The analysis will be structured according to the hypotheses put forth in Chapter 2, namely productivity, quality and translator satisfaction.

### 4.1 Productivity

In section 2.1, it was hypothesized that translators working in pre-translation mode would take longer to produce a translation than would translators working in interactive mode.

#### 4.1.1 Data

As mentioned in section 3.1.7, each of the translators was asked to record their start and finish time for each translation on the Time Sheets given to them prior to commencing the experiment (see Appendix E for Time Sheets). Table 4.1 shows the times each translator took for each translation method.

| Translators   | Interactive Translation | Pre-Translation |
|---------------|-------------------------|-----------------|
| <i>Blue</i>   | 49 min (ST2)            | 49 min (ST1)    |
| <i>Red</i>    | 64 min (ST2)            | 55 min (ST1)    |
| <i>Green</i>  | 51 min (ST1)            | 60 min (ST2)    |
| <i>Yellow</i> | 46 min (ST1)            | 46 min (ST2)    |

**Table 4.1**

This table shows how much time each translator required for each method of translation.

#### *4.1.2 Discussion*

As shown in Table 4.1, the time required for both Blue and Yellow to translate a text in interactive mode was exactly the same amount as that required to translate a text in pre-translation mode.

The remaining two translators produced contrasting results. In the case of Red, it took this translator nine minutes longer to translate a text in interactive mode than it took to translate a text in pre-translation mode. However, the case of Green was exactly the opposite. It took Green nine minutes more to translate a text in pre-translation mode than it took to translate a text in interactive mode.

Given that two of the four translators took precisely the same amount of time for both modes, and that one required more time for interactive translation and the other for pre-translation, it would seem that the mode of translation does not have any significant impact on translator productivity. Therefore, based on the data used in this experiment, it would seem that the hypothesis presented in section 2.1 is not supported. However, it may be worth considering some other factors that might have had an influence on the results.

##### *4.1.2.1 Texts*

Although every effort was made to select comparable texts for the experiment, as described in section 3.1.5, it is still possible that one text may have been more difficult

than the other and thus taken longer to translate regardless of the translation method used. One translator – Green – did comment in the questionnaire that she/he found ST2 to be more difficult than ST1; however, the other three translators did not make any comments to this effect. In keeping with her/his opinion that ST2 was more difficult, Green also took longer with the translation of this text, which in her/his case was the one done in pre-translation mode.

Since Green is the only one who found ST2 to be more difficult, it is possible that Green is simply an outlier in this respect.<sup>28</sup> For instance, consider the fact that both Blue and Yellow took precisely the same amount of time for their two translations (49 minutes for Blue to translate each text, and 46 minutes for Yellow). This would seem to argue against a general difference in the difficulty level of the text since Blue translated ST2 in interactive mode and ST1 in pre-translation mode, while Yellow translated ST1 in interactive mode and ST2 in pre-translation mode. If one text had been significantly more difficult than the other, we might expect to see the time required to translate that text being longer, regardless of the mode. Instead, we see that both Blue and Yellow completed the two translations in exactly the same amount of time using opposite modes for each text.

---

<sup>28</sup> An outlier is a data point that is located far from the rest of the data. It is common for some results to deviate from the norm when authentic data is involved; nevertheless, outlier data has a distorting effect on the results of a small-scale test such as the one reported here. This underlines the need for testing to be carried out on larger scale, where the effects of such outlier data may not be felt as acutely (see section 5.2.1).

#### 4.1.2.2 *Technical difficulties*

Another issue that may have affected the time required to complete the translations was raised by Red in the questionnaires filled out at the time of the experiment (see Appendix G for Completed Questionnaires).

Red took longer to complete the translation in interactive mode (64 minutes) than to complete the translation in pre-translation mode (55 minutes); a finding that did not support the hypothesis that interactive translation would be faster. However, in the questionnaire, Red stated that she/he was having technical problems with the Fusion Translate software when working in interactive mode. The issue was that the ‘copy to Word’ function was not working properly, meaning that the Fusion Translate Companion would not send selected terms to the Word document where the translation was being done. Thus, the time required for this translator to complete the translation using interactive mode may have been adversely affected by the fact that the translator was spending time attempting to correct the technical problem. Furthermore, had the ‘copy to Word’ command worked properly, it could have sped up the translation process because the translator would have saved time spent typing the words that could have been sent to Word automatically. Therefore, in the case of Red, it is possible that the time required for interactive translation could have been somewhat lower if technical difficulties had not been encountered. This would bring Red closer in line with Blue and Yellow, who both took the same amount of time for both translation methods.

#### 4.1.2.3 *Familiarity with software*

In addition to the specific technical difficulties encountered by Red, the general level of familiarity that all the translators had with the software could have been a factor. Although the participants who acted as translators for this experiment had worked with TMs in class and were given an additional tutorial on Fusion Translate in particular, they still did not have a significant amount of experience in this regard. Blue, for example, stated in the questionnaire that when working interactively, “it took me a while to find English equivalent, as it was not highlighted”, meanwhile, Green observed “I think the TM would have been more efficient if I had had more practice using it”.

Working with a TM in interactive mode requires a greater degree of technological background knowledge and experience than working in pre-translation mode. Therefore, translating a text in interactive mode may prove to be more time consuming for translators in the beginning, than translating a pre-translated text. It may be that translators will only see an increase in productivity once they become more experienced with the tool, learning the shortcuts and tricks that can be used, knowing how to avoid or troubleshoot technical problems, and simply having a higher comfort level with the tools. This is in keeping with an observation made by Bowker (2002a:115) that there is a high learning curve associated with sophisticated tools such as TMs, and as such “translators may actually see an initial drop in productivity during this learning phase, although productivity will begin to increase as they become habituated to the software.”

#### 4.1.2.4 *More results to evaluate in interactive mode*

A related point is that translators working in interactive mode may have been initially overwhelmed by the number of options presented to them. As pointed out by Brockmann (1999:9), and as discussed in section 1.3.2, when working in interactive mode, the TM displays many different suggestions for matches. Although these matches are ranked by the system, Bowker (1999) observes that when translators are inexperienced users, they sometimes “go overboard” investigating all the possible options contained in the database, when really it would have been enough to stop after one or two; however, as they become more experienced and confident, users are able to make decisions without needing to verify every small detail in the database.<sup>29</sup> This situation does not present itself in pre-translation mode, however, because the translator is provided with only one match (i.e., the one rated by the system as the “best” match). Given the translators’ limited experience with this type of software as mentioned above, the time spent looking through all potential matches may have adversely affected the overall productivity of translators working in interactive mode.

#### 4.1.2.5 *Quality of the finished translations*

Another factor that may have influenced the time taken by the translators was the quality of the finished product. The translators were asked to produce translations of a quality that they were satisfied with, but based on the comments made in the questionnaire, it

---

<sup>29</sup> Although Bowker is not discussing TMs specifically, she is discussing corpus-based resources. Since a TM is a type of corpus-based resource, I feel this point can be applied equally well to TMs.

seems that at least one translator – Green – did not necessarily do this. Green specifically stated that she/he should have taken more time to revise the translation done in pre-translation mode. She/he explained that taking more time to re-word some sentences and revise the active and passive voice used in certain sentences would have made the overall translation more consistent. Had this translator taken the necessary time to revise the document, the time required to complete a translation in pre-translation mode would have increased. Since Green took 51 minutes for interactive translation and 60 for pre-translation, this would have meant an even greater margin of difference between the two, suggesting that, for Green at least, working in interactive translation mode is considerably faster than working in pre-translation mode. However, we should also take into account the findings reported in section 4.1.2.1 above, where it was noted that Green found ST2 – the text translated in pre-translation mode – to be more difficult, which could explain why it took longer for her/him to translate this text.

#### *4.1.3 General Conclusions Regarding Productivity*

Since two of the translators recorded precisely the same time for their pre-translation and interactive translation tasks, and since only one of the four translators showed any improvement in productivity using interactive mode, while the remaining translator showed a drop in productivity using interactive mode, there is no firm support for the hypothesis that interactive mode is faster than pre-translation mode. Therefore, the general conclusion based on this data is that the translation method – whether interactive

translation or pre-translation – has no significant impact on the productivity of a translator.

## **4.2 Quality**

In terms of quality, the hypothesis presented in section 2.2 stated that translators working in interactive mode would produce higher quality work than translators working in pre-translation mode.

### *4.2.1 Data*

As outlined in section 3.1.1.2, the quality of the translations produced during the experiment was evaluated by two experienced professors from the University of Ottawa School of Translation and Interpretation who also act as examiners for the CTTIC certification exams. Table 4.2 shows the scores (expressed as a percentage) given by the evaluators to each of the translations, noting which translation method was used.

|                    | <b>Interactive translation (IT)</b> | <b>Pre-translation (PT)</b> | <b>Method resulting in higher quality</b> | <b>Margin of improvement</b> |
|--------------------|-------------------------------------|-----------------------------|---|------------------------------|
| <b>Blue</b>        |                                     |                             |   |                              |
| <i>Evaluator 1</i> | 0%                                  | 0%                          | No difference                             | No difference                |
| <i>Evaluator 2</i> | 31%                                 | 27%                         | IT  | Insignificant (4%)           |
| <b>Red</b>         |                                     |                             |   |                              |
| <i>Evaluator 1</i> | 35%                                 | 27%                         | IT  | Insignificant (8%)           |
| <i>Evaluator 2</i> | 78%                                 | 25%                         | IT  | Significant (53%)            |
| <b>Green</b>       |                                     |                             |   |                              |
| <i>Evaluator 1</i> | 57%                                 | 60%                         | PT  | Insignificant (3%)           |
| <i>Evaluator 2</i> | 45%                                 | 73%                         | PT  | Significant (28%)            |
| <b>Yellow</b>      |                                     |                             |   |                              |
| <i>Evaluator 1</i> | 21%                                 | 14%                         | IT  | Insignificant (7%)           |
| <i>Evaluator 2</i> | 65%                                 | 40%                         | IT  | Significant (25%)            |

**Table 4.2**  
Results of translation evaluations

The first thing that I observed when looking at the results of the evaluations was that the majority of scores assigned by the evaluators were extremely low. At first, I was worried about this; however, it is important to keep in mind the purpose of this experiment. The aim is not to see which of the participants is the best translator, or to determine if these translators are ready for the job market. Rather, the goal is to compare the results of interactive translation to those of pre-translation. Accordingly, the analysis of the data in this section is based primarily on a comparison between interactive mode and pre-translation mode *for each individual translator* and is not meant to compare the particular scores produced by each translator against those produced by any other translator.

Nevertheless, I will return to the issue of the low scores assigned, and possible reasons for this, in section 5.1.2.

#### 4.2.2 *Discussion*

In the case of Blue, Evaluator 1 gave both the interactive translation and the pre-translation a score of 0, so it is not possible to compare these two methods. Evaluator 2 awarded a slightly higher score to the interactive translation, but it was not enough to be significant.

Red received a higher score from both evaluators for the interactive translation. In the case of Evaluator 1, it was only slightly higher, but Evaluator 2 gave the interactive translation a significantly higher score than the pre-translation (53% higher).

The opposite trend was observed for Green, where both evaluators gave a higher score to the texts produced using pre-translation. Again, the score assigned by Evaluator 1 was only slightly higher, whereas that given by Evaluator 2 was considerably higher (28%).

The case of Yellow closely resembles that of Red, where the interactive translations received higher scores from the two evaluators. As was the case with Red, Evaluator 1 gave only a slightly higher score, whereas Evaluator 2 gave a significantly higher score (25%).

If we take a more global look at the results, we can see that slightly more than half the time (5 times out of 8), the evaluators awarded higher scores to the texts produced using interactive translation than to those produced in pre-translation mode. However, on 3 out of these 5 occasions, the results were only slightly better. Meanwhile, pre-translation produced a higher quality result on only 2 out of 8 occasions, one of which was only slightly better. In the final case, it was not possible to determine whether interactive or pre-translation resulted in a better translation since both texts received a score of 0 from the evaluator.

Now I will focus on some other factors that may have influenced the results with regard to translation quality.

#### *4.2.2.1 Perceived difficulty of texts and quality of the finished translations*

Once again, it seems that Green is somewhat of an outlier in that she/he was the only translator to perform better using pre-translation – a judgment that was made by both evaluators. This is particularly interesting in light of Green’s comments on the questionnaire. In the case of Green, ST2 was the text that was translated in pre-translation mode, yet Green commented that she/he found ST2 more difficult than ST1 and that she/he “would have been more satisfied w/ Text 2 if I had taken more time at the end to re-word, change sentences from passive to active, etc.” In spite of the fact that Green found the text used for pre-translation to be more difficult, and that she/he felt that the translation was not as good as it could have been, this was the text that received higher

scores from both evaluators. The results of Green therefore seem to be anomalous. As noted above, it is perfectly normal to encounter outliers when dealing with authentic data, but unfortunately, this really distorts the results of a small study such as the one reported here.

#### *4.2.2.2 Amenability of one text for interactive translation*

Since two different texts were used for this experiment, it is possible that one was more amenable to interactive translation than the other. However, looking at the two cases where interactive translation produced a considerably higher quality text than pre-translation mode, we can see that one of these texts was produced by Red and the other by Yellow. It is important to note here that Red translated ST2 in interactive mode, while Yellow translated ST1. This would seem to indicate that it was not simply a case of one text being more conducive to interactive mode, but rather that the method of working in interactive mode can potentially contribute to the production of a higher quality text.

#### *4.2.3 General Conclusions Regarding Quality*

In conclusion, the data collected for this experiment provide only weak support for the hypothesis that working in interactive mode helps translators to produce higher quality texts than they would when working in pre-translation mode. Slightly more than half of the texts (5 out of 8 or 62.5%) that were judged by the evaluators to be of a better quality were produced using interactive mode, though only 2 out of these 5 were rated as being

significantly better, while the other 3 were considered to be of only slightly higher quality.

### **4.3 Translator Satisfaction**

With regard to translator satisfaction, the hypothesis put forth in section 2.3 was that translators working interactively with a TM system will have a higher level of job satisfaction than translators working with pre-translated texts.

#### *4.3.1 Discussion*

As part of the experiment described in Chapter 3, each of the translators completed a post-translation questionnaire which asked them questions about both methods of translation using a TM and their feelings towards each method as well as their feelings towards TMs in general (see Appendix G). Based on the completed questionnaires, I will outline the prominent points raised with regard to the translators' general satisfaction with both pre-translation and interactive translation.

##### *4.3.1.1 Pre-translation mode*

At least once in each of the four completed questionnaires, the translators noted that pre-translation mode seemed time consuming and not as efficient as interactive mode. Specific comments included statements such as:

- “*Pleasantly surprised @ TM = Biggest reason: Time effective.*” (Blue)
- “*Quick answers [in interactive mode]*” (Red)
- “*I found it time-consuming to check the pre-translated text.*” (Green)
- “*The first one (w/ TM) I preferred. I felt like I translated faster, & w/ less errors.*” (Yellow)
- “*I think the TM made me more productive.*” (Yellow)

It is very interesting to note that while working in pre-translation mode, none of the translators used the pre-translated text as the starting point for their target text. Instead, they created a brand new target text and simply used the pre-translated text as a reference. This led to the complaint by several of the translators that having three documents on the screen (i.e., source text, pre-translated text and target text) was too many and involved too much switching back and forth because all of the documents could not be seen on the screen at the same time.<sup>30</sup> Similarly, one translator stated that it took too much time to check the pre-translated text for a term instead of just looking up the term in a dictionary or online, and most stated that they did not use the pre-translated text as their principle resource, preferring to first consult other resources, such as *Termium* or *Le grand dictionnaire terminologique*.

Another element that two out of the four translators – Red and Green – found unpleasant about the pre-translation mode was that it resulted in an inconsistent style such as switching back and forth between active and passive voice.

---

<sup>30</sup> It should be noted here that for the pre-translation mode the translators were provided with the original ST as well as the pre-translated text.

#### 4.3.1.2 *Interactive mode*

The main complaint that the translators had when working in interactive mode was that there was not enough screen space. For example, when another window is open (e.g., a termbase window) it covers up the Fusion Translate Companion window, so it is not always possible to view the matches that are being proposed. One translator (Green) admitted that when she/he could not see the Companion, she/he forgot about it and had to remind her-/himself to go back to it. Otherwise, the translators enjoyed working interactively with the TM and three out of the four stated that it seemed faster and more efficient and they felt they had better productivity working interactively with the TM. When working in interactive mode, all of the translators used the TM as their principle resource and only had to resort to other resources if the TM did not give them what they were looking for.

#### 4.3.3 *General Conclusions Regarding Translator Satisfaction*

Overall, the general consensus was that the interactive method of working with a TM is superior to working in pre-translation mode. Two of the four translators claimed to be equally satisfied/dissatisfied with the results of both methods of translation, while the other two professed to be more satisfied with the translation that resulted from interactive mode, stating:

- *“I was not as confident w/ the pre-translated text as I was w/ the TM text.”*  
(Yellow)
- *“I think the TM [...] helped me translate w/ better quality”* (Yellow)

- *“I think TM [translation] might be a bit more accurate.”* (Blue)

Moreover, all of the translators stated that they preferred working with the TM to working without it, and they all agreed that if asked by a client to work with a TM they would choose to work in interactive mode over pre-translation mode if given the option.

Therefore, the data gathered from the questionnaires filled out by the translators provide relatively strong support for the hypothesis that translators working in interactive mode would have a greater degree of job satisfaction than translators working in pre-translation mode.

#### **4.4 Concluding remarks**

In this chapter I have analyzed the data gathered from the results of the experiment to determine whether or not they provide support for the hypotheses set forth in Chapter 2. To summarize what has been determined by this particular experiment, firstly, in contrast to the hypothesis put forward regarding productivity, the method of translation used does not seem to have any notable impact on translator productivity. Secondly, with regard to quality, interactive translation seems to result in only a slight improvement in the quality of the translation. Finally, in terms of job satisfaction, translators seem to be considerably happier working in interactive mode, which supports the hypothesis put forward in this regard. In the following Chapter, I will proceed to evaluate the

experimental process in general, and offer some suggestions for future research in this field.

***PART III:  
EVALUATION AND CONCLUSION***

## **Chapter 5 – Concluding Remarks**

As stated in section 0.1, the main objective of this thesis was to investigate whether the two modes of working with a TM – interactive mode and pre-translation mode – had a different impact on translator productivity, translation quality and translator satisfaction. This investigation was carried out by first designing and then pilot testing a methodology that could be used to determine which of two modes seems preferable with regard to productivity, quality and translator satisfaction. Now that this methodology and experiment have been implemented, some general comments on the project can be made. The following section is an evaluation of the design and testing of the methodology, followed by suggestions for future research in this area, and finally, some general concluding remarks.

### **5.1 General Evaluation**

Overall, I feel that the methodology used to complete the pilot study for the purpose of this thesis was largely successful. I was able to test my hypotheses and, even though the data set was quite small, I could draw some very preliminary conclusions. However, if I were to do the experiment a second time, there are certain aspects of the process that could be improved.

### *5.1.1 Database Size*

As noted in section 3.1.4.1, I invested a considerable amount of time – approximately 25 hours – assembling a bilingual corpus that would make up my TM database in preparation for the experiment, as well as approximately 3 hours building up the termbase. I feel that this was a reasonable amount of time to invest in database and termbase building for an MA thesis, but there is no doubt that a larger TM database and termbase would likely have generated a greater number of matches. As pointed out by Bowker (2002a:116), the larger the database and termbase resources, the greater the potential for generating useful matches.

In particular, having a larger TM database for this project would have been especially useful in terms of pre-translation because the pre-translated texts that I produced contained mainly term replacements, with only a few replacements of “chunks of language” (see Appendix J). In this sense, our pre-translated texts were more reminiscent of those proposed by Bédard (1990; 1992), as discussed in section 1.4.2.1, rather than fully representing what the modern-day pre-translation option can produce. Because the pre-translated texts used in the experiment were not as ‘hybrid’ as they could have been, this may have affected the outcome of the experiment. For example, it is possible that a text with a greater degree of hybridity might have made pre-translation either more or less attractive to the translators, or it may have been more or less time consuming to process.

### 5.1.2 *Nature of Participants*

Regarding the participants, it would have been more desirable to recruit professional translators, rather than student translators, for a number of reasons. Firstly, as mentioned in section 3.1.1.1, the students had only a limited (and somewhat artificial) experience using the TM tool in their translation technology classes. A group of professional translators who had a real working knowledge of the tool may have produced different results.

Secondly, it is possible that some of the student participants may not have taken the task very seriously. Although they were asked to produce the translations to the best of their ability, the low marks awarded to the work by the evaluators (including two scores of 0), along with some comments on the questionnaires (e.g., Green's statement that it would have improved the quality of the text if she/he had spent more time revising) calls into question whether or not the student translators really gave their best effort. This is not a criticism of the student participants – after all, they were acting in a volunteer capacity. However, if a group of professionals were paid to participate in this type of activity, the results might have been different. Unfortunately, as an MA student with limited resources, it was not possible for me to pay professionals, so I was restricted to working with students, but I recognize that this was not ideal.

Finally, in hindsight I can see that it would have been preferable to limit the amount of information given to the translators participating in the experiment. For this pilot study, the translators were aware that the aim of the experiment was to compare the two

methods of translation. Had the translators completed the experiment not knowing exactly what it was for, there may have been a different outcome.

### *5.1.3 Instructions for Translators*

In section 4.3.1.1 it was noted that none of the translators used the pre-translated text as the starting point for their translation. Rather, they created a new target text from scratch, and only used the pre-translated text as a reference. In retrospect, I feel that this decision was possibly directly influenced by the nature of the instructions given in Part 2 (steps 3 and 4) on the Instruction Sheet (see Appendix D), which prompted them to open a new Word document for their translation. If the translations had been done using the pre-translated document as a base, it would have been easier to determine whether or not the pre-translated segments had an influence on the syntax of the target text, and it may have changed the way the translators felt about pre-translation mode in general.

### *5.1.4 Evaluation Phase*

As mentioned in section 3.1.10, a translator who was not part of the actual experiment did some pre-testing of the experiment beforehand. This was done to ensure that the experiment would run smoothly and that the instructions for the translators and all of the material involved in the actual pilot study were clear. This same type of pre-testing was not carried out for the evaluation phase of the experiment, but I can see that it would have been beneficial to do so. The fact that the two evaluators assigned very different marks to

some of the translations – in some cases differing by more than 40% – could be a sign that the instructions and/or the evaluation grid were not as clear as they could have been. To conduct an effective comparative evaluation, there can be no room for subjective application of the evaluation grid. Rather, it is necessary for the evaluators to apply the evaluation grid in precisely the same way. The difference in the scores assigned by the evaluators would seem to indicate that this did not happen. Of course, it is possible that one of the evaluators may have been an outlier; however, to know this for certain, it would be necessary to repeat the experiment with different and/or a much larger group of evaluators.

In addition, given that I was working with student translators rather than professionals, it might have been a good idea to create a less punitive evaluation grid.<sup>31</sup> Although I did make some effort to lessen the expectations that were present in the original CTTIC guidelines, as described in section 3.1.8, it may have been wise to go even further in this regard. For instance, in the case of Blue, Evaluator 1 assigned a score of 0 to both texts, which meant that it was impossible to conduct a comparative evaluation of the effects of interactive translation versus pre-translation on the work of this translator. Unfortunately, this resulted in a further reduction of the already small pool of data available for analysis.

Lastly, as mentioned in section 3.1.7, a questionnaire was developed for the translators in order to get feedback about the two different methods of translation. In retrospect, it would have also been useful to have had the evaluators fill out a questionnaire to get their

---

<sup>31</sup> Of course, if professional translators are used in future experiments, a more stringent evaluation grid could be used.

views on the translations in general as well as any major problems that they may have noticed. Of course, the evaluators were free to leave any comments on the evaluation grids, however, a detailed questionnaire would have allowed for a more detailed analysis of the data gathered. A questionnaire may also have helped to explain the marking discrepancies between the two evaluators.

#### *5.1.5 Analysis of Source Texts*

As mentioned in section 3.1.5, the source texts used for the pilot study were both taken from one larger text so that it was certain that these texts were at the same level of difficulty. However, no true text analysis was done on these texts. One translator, Green, mentioned in the questionnaire that she/he found ST2 more difficult than ST1. Had an analysis of the source texts been done, and Green had said that she/he found one more difficult than the other, it might have shown that what Green found more difficult was in fact the method of translation, and not the source text used.

#### *5.1.6 Order of Operation*

A final modification that could be useful to make in future experiments concerns the order in which the operations are carried out. As indicated previously in Table 3.1, which I have reproduced below for the convenience of readers, all translators worked in interactive mode first and in pre-translation mode second.

| <b>Translator</b> | <b>Method and text to translate first</b> | <b>Method and text to translate second</b> |
|-------------------|---|--|
| <i>Blue</i>       | ST2 using interactive translation         | ST1 using pre-translation                  |
| <i>Red</i>        | ST2 using interactive translation         | ST1 using pre-translation                  |
| <i>Green</i>      | ST1 using interactive translation         | ST2 using pre-translation                  |
| <i>Yellow</i>     | ST1 using interactive translation         | ST2 using pre-translation                  |

**Table 3.1**

This table shows the method of translation for each individual translator for ST1 and ST2

To ensure that the order of operation did not affect the outcome, it could be a good idea to have half the translators work in interactive mode first followed by pre-translation mode, while the other half would do the reverse. This would help to ensure that translators were not simply “bored” by the time it came to pre-translation, or that they were not simply remembering solutions that they had encountered when working in interactive mode. Since I was working with such a small group, it was hard to compensate for all these possible influences, but with a larger group, it would be easier to do so.

## **5.2 Suggestions for Future Research**

As outlined in section 0.5, certain constraints were put on the pilot experiment that was carried out in the context of this thesis. Therefore, there are a number of ways in which this research could be expanded in the future.

### *5.2.1 Increase Number of Participants*

Firstly, it would be of great interest to expand the study to include more participants, both translators and evaluators. As emphasized in section 0.5.3, and confirmed in the data analysis presented in Chapter 4, the number of participants in this study was clearly far too small to permit any firm conclusions to be drawn with regard to the benefits of one translation method over the other. Nevertheless, although it did not provide enough data to draw any solid conclusions, I believe that the pilot study experiment nevertheless produced interesting results and raised a number of pertinent questions which would seem to merit further investigation. Of course, as noted above, in any scaled up version of this experiment, it would be preferable to not only increase the number of participants, but also use professional translators instead of students.

### *5.2.2 Add New Languages and Language Directions*

For this pilot study, the translations were done from French to English simply because that is the language pair with which I normally work. In principle, however, the methodology that I developed is language independent, so it would be interesting to do the same experiment using the opposite language direction (i.e., English to French), or even other language pairs. Such investigations might make it possible to determine, for example, if different languages or language combinations are more amenable to one translation method or the other (i.e., interactive translation or pre-translation).

### *5.2.3 Investigate Different Text Types*

Although there has been some research conducted to see what types of texts work best with TMs in general (see section 3.1.2), it could be interesting to re-do the experiment using different text types (e.g., annual reports, research papers, business correspondence), or even different subject fields (e.g., technology, law, literature) to see if a given method of translation (interactive vs pre-translation) works better for a given type of text or subject field.

### *5.2.4 Experiment with Different TM Tools*

Lastly, this pilot study experiment was carried out using a single tool: Fusion Translate. However, as mentioned in the introduction, there are many competing tools available on the market. It would be interesting to see whether the results of the experiment would differ depending on the tool used, or whether all TM tools would produce similar results.

## **5.3 Concluding Remarks**

As mentioned in the introduction, the increasing volume of texts to be translated, together with the lack of professionals available to translate these texts, is leading to a tremendous backlog in the translation industry. Using translation technology tools such as TM systems to their highest potential will help translators to make inroads into the rapidly increasing demand for translation. However, in the words of Claude Bédard (1990:447),

...la technologie n'a de sens que si elle rend effectivement de bons services. Et qu'à moins d'être supérieurement performante, la technologie dépend toujours des méthodes de travail dont on l'entoure dans le contexte opérationnel.

This thesis represents an attempt to learn more about how the different methods of using a TM system can affect productivity, quality and job satisfaction. The results of this small pilot experiment show that while productivity seems comparable across the two methods, the quality of the texts appears to be slightly higher when using interactive translation, and the job satisfaction of translators is considerably higher when using interactive translation. This is something that clients might like to take into account when they are seeking to have their products translated. As noted in section 0.2.3, some clients are choosing pre-translation over interactive translation because they are afraid of giving translators access to their TM databases, but what they may not realize is that they could be getting lower quality texts as a result. Moreover, with so much translation work available, and relatively few translators to do it, translators are in a position to choose which jobs they accept, and since it appears that they do not particularly like working with pre-translated texts, clients who do not permit interactive translation may eventually have trouble finding translators to work for them.

## References

- Allen, Jeffrey. "Post-editing". *In*: Harold Somers, ed. Computers and Translation: A Translator's Guide. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2003, 297-317.
- Anobile, Michael. "Foreword". *In*: Robert C. Sprung, ed. Translating into Success: Cutting-edge Strategies for Going Multilingual in a Global Age. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2000, vii.
- Austermühl, Frank. Electronic Tools for Translators. Manchester: St. Jerome Publishing, 2001.
- Bédard, Claude. "Mémoire de traduction cherche traducteur de phrases." Traduire, 2000, 186: 41-49. URL: [http://www.terminotix.com/t\\_fr/index.htm](http://www.terminotix.com/t_fr/index.htm). [Accessed May 23, 2006].
- Bédard, Claude. "La prétraduction automatique, outil de productivité et d'évolution professionnelle." Meta XXXVII, 4: 1992, 738-760.
- Bédard, Claude. "La prétraduction automatique (PTA): un pas en arrière dans la bonne direction?" Actes du colloque Les industries de la langue – Perspectives des années 1990. Montréal : Office de la langue française and Société des traducteurs du Québec, 1990, 435-448.
- Bowker, Lynne. "Productivity vs Quality? A Pilot Study on the Impact of Translation Memory Systems," Localisation Focus 4(1): 2005, 13-20.
- Bowker, Lynne. "Investigating 'reversible' translation resources: Are they equally useful in both directions?" *In* L. Pérez-González, ed. Speaking in Tongues: Language across contexts and users. Valencia: Universidad de Valencia, 2003a, 201-224.

- Bowker, Lynne. "Terminology Tools for Translators". In: Harold Somers, ed. Computers and Translation: A Translator's Guide. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2003b, 49-65.
- Bowker, Lynne. Computer-Aided Translation Technology: A Practical Introduction. Ottawa: University of Ottawa Press, 2002a.
- Bowker, Lynne. "Information Retrieval in Translation Memory Systems: Assessment of Current Limitations and Possibilities for Future Development." Knowledge Organization Journal 29.3, 4: 2002b, 198-203.
- Bowker, Lynne. "Exploring the potential of corpora for raising language awareness in student translators," Language Awareness 8.3,4: 1999, 160-173.
- Bowker, Lynne and Michael Barlow. "Bilingual concordancers and translation memories: A comparative evaluation" In Proceedings of the Second International Workshop on Language Resources for Translation Work, Research and Training, 2004, 52-61. URL: [http://www.ifi.unizh.ch/cl/yuste/lr4trans-2/WKS\\_PAPERS/8.pdf](http://www.ifi.unizh.ch/cl/yuste/lr4trans-2/WKS_PAPERS/8.pdf) [Accessed June 16, 2005]
- Brockmann, Daniel. "Translation Memories as True Databases: Present and Future". The European Language Resources Association Newsletter 4.3: July – September 1999, 9-11.
- Canadian Translation Industry Sectoral Committee (CTISC). "Survey of the Canadian Translation Industry". URL: <http://www.uottawa.ca/associations/csict/principle.htm>. September 30<sup>th</sup>, 1999. [Accessed April 27<sup>th</sup>, 2006]
- Charron, Marc. "Plus vite, encore plus vite: la traduction à l'heure de la mondialisation". Translation Studies in the New Millennium: An International Journal of Translation and Interpreting 3: 2005, 15-27.

- Cheng, Susan. "Globalizing an e-Commerce Web Site". *In*: Robert C. Sprung, ed. Translating into Success: Cutting-edge Strategies for Going Multilingual in a Global Age. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2000, 29-42.
- Circé, Karine. Traduction automatique, mémoire de traduction ou traduction humaine? Proposition d'une approche pour déterminer la meilleure méthode à adopter, selon le texte. MA Thesis, University of Ottawa, 2005.
- Cronin, Michael. Translation and Globalization. London/New York: Routledge, 2003.
- CTTIC Standard Certification Translation Examination. Marker's Guide: Official Languages. [unpublished manuscript made available to me by Professor Malcolm Williams in April 2006].
- Delisle, Jean. "Criticizing Translations: The Notion of Disparity". *In*: Bowker, Lynne, ed. Text-based Studies: Lexicography, Terminology, Translation. In Honour of Ingrid Meyer. Ottawa: University of Ottawa Press, 2006 (in press).
- Esselink, Bert. A Practical Guide to Localization. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2000.
- Gow, Francie. Metrics for Evaluating Translation Memory Software. MA thesis, University of Ottawa, 2003. URL: <http://www.chandos.ca/thesis.html> [Accessed June 15th, 2006]
- Heyn, Matthias. "Translation Memories: Insight and Prospects". *In*: Bowker, Lynne, Michael Cronin, Dorothy Kenny and Jennifer Pearson, eds. Unity in Diversity? Current Trends in Translation Studies. Manchester: St. Jerome Publishing, 1998, 123-136.

- House, Juliane. A Model for Translation Quality Assessment. Tübingen: TBL-Verlag Narr, 1977.
- House, Juliane. Translation Quality Assessment: A Model Revisited. Tübingen: Narr, 1997.
- Kay, Martin. "The Proper Place of Men and Machines in Language Translation". Machine Translation. 12.1/2: 1997, 3-23.
- L'Homme, Marie-Claude. Initiation à la traductique. Brossard, Québec : Linguatex, 1999.
- Lange, Carmen Andrés and Winfield Scott Bennett. "Combining Machine Translation with Translation Memory at Baan". In: Robert C. Sprung, ed. Translating into Success: Cutting-edge Strategies for Going Multilingual in a Global Age. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2000, 203-218.
- Loffler-Laurian, Anne-Marie. La traduction automatique. France: Presses Universitaires du Septentrion, 1996.
- Macklovitch, Elliott and Graham Russell. "What's Been Forgotten in Translation Memory" In: John S. White, ed. Envisioning Machine Translation in the Information Future. Mexico: Springer, 2000, 137-146.
- Melby, Alan K. The Possibility of Language. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1995.
- Merkel, Magnus. "Consistency and Variation in Technical Translation: A Study of Translators' Attitudes" In: Bowker, Lynne, Michael Cronin, Dorothy Kenny and Jennifer Pearson, eds. Unity in Diversity? Current Trends in Translation Studies. Manchester: St. Jerome Publishing, 1998, 137-149.

- Mogensen, Else. "Orwellian Linguistics: How using computer-aided translation tools impacts the target language". *Language International* 12(5): 2000, 28-31. URL: [http://www.erikseninc.com/pdf/Orwell\\_Mogensen.pdf](http://www.erikseninc.com/pdf/Orwell_Mogensen.pdf) [Accessed June 15th, 2006]
- Mossop, Brian. Revising and Editing for Translators. Manchester UK; Northampton MA: St. Jerome Publishing, 2001.
- Nyberg, Eric, Mitamura, Teruko and Huijsen, Willem-Olaf. "Controlled language for authoring and translation" *In*: Harold Somers, ed. Computers and Translation: A Translator's Guide. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2003, 245-281.
- O'Brien, Sharon. "Practical Experience of Computer-Aided Translation Tools in the Software Localization Industry". *In*: Bowker, Lynne, Michael Cronin, Dorothy Kenny and Jennifer Pearson, eds. Unity in Diversity? Current Trends in Translation Studies. Manchester: St. Jerome Publishing, 1998, 115-122.
- Pinto, Claudia and Andrew Draheim. "Translation Business Practices Report". International Bank for Reconstruction and Development, The World Bank, 2004. URL: [http://www.sign.ro/traducatori/articole/wb\\_survey.pdf?SIGN=fc00b2e2a1d4a577989147d97eec5720](http://www.sign.ro/traducatori/articole/wb_survey.pdf?SIGN=fc00b2e2a1d4a577989147d97eec5720) [Accessed June 15<sup>th</sup>, 2006]
- Schäler, Reinhard (2001) "Beyond Translation Memories". *In*: Carl, M. and Way, A., eds. Proceedings of the Workshop on Example-Based Machine Translation, held at the Machine Translation Summit VIII in Santiago de Compostela, Spain, 18 September 2001. Proceedings published on CD-ROM.
- Somers, Harold. "Translation Memory Systems". *In*: Harold Somers, ed. Computers and Translation: A Translator's Guide. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2003a, 31-47.

- Somers, Harold. "Sublanguage". *In*: Harold Somers, ed. Computers and Translation: A Translator's Guide. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2003b, 283-295.
- Thibodeau, Ricky P. "Making a Global Product at MapInfo Corporation". *In*: Robert C. Sprung, ed. Translating into Success: Cutting-edge Strategies for Going Multilingual in a Global Age. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2000, 127-146.
- Topping, Suzanne. "Sharing Translation Database Information: Considerations for developing an ethical and viable exchange of data." Multilingual Computing and Technology, 11.5: 2000, 59-61.
- Trujillo, Arturo. Translation Engines: Techniques for Machine Translation. London: Springer, 1999.
- Tunick, Laraine. "Language Translation, Localization, and Globalization: World Market Forecasts, Industry Drivers, and eSolutions". Hamilton: Allied Business Intelligence Inc., 2002.
- Webb, Lynn E. Advantages and Disadvantages of Translation Memory: A Cost/Benefit Analysis. MA Thesis. Monterey, CA: Monterey Institute of International Studies, 1998. URL: <http://www.webbsnet.com/translation/thesis.html> [Accessed June 15th, 2006]
- Williams, Malcolm. Translation Quality Assessment: An Argumentation-Centered Approach. Ottawa: University of Ottawa Press, 2004.
- Yunker, John. Beyond Borders: Web globalization strategies. Indianapolis: New Riders Publishing, 2003.

*APPENDICES*

## Appendix A: Source Texts

### ST1

#### DISTRIBUTION DE L'OZONE ET DES SUBSTANCES QUI APPAUVRISSENT L'OZONE DANS L'ATMOSPHERE ET CHANGEMENTS OBSERVÉS

Les principales caractéristiques de la climatologie de l'ozone total ont été découvertes avant 1930 par G.M.B. Dobson. Il s'agissait de sa dépendance à la latitude et à la saison, et des changements quotidiens liés aux conditions météorologiques. Grâce à l'utilisation de nombreuses techniques modernes, on dispose maintenant d'une climatologie plus complète, s'étendant du sol à tous les niveaux de la stratosphère. On peut ainsi mesurer la distribution de l'ozone avec l'altitude, ainsi que la quantité d'ozone total, à partir du sol et de diverses plates-formes, comme des ballons, des aéronefs, des fusées et des satellites. Le grave appauvrissement qui se produit au printemps dans l'Antarctique, le « trou d'ozone », a été découvert en 1985; quelques années plus tard, on constatait un appauvrissement aux latitudes moyennes. Ces dernières années, on a noté une perte significative et croissante au printemps dans l'Arctique. Il y a déjà quelques décennies que l'on mesure les concentrations atmosphériques des substances qui appauvrissent l'ozone, et l'on y détecte maintenant des indices probants de l'effet du Protocole de Montréal.

- Le Canada a commencé à faire des mesures de l'ozone total à cinq stations au début des années 60; depuis 1992, il a douze stations au sol d'observation de l'ozone, équipées de spectrophotomètres Brewer. Les données sur l'ozone total sont utilisées dans les études des tendances, les prévisions d'ozone et de rayonnement UV, et pour la caractérisation et l'étalonnage des mesures satellitaires.
- Le réseau canadien d'ozonosondage a été créé au milieu des années 60, avec quatre stations. À l'heure actuelle, des sondes sont lâchées au moins une fois par semaine à six endroits. Toutes les données des sondages et des mesures au sol sont disponibles auprès du Centre de données mondiales sur l'ozone et sur le rayonnement ultraviolet, qu'exploite le Canada pour l'OMM.
- L'Observatoire de l'ozone stratosphérique d'Eureka (80° N, 86° O), qui fait partie du Réseau OMM/PNUE pour la détection de changements stratosphériques (NDSC), a ouvert ses portes en 1992. On y trouve des lidars de mesure de l'ozone et des aérosols, des spectromètres à transformée de Fourier et d'autres spectromètres destinés à mesurer la composition de la stratosphère. L'Institut de science terrestre et spatiale (ISTS) de l'Ontario, ainsi que la Meteorological Agency and Communications Research Laboratory du Japon collaborent avec Environnement Canada dans la réalisation de ces expériences.
- Les valeurs moyennes annuelles de l'ozone aux stations canadiennes ont baissé d'environ 6 % par rapport aux moyennes d'avant 1980. La baisse atteint son maximum au printemps et son minimum à l'automne. Les mesures des sondes montrent que la plus grande partie de la perte survient dans la basse stratosphère.

## ST2

### Écosystèmes marins

- La sensibilité des processus physiologiques aux UVB solaires exige une fonction de pondération biologique (dite aussi spectre des effets biologiques) pour quantifier l'irradiance efficace. Il faut déterminer les spectres pour un plus grand nombre de fonctions biologiques et faire des estimations de leurs plages de variation pour chaque processus, qui semblent hautement variables d'une espèce à l'autre.
- On a détecté les effets biologiques des UVB jusqu'à plusieurs dizaines de mètres dans la colonne d'eau, et en particulier l'inhibition de la photosynthèse à court terme dans le phytoplancton de l'Antarctique. À mesure que les recherches se poursuivent, on voit converger les estimations quantitatives de l'inhibition de la photosynthèse par le rayonnement UV.
- Certains dinoflagellés toxiques présentent des mécanismes de photoprotection contre les UV qui pourraient leur donner un avantage compétitif susceptible d'entraîner une prédominance accrue des algues toxiques ou nuisibles.
- La hausse des niveaux d'UVB pourrait favoriser les espèces de phytoplancton qui produisent du sulfure de diméthyle, un gaz participant à la formation des nuages, et donc contribuer au changement climatique.
- Pour interpréter les effets des UV dans les eaux canadiennes, il faut reconnaître les différences fondamentales entre l'Antarctique et les eaux de l'hémisphère Nord, dont l'Arctique. Des recherches de base sur les écosystèmes marins du Canada sont donc essentielles.

### Matériaux

Le rayonnement UV a des effets nocifs significatifs sur nombre des matériaux utilisés à l'extérieur. Toute augmentation des UV qui touchent la surface de la Terre entraînera une dégradation de l'infrastructure et, donc, des coûts élevés de réparation et de remplacement.

- Des recherches canadiennes avaient examiné les effets des UV sur les polymères, le bois et le papier, les matériaux de construction, les peintures et revêtements, ainsi que les textiles et les vêtements, mais on ciblait davantage l'évaluation de la résistance au rayonnement des matériaux utilisés dans l'espace et des tissus pour vêtements.
- Les rayons UVB endommagent les polymères synthétiques et d'autres matériaux, mais on n'en comprend pas encore bien les mécanismes au niveau moléculaire, et la combinaison des impacts des rayonnements de courte et de grande longueur d'onde avec d'autres variables environnementales complique encore la situation.
- La couleur des produits de pâte et papier blanchi issus de procédés peu coûteux est altérée par le rayonnement UV. Des chercheurs canadiens ont fait de grands progrès dans la compréhension de ce phénomène. La capacité de réduire cette altération de la couleur pourrait entraîner une grande expansion du marché pour cette catégorie de produits.
- Les matériaux de construction non plastiques, comme les revêtements d'étanchéité et les produits de scellement pour l'extérieur sont actuellement l'objet de recherches quant à leur résistance aux UV, mais pas précisément dans le contexte d'une irradiation accrue liée à l'ozone.

## Appendix B: Pre-translated Source Texts

### ST1

#### THE DISTRIBUTION OF OZONE AND OZONE-DEPLETING SUBSTANCES IN THE ATMOSPHERE AND OBSERVED CHANGES

Les principales caractéristiques de la climatologie de l’ozone total ont été découvertes avant 1930 par G.M.B. Dobson. Il s’agissait de sa dépendance à la latitude et à la saison, et des changements quotidiens liés aux conditions météorologiques. Grâce à l’utilisation de nombreuses techniques modernes, on dispose maintenant d’une climatologie plus complète, s’étendant du sol à tous les niveaux de la stratosphère. On peut ainsi mesurer la **ozone distribution** avec l’altitude, ainsi que la **total ozone amount**, à partir du sol et de diverses plates-formes, comme des ballons, des aéronefs, des fusées et des satellites. Le **severe depletion** qui se produit au printemps dans l’Antarctique, le « **ozone hole** », a été découvert en 1985; quelques années plus tard, on constatait un **midlatitude depletion**. Ces **recent years**, on a noté une perte significative et croissante au printemps dans l’Arctique. Il y a déjà quelques décennies que l’on mesure les concentrations atmosphériques des substances qui appauvrissent l’ozone, et l’on y détecte maintenant des indices probants de l’effet du **Montreal Protocol**.

- Le Canada a commencé à faire des mesures de l’ozone total à cinq stations au début des années 60; depuis 1992, il a douze stations au sol d’observation de l’ozone, équipées de spectrophotomètres Brewer. Les données sur l’ozone total sont utilisées dans les études des tendances, les prévisions d’ozone et de **UV radiation**, et pour la caractérisation et l’étalonnage des **satellite measurements**.
- Le réseau canadien d’ozonosondage a été créé au milieu des années 60, avec quatre stations. À l’heure actuelle, des sondes sont lâchées au moins une fois par semaine à six endroits. Toutes les données des sondages et des mesures au sol sont disponibles auprès du **World ozone and ultraviolet radiation data centre (woudc)**, qu’exploite le Canada pour l’OMM.
- L’Observatoire de l’ozone stratosphérique d’Eureka (80° N, 86° O), qui fait partie du Réseau OMM/PNUE pour la détection de changements stratosphériques (NDSC), a ouvert ses portes en 1992. On y trouve des lidars de **measuring ozone** et des aérosols, des spectromètres à transformée de Fourier et d’autres spectromètres destinés à mesurer la composition de la stratosphère. L’Institut de science terrestre et spatiale (ISTS) de l’Ontario, ainsi que la Meteorological Agency and Communications Research Laboratory du Japon collaborent avec Environnement Canada dans la réalisation de ces expériences.
- Les valeurs moyennes annuelles de l’ozone aux **canadian stations** ont baissé d’environ 6 % par rapport aux moyennes d’avant 1980. La baisse atteint son maximum au printemps et son minimum à l’automne. **Additionally, the bulk of this decline was predicted by those models to occur in the upper stratosphere, near 40 km.**

## ST2

### Écosystèmes marins

- La sensibilité des processus physiologiques aux **solar UV-B** exige une fonction de pondération biologique (dite aussi spectre des effets biologiques) pour quantifier l'irradiance efficace. Il faut déterminer les spectres pour un plus grand nombre de fonctions biologiques et faire des estimations de leurs plages de variation pour chaque processus, qui semblent hautement variables d'une espèce à l'autre.
- On a détecté les effets biologiques des UVB jusqu'à plusieurs dizaines de mètres dans la colonne d'eau, et en particulier l'inhibition de la **short-term photosynthesis** dans le phytoplancton de l'Antarctique. À mesure que les recherches se poursuivent, on voit converger les estimations quantitatives de l'inhibition de la photosynthèse par le **UV radiation**.

**Some toxic dinoflagellates show UV-photoprotective mechanisms [Carreto et al. 1989] which might give them a competitive edge in an enhanced-UV environment.**

**Increases in UV-B could also favour species of phytoplankton that produce dimethyl sulphide (DMS), a reactive gas implicated in cloud formation and, hence, local climate and heat balance.**

- Pour interpréter les **UV effects** dans les eaux canadiennes, il faut reconnaître les différences fondamentales entre l'Antarctique et les eaux de l'hémisphère Nord, dont l'Arctique. Des recherches de base sur les écosystèmes marins du Canada sont donc essentielles.

### Matériaux


Le **UV radiation** a des effets nocifs significatifs sur nombre des matériaux utilisés à l'extérieur. Toute augmentation des UV qui touchent la **earth's surface** entraînera une dégradation de l'infrastructure et, donc, des coûts élevés de réparation et de remplacement.

- Des recherches canadiennes avaient examiné les **UV effects** sur les polymères, le bois et le papier, les **building materials**, les peintures et revêtements, ainsi que les textiles et les vêtements, mais on ciblait davantage l'évaluation de la résistance au rayonnement des matériaux utilisés dans l'espace et des tissus pour vêtements.
- Les rayons UVB endommagent les polymères synthétiques et d'**other materials**, mais on n'en comprend pas encore bien les mécanismes au niveau moléculaire, et la combinaison des impacts des **short and long wavelength radiation** avec d'autres variables environnementales complique encore la situation.
- La couleur des produits de pâte et papier blanchi issus de procédés peu coûteux est altérée par le **UV radiation**. Des chercheurs canadiens ont fait de grands progrès dans la compréhension de ce phénomène. La capacité de réduire cette altération de la couleur pourrait entraîner une grande expansion du marché pour cette catégorie de produits.
- Les **building materials** non plastiques, comme les revêtements d'étanchéité et les produits de scellement pour l'extérieur sont actuellement l'objet de recherches quant à leur résistance aux UV, mais pas précisément dans le contexte d'une irradiation accrue liée à l'ozone.

## Appendix C: Tutorial on Fusion Translate

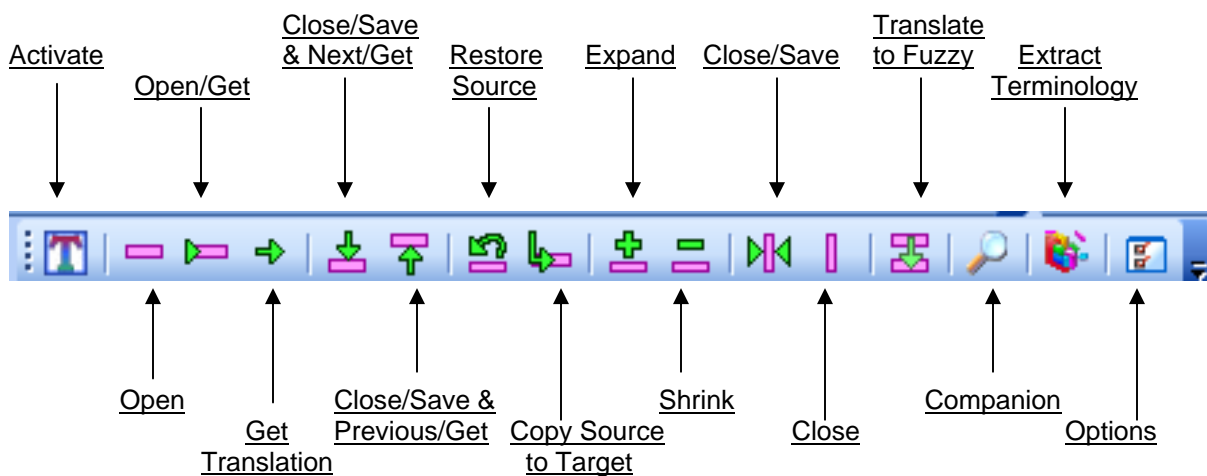
### Practice for Fusion Translate

#### 1. *Opening and Saving Files*

- Along with this document in the email you received there are four Fusion files as well as the text to translate. Save the files to your computer.
- Open **Text to translate - Fusion Practice**. If you do not see the Fusion toolbar, right-click on any toolbar and select **Fusion Translate**.
- On the Fusion toolbar, select **Activate**  and Fusion Translate will load.
- From the **File** menu in Fusion Translate select **Open** and open the files that you saved from the email:
  - Fusion Test – Obesity** (this is a Project file)
  - Fusion Test TM – Obesity** (this is a Translation Memory file)
  - Fusion Test Terminology – Obesity** (this is a Terminology file)
  - Fusion Test Dictionary – Obesity** (this is a Dictionary file)

#### 2. *Interactive Translation with the TM*

- In the **Text to translate - Fusion Practice** document, put your cursor anywhere in the first sentence, and select **Open/Get** from the Fusion toolbar:



- The **Fusion Companion** will open. It is important that you are able to see the Companion while you are translating. It will show you term equivalents, and

will indicate what needs to be modified in the target text in the case of a partial match.

You can choose how to arrange the windows from the **Window** menu in the Companion (**Align Left, Align Top, Align Right, Align Bottom**).

- You will also notice that the first sentence in the source text has been encased in a blue box, and there is a yellow box below it in which the TM has pasted a partial match.

The Companion will show you what needs to be changed in the target text to make it correspond to the source text by highlighting the part(s) of the sentence which do not correspond.

In the case of an exact match, the target text box will be green (or in some cases may look the same color as the source text box).

In the case of no match at all, the target text box will be orange and there will be no text in the box. It is up to you to enter a translation.

- In the case of this first sentence, a fuzzy match for the entire sentence has been proposed. However, if you look at the left-hand side of the Companion window, you will see that other types of matches are also available for your inspection. All types of matches that are found by the system, including exact and fuzzy matches, but also sub-segment or dictionary matches, will be presented for you to consult. Although Fusion tries to rank the matches in an order from most to least helpful, you as the translator are the best judge of what is useful to you, so don't feel obliged to consult only the first match that is presented.
- When you have finished translating one sentence, select **Close/Save & Next/Get** on the Fusion toolbar to move on to the next sentence and close the current one. The source text sentence will disappear and you will be left with only the target text sentence.
- When you get to sentence 3, you will notice that you are presented with an orange box, which means that no exact or fuzzy sentence matches were found in the TM. However, in the Companion, you can see that a number of expression matches were found. Start typing your own translation of the text directly into the orange box until you reach the term "graisses saturées". The Companion has some information that might help you translate this expression.
  - First, click on "graisses saturées" in the left-hand pane of the Companion window. You will see that the contents of the Dictionary are displayed, which indicates that the translation is "saturated fats"
  - You can now type this expression directly into your TT, or, if you prefer, you can try to copy and paste it into the text by doing the following.
  - If you click on "Select translation" in the left-hand pane, you will see that on the right-hand side, the English sentence containing the term "saturated fats" is displayed. Use your cursor to select "saturated fats" in the English sentence and then click on "Send to Word" (far right) to

paste the term directly in to your text.

- Translate the entire document using the TM. When you get to the last sentence, you can either choose **Close/Save & Next/Get** or **Close**.

### **3. Clean-Up**

You will need to clean up your document to remove source text and segment delimiting marks.

- Return to the main Fusion window (not the Companion). From the **Tools** menu, select **Clean-up...** In the **Clean up Documents** window select **Add** and select the **Text to translate – Fusion Practice** document and click **Open**.
- Now click **Next** in the **Clean up Documents** window.
- In the **Destination** section, uncheck the *Move cleaned documents to folder* box.
- Click **Next**.
- In the **Clean up and Updating** section, select: *Don't Update the TM but clean up the translation unit in the document.*
- Click **Next** then **Start**.
- When the clean up is complete, click **Done**.

## Appendix D: Instructions for Translators, Time Sheet, Resources Sheet and Translator Questionnaire

### Instructions<sup>32</sup>

*NOTE: In order to comparatively assess the quality of the translations, it is important that each translator has access to the same resources. For the purpose of this experiment, you may use the following online resources:*

*Termium (accessed through <http://www.biblio.uottawa.ca/index-e.php>)*

*Le grand dictionnaire terminologique ([www.granddictionnaire.com](http://www.granddictionnaire.com))*

*Merriam-Webster Online ([www.m-w.com](http://www.m-w.com))*

*You can also use any resources available in the Writing Centre.*

*From the **Start** menu, select **All Programs**. Under **English**, select **Dictionaries**. From there you have access to:*

*Hachette Oxford (English-French)*

*Random House Webster (English)*

*Under **Français/Dictionnaires** you have access to:*

*Le Dictionnaire Hachette-Oxford (French-English - you may need to change the language direction by clicking F→E at the top right)*

*Le Petit Robert*

*You also have access to any Microsoft Word functions.*

*Please do not perform searches online, or use resources other than those listed above.*

### Part 1 - Interactive Translation Using Fusion Translate

**For this section you will use Fusion Translate to translate your text.**

1. From **u:\Julian Wallis\Fusion Experiment\Source Texts**, copy **Blue ST2** and paste it into the following folder:  
**u:\Julian Wallis\Fusion Experiment\TMtranslations**. Now rename the copied file **TMtranslation\_Blue**. This is the file you will use with the TM.
2. Open the **TMtranslation\_Blue** file. If a dialog box appears asking if you would like to deactivate the TRADOS toolbar and activate the Fusion toolbar, click **Yes**.
3. Now open the following files that you will find in  
**u:\Julian Wallis\Fusion Experiment\Translation Memories**

---

<sup>32</sup> Each translator was given a separate set of instructions. Red and Blue completed the translations following this outline (translate ST1 in pre-translation mode and ST2 in interactive mode), and Green and Yellow did the opposite (translate ST1 in interactive mode and ST2 in pre-translation mode).

**Blue Project**  
**Blue TM**  
**Blue Dictionary**  
**Blue Terminology**

4. In Fusion Translate, go to **Tools\Options...**. The **Project Options** dialog box will appear. Under the **Translate** tab, in the **General** section, make sure the "Copy source into target segment when no matches are found" box is not checked. In the **Translation Memory** section, make sure the "Translate segment with \_\_\_\_% or higher match value" is set at 50%. Click **Ok**.
5. There is a clock in the bottom corner of the computer screen. Record the Start time on your *Time Sheet* NOW. Be sure you are recording this information in the appropriate place (i.e., ST2).
6. Place your cursor in the first sentence of the Word document. On the Fusion toolbar in your Microsoft Word file, click the **Open/Get** command. The **Fusion Companion** will open, and your text will be ready for translation. Make sure you can see the Companion while you are translating. Also be sure to save regularly throughout your translation (Ctrl-s).
7. When you have finished your translation, use the **Clean up** option in **Fusion Translate's Tools** menu. (In the **Clean up Documents** window select **Add** and select **TMtranslation\_Blue** from:  
**u:\Julian Wallis\Fusion Experiment\TMtranslations**. Click **Open**. Now click **Next** in the **Clean up Documents** window. In the **Destination** section, uncheck the *Move cleaned documents to folder* box. Click **Next**. In the **Clean up and Updating** section, select: **Don't Update the TM but clean up the translation unit in the document**. Click **Next** then **Start**. When the clean up is complete, click **Done**.)
8. Record the Finish time on your *Time Sheet* NOW. Be sure you are putting the information in the correct place (i.e., ST2).
9. Make a note of any additional resources you used to help you with your translation on the *Resources Sheet* NOW. Be sure you are putting the information in the correct place (i.e., ST2)
10. You may now close Fusion Translate. When you are asked if you would like to save the changes, click **Yes**.

## **Part 2 - Translating a Pre-translated Text**

**For this section, you will not use Fusion Translate. You will use the pre-translated text specified below.**

1. Open **Blue ST1** from **u:\Julian Wallis\Fusion Experiment\Source Texts**. You will also have this document in hard copy.
2. Record the Start time on your *Time Sheet* NOW. Be sure you are recording the information in the appropriate place (i.e., ST1).
3. Open a new Microsoft Word document and save it as **Pretranslation\_Blue** in **u:\Julian Wallis\Fusion Experiment\Pre-translations**.
4. Translate **Blue ST1** in the new Word document using the pre-translated text you will find in **u:\Julian Wallis\Fusion Experiment\Pre-translated Texts\Blue PT**. You will also have this document in hard copy. Be sure to save regularly throughout your translation (Ctrl-s).
5. Record the Finish time on your *Time Sheet* NOW. Be sure you are putting the information in the correct place (i.e., ST1).
6. Make a note of any additional resources you used to help you with your translation on the *Resources Sheet* NOW. Be sure you are putting the information in the correct place (i.e., ST1)
7. Now you may close all documents (make sure they are saved!) and fill out the *Translator Questionnaire*.
8. That's it! Please leave your *Time Sheet*, *Resources Sheet*, and *Translator Questionnaire* with me.

## **Time sheet**

It is very important to keep an accurate record of the time it takes you to complete each of your translations. Before you begin translating, record the Start time below. Once you are satisfied with your translation, record the Finish time. You can use the clock in the bottom left-hand corner of the computer screen.

### **ST1**

Start time: \_\_\_\_\_

Finish time: \_\_\_\_\_

### **ST2**

Start time: \_\_\_\_\_

Finish time: \_\_\_\_\_

**Resources sheet**

Please make a brief note of any additional resources that you used to help with the translation of each text (e.g. Termium, Le Grand dictionnaire terminologique, Robert-Collins dictionary, etc.).

**ST1**

---

---

---

---

---

---

---

---

---

---

---

---

**ST 2**

---

---

---

---

---

---

---

---

---

---

---

---

## **Translator Questionnaire**

**Please answer the following questions:**

1. Please note what you liked and disliked, if anything, about working interactively with the Translation Memory.
  
2. When working interactively with the Translation Memory, did you use the TM as your principal resource? If not, what other resource(s) ranked ahead of the TM with regard to usefulness for this translation?
  
3. Please note what you liked and disliked, if anything, about working with the pre-translated text.
  
4. When working with the pre-translated text, did you use the pre-translated material as your principal resource? If not, what other resource(s) ranked ahead of the pre-translated material with regard to usefulness for this translation?
  
5. Did you feel that one method of translation worked better for you than the other? Why?
  
6. Were you more satisfied with one translation over the other? If so, please explain why and indicate which method was used.

7. What resources do you normally use to aid you in your translations (i.e. dictionaries, Termium, GDT etc...)?
  
8. Did you find either method of translation (interactive or pre-translation mode) more efficient than the method of translation you are used to?
  
9. Did you experience any difficulty with either method of translation? If so, please explain.
  
10. If you had the choice, would you choose to work with translation memories on a regular basis – either in interactive or pre-translation mode - or would you prefer not to use these tools at all? Explain.
  
11. If a client absolutely required you to work using a translation memory, but gave you a choice between working interactively with a TM or using a pre-translated text, which option would you choose and why?

## Appendix E: Completed Time Sheets

### Time sheet - Blue

It is very important to keep an accurate record of the time it takes you to complete each of your translations. Before you begin translating, record the Start time below. Once you are satisfied with your translation, record the Finish time. You can use the clock in the bottom left-hand corner of the computer screen.

#### ST1

Start time: \_\_\_\_\_ 4:21 \_\_\_\_\_      Finish time: \_\_\_\_\_ 5:10 \_\_\_\_\_

#### ST2

Start time: \_\_\_\_\_ 3:28 \_\_\_\_\_      Finish time: \_\_\_\_\_ 4:17 \_\_\_\_\_

## **Time sheet - Red**

It is very important to keep an accurate record of the time it takes you to complete each of your translations. Before you begin translating, record the Start time below. Once you are satisfied with your translation, record the Finish time. You can use the clock in the bottom left-hand corner of the computer screen.

### **ST1**

Start time: 4:33pm

Finish time: 5:28pm

### **ST2**

Start time: 3:26pm

Finish time: 4:30pm

## **Time sheet - Green**

It is very important to keep an accurate record of the time it takes you to complete each of your translations. Before you begin translating, record the Start time below. Once you are satisfied with your translation, record the Finish time. You can use the clock in the bottom left-hand corner of the computer screen.

### **ST1**

Start time: 3:26pm

Finish time: 4:17pm

### **ST2**

Start time: 4:23pm

Finish time: 5:23pm

## **Time sheet - Yellow**

It is very important to keep an accurate record of the time it takes you to complete each of your translations. Before you begin translating, record the Start time below. Once you are satisfied with your translation, record the Finish time. You can use the clock in the bottom left-hand corner of the computer screen.

### **ST1**

Start time: \_\_\_\_\_ 15:37 \_\_\_\_\_

Finish time: \_\_\_\_\_ 16:23 \_\_\_\_\_

### **ST2**

Start time: \_\_\_\_\_ 16:27 \_\_\_\_\_

Finish time: \_\_\_\_\_ 17:13 \_\_\_\_\_



## **Resources sheet - Red**

Please make a brief note of any additional resources that you used to help with the translation of each text (e.g. Termium, Le Grand dictionnaire terminologique, Robert-Collins dictionary, etc.).

### **ST1**

Termium

---

Granddictionnaire

---

Oxford Hachette

---

m-w.com

---

---

---

---

---

---

---

---

### **ST 2**

Termium

---

Granddictionnaire

---

Oxford-Hachette

---

---

---

---

---

---

---

---

**Resources sheet - Green**

Please make a brief note of any additional resources that you used to help with the translation of each text (e.g. Termium, Le Grand dictionnaire terminologique, Robert-Collins dictionary, etc.).

**ST1**

Termium  
\_\_\_\_\_  
Le Grand Dictionnaire  
\_\_\_\_\_  
Hachette Oxford Dictionary  
\_\_\_\_\_  
m-w.com  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

**ST 2**

Termium  
\_\_\_\_\_  
Le Grand Dictionnaire  
\_\_\_\_\_  
Hachette Oxford Dictionary  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

**Resources sheet - Yellow**

Please make a brief note of any additional resources that you used to help with the translation of each text (e.g. Termium, Le Grand dictionnaire terminologique, Robert-Collins dictionary, etc.).

**ST1**

Hachette-Oxford Eng/Fr Dictionary (electronic) \_\_\_\_\_

Termium Plus \_\_\_\_\_

Le grand dictionnaire terminologique \_\_\_\_\_

~~Random House Webster's Unabridged Dictionary~~ \_\_\_\_\_

MS Word thesaurus \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

**ST 2**

Hachette-Oxford E/F Dictionary \_\_\_\_\_

Termium Plus \_\_\_\_\_

Le GDT \_\_\_\_\_

MS Word Thesaurus \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

## Appendix G: Completed Questionnaires

### Translator Questionnaire - Blue

Please answer the following questions:

1. Please note what you liked and disliked, if anything, about working interactively with the Translation Memory.

***Liked it but it took me a while to find English equivalent, as it was not highlighted.***

2. When working interactively with the Translation Memory, did you use the TM as your principal resource? If not, what other resource(s) ranked ahead of the TM with regard to usefulness for this translation?

***I mostly used TM but still had to resort to Termium but not as much as ST1 (pre-translation).***

3. Please note what you liked and disliked, if anything, about working with the pre-translated text.

***- Disliked, had to keep running to Termium/Webster many more times than TM.  
- Enjoyed working more with TM.***

4. When working with the pre-translated text, did you use the pre-translated material as your principal resource? If not, what other resource(s) ranked ahead of the pre-translated material with regard to usefulness for this translation?

***Really had to rely on Termium for many technical concepts I think would have been more readily available in TM so it seemed more tedious than TM. Also had to use Webster which I did not use in TM.***

5. Did you feel that one method of translation worked better for you than the other? Why?

***TM worked better. It seemed less long as info was readily available.  
- Did not use Termium as much  
Also easy to use for first time with no major problems.***

6. Were you more satisfied with one translation over the other? If so, please explain why and indicate which method was used.

***Tried my best for both :) I think TM might be a bit more accurate because the 'troublesome terms' were usually in TM while for the other one, if it was not in Termium I struggled with certain terms.***

7. What resources do you normally use to aid you in your translations (i.e. dictionaries, Termium, GDT etc...)?

***Termium  
Dictionaries  
Specialized websites  
\*Never any TM.***

8. Did you find either method of translation (interactive or pre-translation mode) more efficient than the method of translation you are used to?

***Pleasantly surprised @ TM = Biggest reason: Time effective instead of running to a dictionary each time. You still need to use termium but I relied on it much less than ST-1 or than I usually do. Same thing for the dictionary and even managed without having to look through countless websites.***

9. Did you experience any difficulty with either method of translation? If so, please explain.

***ST-1. Lack of certain terms in Termium left me with terms I did not how to translated. ST-2: Frustrating having to search for Eng. Equivalent as it was not hightlighted.***

10. If you had the choice, would you choose to work with translation memories on a regular basis – either in interactive or pre-translation mode - or would you prefer not to use these tools at all? Explain.

***I would use TM only if it did have a "good memory" but would still have to resort to other aids such as Termium.***

11. If a client absolutely required you to work using a translation memory, but gave you a choice between working interactively with a TM or using a pre-translated text, which option would you choose and why?

***I would choose TM only if I knew the memory was “good”. I really enjoyed TM! You made me a believer :) I also liked the TM outline, easy to use. Easier to separate and translate sentence by sentence than side by side. “Cleaner” translation and seemed quicker than pre-translated text.***

## Translator Questionnaire - Red

Please answer the following questions:

1. Please note what you liked and disliked, if anything, about working interactively with the Translation Memory.

**like: quick answers**

**dislike: half of the searches were barely relevant anyway.  
technical problems with copy&paste  
changes had to made after clean-up**

2. When working interactively with the Translation Memory, did you use the TM as your principal resource? If not, what other resource(s) ranked ahead of the TM with regard to usefulness for this translation?

**Yes. the second most used was Termium**

3. Please note what you liked and disliked, if anything, about working with the pre-translated text.

**like: The names of organizations ~~prev~~ being provided.**

**dislike, having to work w 3 documents (source, target, PT)**

4. When working with the pre-translated text, did you use the pre-translated material as your principal resource? If not, what other resource(s) ranked ahead of the pre-translated material with regard to usefulness for this translation?

**no. I did a free translation using other dictionaries until I came across something I couldn't find. At that point I referred myself to the PT text.**

5. Did you feel that one method of translation worked better for you than the other? Why?

**no. I don't think it made much difference because the PT text had about as many matches as the TM. Given that the subject matter was so technical anyway, I based myself on Termium & Granddictionarie more.**

6. Were you more satisfied with one translation over the other? If so, please explain why and indicate which method was used.

***I didn't particularly like either. The PT one allowed me to phrase things more in my own way, but it resulted in an inconsistent style.***

7. What resources do you normally use to aid you in your translations (i.e. dictionaries, Termium, GDT etc...)?

***for technical translation, Termium and websites.***

8. Did you find either method of translation (interactive or pre-translation mode) more efficient than the method of translation you are used to?

***Both were more efficient than translating from scratch (my usual method)***

9. Did you experience any difficulty with either method of translation? If so, please explain.

***TM, wouldn't copy and paste or 'Send to Word' my answers***

***PT, with my own wording if switched from active to passive voice, leading to an inconsistent style.***

10. If you had the choice, would you choose to work with translation memories on a regular basis – either in interactive or pre-translation mode - or would you prefer not to use these tools at all? Explain.

***Yes. They are more efficient when it comes to technical translations. Plus, if I were to work w a TM that had primarily my own translations in it, I would be more comfortable accepting suggestions.***

11. If a client absolutely required you to work using a translation memory, but gave you a choice between working interactively with a TM or using a pre-translated text, which option would you choose and why?

***Working with the TM. Even though I enjoyed the PT text more for this experiment, the TM would allow me more of the opportunity to know the contexts from which matches were pulled.***

## Translator Questionnaire - Green

Please answer the following questions:

1. Please note what you liked and disliked, if anything, about working interactively with the Translation Memory.

***I enjoyed all of the expression match options provided by the TM. Unfortunately there just wasn't enough screen space, so if I started using other resources (e.g. Termium) and I couldn't see the Fusion companion, sometimes I would forget to look for matches.***

2. When working interactively with the Translation Memory, did you use the TM as your principal resource? If not, what other resource(s) ranked ahead of the TM with regard to usefulness for this translation?

***Very difficult to say... I tried to always start with the TM, but if I started using another resource for a bit I had to remind myself to return to the TM. Overall though, I used the TM the most, then Termium, then the Oxford Hachette, then the GDT.***

3. Please note what you liked and disliked, if anything, about working with the pre-translated text.

***I found it time-consuming to check the pre-translated text and I was frustrated that I couldn't fit more on the screen. I had the ST2 open and my TT doc, and to check the pre-translated text I had to check around too much.***

4. When working with the pre-translated text, did you use the pre-translated material as your principal resource? If not, what other resource(s) ranked ahead of the pre-translated material with regard to usefulness for this translation?

***No. My principle resource was Termium. However, the pre-translated text was useful for select wording/terminology.***

5. Did you feel that one method of translation worked better for you than the other? Why?

***Very difficult to say. I found ST2 much more difficult to translate than ST1, but science is not my forte... I feel like much of the terminology in ST1 could have***

***been found easily in Termium (and was!) Overall though, TM was easier. I thought the ST was easier, though.***

6. Were you more satisfied with one translation over the other? If so, please explain why and indicate which method was used.

***No, I feel similarly about both. However, I think I would have been more satisfied w/ Text 2 if I had taken more time at the end to re-word, change sentences from passive to active, etc.***

7. What resources do you normally use to aid you in your translations (i.e. dictionaries, Termium, GDT etc...)?

***I usually rely on Termium, GDT, M-W.com, TV5 Mediadico, and Government of Canada sites. I also sometimes use paper reference (e.g. le Petit Robert, Collins Fr./Eng./Eng./Fr. Dictionary.***

8. Did you find either method of translation (interactive or pre-translation mode) more efficient than the method of translation you are used to?

***I think the TM would have been more efficient I had had more practice using it or if I were an expert user with a decent TM. I find pre-translation to be a bit restrictive. That said, there is always the inclination to use the match that's provided, particularly if you're not overly familiar with the subject matter.***

9. Did you experience any difficulty with either method of translation? If so, please explain.

***My main frustration was lack of screen space to use all the tools (word, dictionaries, etc.) This affected both methods equally.***

10. If you had the choice, would you choose to work with translation memories on a regular basis – either in interactive or pre-translation mode - or would you prefer not to use these tools at all? Explain.

***I think TM's are definitely beneficial. If you have a good TM w/ relevant texts and are a proficient user, I think you'll have higher productivity.***

11. If a client absolutely required you to work using a translation memory, but gave you a choice between working interactively with a TM or using a pre-translated text, which option would you choose and why?

***I would absolutely prefer working with the TM. Having to use a pre-translated text requires you to use terminology that you may not agree with or that would not be your first choice. The TM provides you options that you can select or ignore. I use it for inspiration or for the matches – either way, it's by far preferable to a pre-translated text, in my opinion.***

## **Translator Questionnaire - Yellow**

Please answer the following questions:

1. Please note what you liked and disliked, if anything, about working interactively with the Translation Memory.

**:) - The TM offered suggestion I didn't think of right away, and facilitated finding the right word/expression.**

**:) - Especially good for finding equivalents to names of organizations.**

**:( - Too many windows open at the same time.**

2. When working interactively with the Translation Memory, did you use the TM as your principal resource? If not, what other resource(s) ranked ahead of the TM with regard to usefulness for this translation?

**I used the TM as my main resource.**

3. Please note what you liked and disliked, if anything, about working with the pre-translated text.

**- The more complicated terms were not offered as fuzzy matches; I couldn't rely on the pre-translated text like I could the TM.**

4. When working with the pre-translated text, did you use the pre-translated material as your principal resource? If not, what other resource(s) ranked ahead of the pre-translated material with regard to usefulness for this translation?

**My principle resource was Termium, GDT, & the Hachette-Oxford Dictionary.**

5. Did you feel that one method of translation worked better for you than the other? Why?

**- The TM helped me concentrate on one sentence at a time – using the pre-translated text, the translation process was more frustrating... too many sentences jumped out at me at the same time.**

6. Were you more satisfied with one translation over the other? If so, please explain why and indicate which method was used.

**- The first one (w/ TM) I preferred.  
I felt like I translated faster, & w/ less errors**

7. What resources do you normally use to aid you in your translations (i.e. dictionaries, Termium, GDT etc...)?

**-Hachette Oxford                      -Google  
-Termium  
-GDT**

8. Did you find either method of translation (interactive or pre-translation mode) more efficient than the method of translation you are used to?

**- I didn't find the "pre-trans mode" particularly different than my traditional (i.e. no TM) mode of translation. But I think the TM made me more productive & helped me translate w/ better quality.**

9. Did you experience any difficulty with either method of translation? If so, please explain.

**- I was not as confident w/ the pre-translated text as I was w/ the TM text.**

10. If you had the choice, would you choose to work with translation memories on a regular basis – either in interactive or pre-translation mode - or would you prefer not to use these tools at all? Explain.

**- Yes. I would definitely work w/ a TM, given I had a high quality corpus from which to work.**

11. If a client absolutely required you to work using a translation memory, but gave you a choice between working interactively with a TM or using a pre-translated text, which option would you choose and why?

**~~TM – at least w/ TM I'd already have an idea of the style/vocab~~ At least w/ TM, I know where the translations have come from (me), but the pre-translated text could be from anyone.**

## **Appendix H: Evaluator Guidelines**

### **Evaluation Guidelines**

*For my M.A thesis, I am doing a study on translation memory software and two different approaches to translating with the help of this software: interactive translation and pre-translation. Your role will be to help assess the quality of the translations produced.*

*I have provided you with two original source texts (ST1 and ST2), four translations of ST1 and four translations of ST2, as well as a Final Score Sheet. I have also attached an Evaluation Grid to the front of each translation for you to use as an evaluation guideline. Feel free to mark up the translations, and add comments or suggestions.*

Please complete the evaluations in the following manner:

- Mark the translations as if they were at the level of fourth year students at the School of Translation and Interpretation (STI) at the University of Ottawa
- Use the Evaluation Grid attached to the front of each translation as an evaluation guideline
- Give each translation a percentage grade following the numerical deductions on the Evaluation Grid
- Record the percentage grade for each translation in each source text group (ST1 and ST2) on the Final Score Sheet

*Thank you very much for volunteering your time!*

**Final Score Sheet**

**ST1**      Blue      Grade: \_\_\_\_\_

Red      Grade: \_\_\_\_\_

Green      Grade: \_\_\_\_\_

Yellow      Grade: \_\_\_\_\_

**ST2**      Blue      Grade: \_\_\_\_\_

Red      Grade: \_\_\_\_\_

Green      Grade: \_\_\_\_\_

Yellow      Grade: \_\_\_\_\_

## **Evaluation Grid**

*Please use the following guidelines to aid you in evaluating the translations. These are general guidelines and are only suggestions. If you feel that a Considerable or Minor Error made a significant change in the meaning or coherence of the text, mark the translation accordingly. All texts begin with 100 percent and then marks are deducted accordingly.*

| <b><u>Major Errors</u></b> (-5)                      | <u>No. of errors</u> | <u>Deduction</u> |
|--|----------------------|------------------|
| Change in meaning (lack of comprehension of text)    | _____                | _____            |
| Opposite meaning                                     | _____                | _____            |
| Nonsense   | _____                | _____            |
| Omission   | _____                | _____            |
| Addition adversely affecting the meaning of the text | _____                | _____            |
| <b><u>Considerable Errors</u></b> (-3)               |                      |                  |
| Imprecise meaning                                    | _____                | _____            |
| Grammar  | _____                | _____            |
| Poor sentence/phrase construction                    | _____                | _____            |
| Poor choice of term                                  | _____                | _____            |
| Ambiguity  | _____                | _____            |
| Gallicism  | _____                | _____            |
| Mistranslation of one word                           | _____                | _____            |
| <b><u>Minor errors</u></b> (-1)                      |                      |                  |
| Spelling   | _____                | _____            |
| Punctuation  | _____                | _____            |
| Incorrect register of language                       | _____                | _____            |
| Awkwardness  | _____                | _____            |
| Redundancy   | _____                | _____            |

### **Other things to consider:**

### Comments

Overall coherence \_\_\_\_\_  
\_\_\_\_\_  
Overall fluency \_\_\_\_\_  
\_\_\_\_\_

**Total:**

\_\_\_\_\_

## Appendix I: Completed Final Score Sheets

### Final Score Sheet – Evaluation 1

**ST1**      Blue      Grade:   0  

Red      Grade:  27 

Green      Grade:  57 

Yellow      Grade:  21 

**ST2**      Blue      Grade:   0  

Red      Grade:  35 

Green      Grade:  60 

Yellow      Grade:  14

**Final Score Sheet – Evaluation 2**

**ST1**      Blue      Grade: 27

Red      Grade: 25

Green      Grade: 45

Yellow      Grade: 65

**ST2**      Blue      Grade: 31

Red      Grade: 78

Green      Grade: 73

Yellow      Grade: 40

## Appendix J: Target Texts

### ST1 - Blue

Distribution of ozone and ozone-depleting substances in the atmosphere and observed changes.

The main characteristics in the climatology of the ozone had been discovered before 1930 by G.M.B. Dobson. It dealt with his dependence on latitude, the season and to daily changes linked to meteorological conditions. Thanks to the use of numerous modern technologies, we now dispose of a more complete climatology, stretching from the ground to all levels of the stratosphere. We can thus measure the ozone distribution with altitude as well as the total ozone amount, from the ground and various platforms, such as balloons, aircrafts, rockets and satellites. The severe depletion that results in the spring in the Antarctic, the «ozone hole» was discovered in 1985. Some years later, we observed a midlatitude depletion. In recent years, we noted a significant loss and growth of springtime in the Arctic. Some decades ago, we started measuring the atmospheric concentration of substances that depleted the ozone, and now detect clues probative of the effect of the Montreal Protocol.

Canada started measuring the total ozone in five stations at the beginning of the sixties, since 1992. There are twelve stations on the ozone's observation ground, equipped with Brewer spectrophotometers. The results on the total ozone were used in tendency studies, ozone forecasts, UV radiation, and for the characterization and calibration of satellite measurements.

The Canadian ozonization network was created in the mid-sixties, complete with four stations. At this very moment, probes are released at least once a week on six stations. All sampling results and ground measurements are available through the World Ozone and Ultraviolet Data Center (WOUDC), which Canada develops for the WMO.

The Stratospheric Ozone Observatory at Eureka (80°N, 86°O), also part of the WMO / UNEP for the detection of stratospheric change (NDSC) opened its doors in 1992. We can find lidars for measuring ozone and aerosol, Fourier transform spectrometers and other spectrometers used to measure the composition of the stratosphere. The Institute for Space and Terrestrial Science (ISTS) of Ontario, as well as the Meteorological Agency and Communications Research Laboratory of Japan both collaborate with Environment Canada in the achievement of these experiments.

The annual average value of the ozone in Canadian stations has decreased by approximately 6% in comparison with averages before 1980. This decrease reached its maximum in the spring and its minimum in the fall. The results of the probes show the bulk of this decline occur in the lower stratosphere.

## ST1 -Red

### THE DISTRIBUTION OF OZONE AND OZONE-DEPLETING SUBSTANCES IN THE ATMOSPHERE AND OBSERVED CHANGES

The principal characteristics of the climatology of the total ozone were discovered before 1930 by G.M.B. Dobson, and depends on the latitude and the season, as well as daily changes linked to meteorological conditions. Thanks to a number of modern techniques, we are now able to get a more complete climatology, extending from the earth's surface through all the levels of the stratosphere. We are also able to measure ozone distribution with altitude, as well as the total ozone amount, from the surface and through the use of various platforms, such as balloons, aircraft, rockets and satellites. In 1985, a severe depletion that occurs in Antarctica in the spring, known as the "ozone hole" was discovered; several years later, a depletion in the midlatitudes was observed. These last years have seen significant and increasing depletion in the Arctic during the spring. Already for several decades, we have measured atmospheric concentrations of substances that deplete the ozone, and now detect there conclusive indications of the effect of the Montreal Protocol.

Since the beginning of the 1960s, Canada has started to take measurements of the total ozone from five stations. Since 1992, Canada has had twelve surface stations equipped with Brewer spectrophotometers to observe the ozone. Data on the total ozone are used in trend studies, ozone and UV radiation forecasts, as well as for the characterization and calibration of satellite measurements.

The Canadian network for ozone surveying was established in the 1960s, with four stations. At present, probes are released at least once a week from six places. All data gathered from the probes and surface measurements are made available to the World Ozone and Ultraviolet Radiation Data Centre, which uses Canada for the WMO.

The Observatory for Stratospheric Ozone of Eureka (80° N, 86° W), which is part of the WMO/UNEP Network for Detection of Stratospheric Change (NDSC), was established in 1992. The Observatory has at its disposal measuring lidars for the ozone and aerosols, and Fourier transform spectrometers to measure the composition of the stratosphere. The Institute for Space and Terrestrial Science (ISTS) of Ontario, as well as the Meteorological Agency and Communications Research Laboratory in Japan collaborate with Environment Canada in the carrying out of its experiments.

The annual average measurements of the ozone in Canadian stations has decreased around 6% as compared to averages before 1980. This decrease reaches its lowest level in spring and in autumn. Probe measurements show that a large part of the decrease occurs in the lower stratosphere.

## ST1 - Green

### THE DISTRIBUTION OF OZONE AND OZONE-DEPLETING SUBSTANCES IN THE ATMOSPHERE AND OBSERVED CHANGES

The main characteristics of the total ozone climatology were discovered by G.M.B. Dobson before 1930. It consists of its dependence on the latitude and the season, as well as day-to-day changes related to weather conditions. The use of many new modern technologies has provided us with a more complete climatology, extending from the ground to all levels of the stratosphere. This means we can measure ozone distribution by altitude, as well as the total ozone amount, from the ground and other, diverse platforms, such as balloons, aircraft, rockets and satellites. The severe depletion that occurs in the spring in the Antarctic, the "ozone hole," was discovered in 1985; a few years later a midlatitude depletion was discovered. In recent years, we have noted a significant and growing depletion in the spring in the Antarctic. A number of decades have passed since we first measured the atmospheric concentrations of ozone-depleting substances, and there is now conclusive evidence of the effect of the Montreal Protocol.

Canada began measuring total ozone at five stations at the beginning of the 1960s; since 1992, there are twelve ground-based ozone observing stations, equipped with Brewer spectrophotometers. Total ozone data is used in studies of trends, ozone and UV predictions, and in the characterization and calibration of satellite measurements. The Canadian ozone-probing network was launched in the mid-sixties, with four stations. Currently, probes are released once a week from six areas. All the probe data and ground measurements are available from the World Ozone and Ultraviolet Radiation Data Centre, which Canada operates on behalf of the WMO.

The Eureka Stratospheric Ozone Observatory (80° N, 86° W), which is part of the WMO/UNEP Network for Detection of Stratospheric Change (NDSC), opened in 1992. It has lidars for measuring ozone and aerosol, Fourier transformation spectrometers and other spectrometers to measure the composition of the stratosphere. The Institute for Space and Terrestrial Science (ISTS) of Ontario, as well as the Meteorological Agency and Communications Research Laboratory of Japan are working with Environment Canada to conduct these experiments.

The average annual ozone values at the Canadian stations decreased by approximately 6% compared to the averages prior to 1980. The decrease was most significant in the spring and least significant in the fall. The probe measurements show that the bulk of this decline is occurring in the lower stratosphere.

## ST1 - Yellow

### THE DISTRIBUTION OF OZONE AND OZONE-DEPLETING SUBSTANCES IN THE ATMOSPHERE AND OBSERVED CHANGES

The main characteristics of total ozone climatology were discovered before 1930 by G.M.B. Dobson. It mainly concerned dependence on latitude and season, as well as day-to-day changes linked to meteorological conditions. We now have a more complete climatology at our disposal thanks to the use of several modern techniques, from the ground to all the levels of the stratosphere. Ozone distribution can be measured with altitude, as well as the ozone total ozone from the ground and from various platforms, such as balloons, aircraft, rockets, and satellites. The “hole in the ozone layer,” a severe depletion of ozone, which happens in the spring in the Antarctic, was discovered in 1985, and several years later depletions at mid latitudinal levels were discovered. In recent years, a significant and increasing loss in the Arctic during the spring has been noted. Atmospheric concentrations of ozone-depleting substances have been detected for several decades, and now convincing evidence of the Montreal Protocol effect has also been detected.

Canada began taking total ozone measurements from five stations in the early 1960s, and since 1992 has used twelve ground-level ozone observation stations equipped with Brewer spectrophotometers. The data on total ozone data are used in studies on trends, ozone and UV radiation forecasts, and for characterization and calibration of satellite-based measurements.

The Canadian network of ozone surveying was created in the mid-1960s, with four stations. Currently, probes are launched at least once a week to six various locations. All the survey and ground-level measurement data are available at the World Ozone and Ultraviolet Radiation Data Centre that Canada has been using for the WMO.

The Eureka Stratospheric Ozone Observatory (80° N, 86° O), part of the WMO/UNEP Network to Detect Stratospheric Changes (NDSC), opened its doors in 1992. There are ozone and aerosol measuring lidars, Fourier Transform Infra-Red (FTIR) spectrometers, and other types of spectrometers for measuring the composition of the stratosphere. The Ontario Institute for Space and Terrestrial Science (ISTS) as well as the Japan Meteorological Agency and Communications Research Laboratory have collaborated with Environment Canada to carry out these experiments.

Annual ozone averages measured at Canadian stations have decreased by 6% compared to averages reported in the mid-1980s. The decrease reached its maximum in the spring, and its minimum in the autumn. Additionally, the probe measurements show that the majority of the loss takes place in the upper stratosphere.

## ST2 - Blue

### Marine Ecosystems

The sensibility of physiological processes to solar UV-B's requires a function of biological weight (also called the spectrum of biological effects) for quantifying effective irradiance. The spectrum must be determined for a large number of biological functions and make estimations on their range of variation for each process that appear to be highly variable from one specie to the next.

The biological effects of UV-B have been detected tens of meters into the water column, and especially the inhibition of short-term in the phytoplankton in the Antarctic. As studies continue, we see quantitative estimations of inhibition converge in UV changes. Some toxic dino-flagellates show UV-photoprotective mechanisms which might give them a competitive edge in an enhanced-UV environment susceptible of bringing a predominant accretion of toxic or noxious algae.

Increases in UV-B could also favour species of phytoplankton that produce dimethyl sulphide, a reactive gas implicated in cloud formation and, hence, local climate and heat balance.

To interpret UV effects in Canadian waters, we must recognize the fundamental differences between the Antarctic and the Northern hemisphere waters, such as the Arctic. Basic research on Canadian marine ecosystems are thus essential.

### Materials

UV radiation has adverse effects on a number of materials on the outside. Each UV augmentation that touches the earth's surface will result degradation of the infrastructure, and so, higher repair and replacement costs.

Canadian research has examined polymeric UV effects on wood and paper, building materials, paints and coatings, as well as textiles and clothing, but largely targeted the evaluation of resistance to the radiation of materials on the outside and clothing fabric. UV-B damages synthetic polymers and other materials, but have yet to well understand the mechanisms at the molecular level, and the combination of short and long wavelength radiation with other environmental variables that add further complexity to the issue. The color of pulp and paper products and bleached paper issued from low cost processes is altered by UV radiation. Canadian researchers have made much progress in the understanding of this phenomenon. The capacity of reducing this color alteration creates a large market expansion for this category of products.

Non plastic building materials, like roofing membranes and outdoor sealants are actually the object of research in their resistance to UV, but not precisely in the accretion radiation context linked to the ozone.

## ST2 - Red

### Marine Ecosystems

The sensitivity of physiological processes to UV-B rays requires biological weighting (also called biological effects spectrum) to quantize efficient irradiance. Spectrums for a larger number of biological functions must be selected and the variable ranges for each process, which appear to be highly variable between species, estimated.

We have detected the biological effects of UV-B rays up to tens of metres into the water column, in particular, the inhibition of short-term photosynthesis in phytoplankton in Antarctica. As research continues, we see quantitative estimates of the inhibition of photosynthesis by UV radiation converge.

Some toxic dinoflagellates show UV-photoprotective mechanisms which might give them an advantage capable of creating an enhanced predominance of harmful or toxic algae.

Increases in UV-B could also favour species of phytoplankton that produce dimethyl sulphide, a gas implicated in cloud formation and, hence, climate change.

In order to interpret the effects of UV radiation on Canadian waters, we must recognize the fundamental differences between Antarctica and the bodies of water in the Northern Hemisphere including the Arctic. It is therefore essential to research Canadian marine ecosystems.

### Materials

UV radiation has significantly harmful effects on a number of materials used in the environment. Any increase in UV radiation that comes into contact with the earth's surface will involve deterioration in its infrastructure and, therefore, high repair and replacement costs.

Canadian researchers have examined the effects of UV radiation on building materials such as polymers, wood and paper, on paints and coatings, as well as on textiles and clothing. However, we further aimed to evaluate the radiation resistance of materials used in space and in materials used for clothing.

UV-B rays damage synthetic polymers and other materials; however, we still do not understand mechanisms at the molecular level well enough, and the combination of the impact of short- and long-wave radiation with other environmental variables continues to complicate the situation.

The colouring from bleached pulp and paper products resulting from cheap procedures is altered by UV radiation. Canadian researchers have made significant progress in our understanding of this phenomenon. The ability to reduce this colour alteration could lead to large market expansion in this category of products.

Non-plastic building materials, such as membrane roofing and outdoor sealants as to their resistance to UV radiation are currently the subject of research, but not exactly in the context of increased irradiation linked to the ozone.

## ST2 - Green

### Marine Ecosystems

The sensitivity of physiological processes to UVB rays requires biological balancing (also called biological action spectrum) to quantify sufficient irradiance. Spectrums must be determined for a larger number of biological functions, and variation ranges must be estimated for each process, which appear highly variable from one species to another.

The biological effects of UVB rays have been detected up to several tens of metres in the water column, particularly the inhibition of short-term photosynthesis in Antarctic phytoplankton. Continuing research has shown the convergence of quantitative estimations on the inhibition of photosynthesis due to UV radiation.

Some toxic dinoflagellates possess UV-photoprotective mechanisms that might give them a competitive edge that may cause a predominant accrual of toxic or noxious algae.

Increases in UV-B levels could favour species of phytoplankton that produce dimethyl sulfide, a reactive gas involved in cloud formation and, therefore, climate change.

To determine UV effects on Canadian waters, the fundamental differences between the Antarctic and the waters of the Northern hemisphere, including the Arctic, must be assessed. Basic research must be done on Canada's marine ecosystems.

### Material

UV radiation has significant harmful effects on many materials used outside. All increases in UV rays reaching the earth's surface cause infrastructure to degrade and costs of repair and replacement to rise.

Canadian studies examined UV effects on polymers, wood and paper, building materials, paintings and coatings, as well as textiles and clothing; however, the primary focus was evaluating the radiation resistance of materials used in space and clothing fabrics.

UV-B rays damage synthetic polymers and other materials, but molecular mechanisms are not yet completely understood. The combination of the impact of short and long wavelength radiation with other environmental variables further complicates the situation.

UV radiation alters the colour of cheaply processed bleached pulp and paper products. Canadian researchers have made significant progress in understanding this phenomenon. The ability to reduce this colour alteration could lead to a considerable expansion in the market for these types of products.

The resistance of non-plastic building materials, such as membrane roof systems and exterior sealants, to UV rays is currently being studied, but not specifically within the context of an accumulated radiation linked to the ozone.

## ST2 - Yellow

### Marine Ecosystems

- The sensibility of solar UVB physiological processes requires a function of biological weighting (also called biological effects spectrum) to quantify the efficient irradiance. The spectrums must be determined for a larger number of biological function and to estimate their variation range for each process that seemly highly variable from one species to the next.
- Biological effects of UVB were detected for several dozen meters in the water column, as were the inhibition of short-term photosynthesis of phytoplankton in the Antarctic. As long as the research continues, we envision the convergence of quantitative estimations of the inhibition of photosynthesis by UV radiation.
- Some toxic dinoflagellates show photoprotective mechanisms against UV radiation which might give them a competitive edge susceptible of causing an increased predominance of toxic or dangerous algae.
- The increase of UVB levels increase the amount of phytoplankton species that produce dimethyl sulfide, a gas involved in cloud formation, and therefore a contributor to climate change.
- To interpret UV effects in Canadian water, the fundamental differences between the Antarctic and the water of the Northern Hemisphere (which includes the Arctic) must be noted. For that reason, basic research on marine ecosystems in Canada is essential.

### Equipment

UV radiation causes significant harmful effects on a number of materials used outside. Any increase in UV radiation that touches the earth's surface will cause infrastructure damage, and therefore, increased repair and replacement costs.

- Canadian research had examined UV effects on polymers, wood, paper, building materials, paint, and coatings, as well as textiles and clothing. However emphasis has been put on the evaluation of the materials' resistance to radiation in space and in materials used in clothing.
- UVB radiation damages synthetic polymers as well as other materials, but the mechanisms on the molecular level are still not well understood, nor are the combination of impacts of short and long wavelength radiation. The situation is complicated by other environmental variables.
- The colour of bleached pulp and paper products manufactured from economical procedures is altered by UV radiation. Canadian researchers have made significant progress in understanding this phenomenon. The ability to reduce this colour alteration could lead to a large expansion of the market for this category of products.
- The non-plastic building materials, such as waterproof coatings and outdoor sealing products are currently the subject of research on their resistance to UV radiation, but not precisely in the context of an increased irradiation linked to the ozone.