© 2012 Errol Marinus O’Neill
THE EFFECT OF ONLINE TRANSLATORS ON L2 WRITING IN FRENCH

BY

ERROL MARINUS O’NEILL

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in French in the Graduate College of the University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Professor Douglas A. Kibbee, Chair
Assistant Professor Randall W. Sadler, Director of Research
Associate Professor Laurence Mall
Assistant Professor Jessica L. Sturm, Purdue University
Online translation (OT) sites such as Free Translation and Babel Fish are freely available to the general public and purported to convert inputted text, from single words to entire paragraphs, instantly from one language to another. Discussion in the literature about OT for second-language (L2) acquisition has generally focused either on the pedagogical and ethical issues surrounding online translators or on anecdotal discussions of errors researchers have encountered on such tools. This dissertation takes a different approach by investigating the quantitative and qualitative effects of OT use on L2 composition writing.

A study conducted among 32 third- and fourth-semester L2 French students compared the results of three conditions: participants who were allowed to use an online translator for writing tasks after attending a training session on OT (+translator, +training); participants who were permitted to use OT for their writing, but who had received no prior training (+translator, -training); and a control group whose participants had no training and were not allowed to use OT (-translator, -training). Each of the 128 compositions collected was scored by raters and evaluated on six features: overall comprehensibility, content, spelling, syntax, remaining grammar, and vocabulary. Additionally, raters were asked to judge whether or not, in their estimation, each composition’s writer had used an online translator.

The results showed that not only did OT use not have a negative effect on the mean scores of those who used a translator during the study, but the online translator groups performed as well as, if not better than, the control group on the experimental writing...
tasks both overall and on specific linguistic features. Specifically, the global scores of the translation group that had received training (+translator +training) were in fact statistically significantly higher ($p < 0.05$) on the second of two experimental writing tasks as compared to the control group (-translator -training). Additionally, on Tasks One and Two, one or both translator groups significantly outperformed the control group on four of the six components evaluated (overall comprehensibility, content, spelling, and remaining grammar). Raters were able to identify correctly, to a statistically-significant level, whether or not online translation was used for compositions; there were however cases where one or both raters scoring a given composition mistakenly suspected OT use, as well as some OT-aided compositions that escaped detection by raters.

This dissertation analyzes global and component scores of compositions written with the use of an online translator and those written without such aid; presents and discusses samples of student writing in light of rater comments and participant self-reports; and explores the implications of the study’s results for L2 learning and instruction.
ACKNOWLEDGMENTS

A person’s life and career are shaped by thousands of people. It would be impossible to list everyone who has helped me get to the point of successfully completing my dissertation. For this reason, I would like first of all thank everyone who has helped me in all ways big and small but who is not mentioned below.

I wish to thank specifically my committee for all they have done to get me to this point. I would not have been able to complete my Ph.D. were it not for Dr. Randall Sadler, who generously agreed to become my thesis advisor. Despite the logistics involved in advising someone hundreds of miles away, he was willing to see my dissertation to completion and offered important suggestions that allowed me to develop my discussion. Dr. Doug Kibbee has been an advisor to me in a number of capacities since I started my M.A. studies under his direction. He kindly served as chair of my doctoral committee for my prelims through my defense, giving generously of his time and advice both before and since his retirement. Dr. Laurence Mall has always been encouraging and supportive, sharing her experiences and important insights on online translation, passing along interesting information and anecdotes, and accepting to participate on my committee and in my defense despite her many obligations. Dr. Jessica Sturm has played a number of different roles through the years I have known her; I will always be grateful to her for all of the advice, support, encouragement, and friendship she has given me through the various stages of my journey.

I would like to thank Dr. Alice Omaggio Hadley, who initially inspired me to go into SLATE and who served on my prelim committee. I have learned a great deal from her as my instructor, supervisor, and mentor. It is in a very large measure thanks to her
example of pursuing excellence in L2 instruction and coordination, her unequaled
dedication, and her confidence in me that I have been able to accomplish what I have.
Thanks are also due to Dr. Peter Golato for serving as the head of my prelim committee
and for his role in helping get my dissertation and study off the ground, as well as Dr. Liz
Martin for her insights both as a member of my prelim committee and as an instructor.

The study itself would not have been possible without the extremely generous help,
work, and patience of all the TAs who proctored tasks, who allowed their classes to be
interrupted for five weeks, and/or who rated the compositions: Sarah Bradley, Tricia
Dimit, Leila Ennaii, Jennifer Folkert, Michael Foster, Stéphanie Gaillard, Jenelle Grant,
Sarah Hauser, Eleonore Le Corvaisier, Jennifer Lonigro, Jessica Nicholas, Daniel
Maroun, Liz Martinez, Rebecca Nevers, Nathan Owens, Anna Philipponneau, Nadège
Saint-Maxent, and Awa Sarr. Without you, my study would not have been possible.

There are many other professors and TAs over the years that I’ve had the pleasure
to work with and know; I would like to thank them all for the role they have played in my
academic and personal life. I owe a great deal to the Department of French, Graduate
College, and the University for all of their support, both academic and financial,
throughout my studies. I am thankful as well to Dr. Patricia Phillips-Batoma for allowing
me to conduct my research while she was the acting Director of Basic Language
Instruction. I would also like to thank the students who participated in my study for all
their work and time both during and outside of class.

I extend a heartfelt thanks to my colleagues past and present in the French section
at the University of Memphis, including Drs. Ralph Albanese, Will Thompson, Denis
Grélé, Sue Johnson, Susan Silver, William Hendel, Matthew Moyle, and Robert Fagley,
as well as Christiane Gilbert and Jan Starczewski, for being supportive as I completed my
dissertation while serving as a full-time faculty member. I would also like to thank Lei
Wang very much for her time and willingness to help me with statistical analyses.

There are a number of people who have touched my life personally — many of
whom were already cited above. Sadly, my father (who died in 1999 after a long battle
with cancer), and my mother (who died quite suddenly on Bastille Day in 2010), never
got to see me get my doctorate, but it is thanks to their love and encouragement that I
learned the importance of pursuing my education and following my dreams; they will
always be a part of who I am. I would also like to thank Don, who as my stepdad has
treated me like a son, for all the love he has shown and everything he has done for me
and our whole family. My brother Treesong and sister Val, through love, humor, and
conversations about both the most essential and silliest parts of life, have been the best
siblings a brother could ever hope to have. Thanks also to the rest of my family, on both
my and my wife’s sides, for all the caring and support they have shown over the years.

I owe so much to Dr. Todd J. Hamilton, my best friend through thick and thin. I
have always been able to count on him for matters both personal and professional. His
advice and companionship have helped me find strength and confidence when I most
needed it. I will forever be in his debt and will always treasure our friendship. And last
but not least, I reserve my final thanks to the love of my life, Melanie, who has put up
with all the craziness that finishing a dissertation involves, and then some; who has stood
beside me all the way, no matter what life has brought our way; who has comforted and
encouraged me in all that I do; who has given me her unconditional love and support, and
made me feel complete. I dedicate this dissertation, and my life, to her.
# TABLE OF CONTENTS

## CHAPTER 1: INTRODUCTION
1.0 General introduction to the topic ................................................................. 1

## CHAPTER 2: SELECTED REVIEW OF THE LITERATURE
2.0 Introduction .................................................................................................. 5
2.1 Reading and its relationship to writing ......................................................... 5
   2.1.1 Research investigating L1 reading and writing................................. 6
   2.1.2 Research investigating L2 reading and writing............................... 9
2.2 Paper-based L2 writing ............................................................................. 14
2.3 Computer-based writing .......................................................................... 21
2.4 Machine translation, or non Web-based MT ............................................ 31
2.5 Online translation and other writing tools ............................................. 43
2.6 Conclusion ................................................................................................. 62

## CHAPTER 3: METHODOLOGY
3.0 Introduction ................................................................................................. 67
3.1 Participants ................................................................................................ 68
3.2 Experimental design .............................................................................. 69
3.3 Materials and procedure .................................................................... 70
3.4 Rating ....................................................................................................... 77
3.5 Conclusion.................................................................................................. 79

## CHAPTER 4: RESULTS
4.0 Introduction ................................................................................................ 81
4.1 CLEP Pretest results ............................................................................. 81
4.2 Interrater reliability ................................................................................ 83
4.3 Written Pretest results .......................................................................... 85
   4.3.1 Written Pretest subscores ............................................................ 86
4.4 Task One results .................................................................................... 88
   4.4.1 Task One Content subscore ....................................................... 89
   4.4.2 Task One Remaining Grammar subscore ................................... 91
   4.4.3 Task One other subscores and summary .................................. 92
4.5 Task Two results .................................................................................... 93
   4.5.1 Task Two Overall Comprehensibility subscore ........................ 95
   4.5.2 Task Two Content subscore ...................................................... 96
   4.5.3 Task Two Spelling and Accents subscore .................................. 97
   4.5.4 Task Two Remaining Grammar subscore ................................ 98
   4.5.5 Task Two other subscores and summary .................................. 100
4.6 Posttest results ....................................................................................... 101
   4.6.1 Posttest subscores .................................................................... 102
4.7 Rater judgements on translator usage .................................................. 104
   4.7.1 Rater judgements for each composition .................................... 105
4.8 Participant self-reports ......................................................................... 109
4.8.1 Background Questionnaire .............................................................. 109
  4.8.1.1 Online translation to or from English ...................................... 111
  4.8.1.2 Online translation to or from French ...................................... 112
  4.8.1.3 Online translation to or from French during classtime ................ 112
  4.8.1.4 Online translation to or from French outside of class ................. 113
  4.8.1.5 Prior online translation to or from other languages .................. 114
  4.8.2 Translator use during the study .............................................. 115
  4.9 Conclusion ................................................................................. 115

CHAPTER 5: DISCUSSION
  5.0 Introduction ................................................................................. 116
  5.1 Effect of translator use on overall scores ...................................... 116
  5.2 Effect of translator use on categories .......................................... 119
    5.2.1 Effect of translator use on Overall Comprehensibility .............. 120
    5.2.2 Effect of translator use on Content ....................................... 122
    5.2.3 Effect of translator use on Spelling and Accents ..................... 123
    5.2.4 Effect of translator use on Syntax ........................................ 125
    5.2.5 Effect of translator use on Remaining Grammar ..................... 126
    5.2.6 Effect of translator use on Vocabulary ................................... 128
    5.2.7 Conclusion ............................................................................. 130
  5.3 Rater judgments on translation use ........................................... 133
  5.4 Participant self-reports ................................................................ 136
    5.4.1 Self-reports on cultural lesson .............................................. 137
    5.4.2 Self-reports on translator training ........................................ 140
    5.4.3 Self-reports on translation use ............................................. 145
    5.4.4 Self-reports on effects of online translator use ....................... 148
  5.5 Samples of participant writing and self-reports, rater judgments and
    comments ....................................................................................... 153
    5.5.1 Compositions with Type II errors ........................................... 155
      5.5.1.1 Compositions with Type II errors by only one rater ............ 155
      5.5.1.2 Compositions with Type II errors by both raters ............. 166
    5.5.2 Compositions with Type I errors .......................................... 176
      5.5.2.1 Compositions with Type I errors for only one rater ......... 177
      5.5.2.2 Compositions with Type I errors by both raters ............ 181
    5.5.3 Compositions correctly judged .............................................. 187
      5.5.3.1 Compositions with translator use correctly judged ........... 187
      5.5.3.2 Compositions with translator non-use correctly judged .... 196
    5.5.4 Summary of writing samples and comments ............................ 202
  5.6 Conclusion .................................................................................. 202

CHAPTER 6: CONCLUSIONS
  6.0 Presentation of chapter .............................................................. 206
  6.1 Limitations ................................................................................ 206
  6.2 Future research ......................................................................... 211
  6.3 Implications and conclusions .................................................... 214
BIBLIOGRAPHY .................................................................................................................. 217
APPENDIX A: CONSENT FORM – ONLINE TRANSLATION STUDY .223
APPENDIX B: BACKGROUND QUESTIONNAIRE .......................................................... 224
APPENDIX C: READING PRETEST (EXCERPT) ............................................................ 229
APPENDIX D: WRITTEN PRETEST .............................................................................. 230
APPENDIX E: TASK ONE ............................................................................................... 231
APPENDIX F: TASK TWO ............................................................................................... 233
APPENDIX G: POSTTEST ................................................................................................. 235
APPENDIX H: PRE-WRITING ACTIVITY I ................................................................. 236
APPENDIX I: PRE-WRITING ACTIVITY II ............................................................... 238
APPENDIX J: TYPING ACCENTS FOR FRENCH CLASS ............................................ 240
APPENDIX K: TRAINING SESSION – LESSON OUTLINE ...................................... 241
APPENDIX L: ONLINE TRANSLATION – SOME POSSIBLE STRENGTHS AND WEAKNESSES ........................................................................................................ 245
APPENDIX M: CULTURAL LESSON ............................................................................ 246
APPENDIX N: SELF-REPORT EXIT QUESTIONNAIRE ................................................ 251
APPENDIX O: DEBRIEFING ......................................................................................... 254
APPENDIX P: SCORING CRITERIA FOR COMPOSITIONS ........................................ 255
Chapter 1

Introduction

1.0 General introduction to the topic

The field of second-language (L2) writing has been evolving over the past several decades due to the introduction, and increasing availability, of personal computers in the laboratories, classrooms, and homes of L2 students. Traditionally, students in L2 courses have had to write compositions to foster and demonstrate their fluency in the target language using pen and paper. Students have long had additional tools available to them such as textbooks, grammars, dictionaries, and other paper-based resources to assist in the process of writing. The advent of the computer age has brought about a change both in the media used for writing as well as the tools available to aid in the writing process. Word processing software allows the student to compose directly on the computer, while resources on the computer such as electronic dictionaries and a seemingly endless amount of material available via the Internet offer more resources to the student than ever before at little or no cost. Students still have access to paper-based resources, but the prevalence of and easy access to computers and the materials that can be used thereon have increasingly made computers dominant in writing done for many L2 courses.

At first glance, combining the medium of writing with the resources used to aid in writing would seem to be beneficial: editing text is a matter of cutting, copying, and pasting; more tools are at the students’ disposal, and so on. There are however several risks involved in using computers for L2 writing. One such risk is directly related to one of the benefits listed above, namely the plethora of available resources. While some resources may serve to help students in their writing, students may discover and attempt
to use tools that may not be beneficial to their writing, at least if not understood correctly. One type of tool that students are increasingly aware of is online translators. There are dozens of websites that claim to offer near-instantaneous translation of text from one language into another. Visitors to the website need only type in or paste some text in one language, and within seconds they receive text purported to be a translation of the original in a second language. The temptation is great for students to simply put words, phrases, or entire paragraphs into the online translator to receive in seconds what it may have taken minutes or hours to compose in the target language on their own without using a translator. Without instruction or advice from teachers, beginning and even more advanced students may not be aware of whether they can or cannot use this tool to aid in their composition writing, and if so in what ways it might be helpful or detrimental. Instructors themselves, either unaware of their students’ use of translators or unsure what impact their use might have on student writing, may not have a policy on this new technology nor know what a sound policy would be.

As we will see in the review of the current literature in L2 writing and related fields, the issue of online translation has been left largely unaddressed, apart from a small collection of articles generally advising against the use of online translators. These papers generally offer brief examples of poorly-translated words or phrases or advice on how to demonstrate to students some potential pitfalls of using translators to create text in the L2. No quantitative studies appear to have been published which would confirm or deny any effect of online translators on L2 learner writing, much less provide definitive substantiation for encouraging, discouraging, or disallowing their use on writing assignments for the foreign language classroom.
This dissertation proposes to address the issue of the use of online translators in L2 writing. Current research and theory on L2 proficiency, including the possible relationship between reading and writing levels, the writing process in general, and more specifically L2 writing, are presented. A theoretical framework useful for considering L2 writing specifically in view of online translators is outlined. Under this framework, the results of a study conducted to measure the effects of online translation on the writing of Novice-High and Intermediate-Low French L2 learners is presented and discussed.

More specifically, the study compared several aspects of student writing, including overall scores on compositions written with and without online translators; any difference in translator-assisted writing on specific features of L2 writing, such as content, vocabulary, grammar, and overall comprehensibility of compositions; any effect of training on students using online translators versus those who received no prior training; and a look at whether foreign-language teacher raters can reliably distinguish between those compositions that were written with the aid of an online translator and those that were not. Students’ backgrounds and abilities in terms of reading and previous technology usage are discussed to find any possible correlation between L2 reading and student performance in L2 writing, both with and without the aid of online translators. In view of the literature review, the study’s results, and an analysis thereof, the aim of this document is to provide those in the field of second language teaching a theoretical and quantitative basis for making informed decisions concerning the use of online translators.

The following organization is used to accomplish this purpose. Chapter 2 presents a selected thematic review of the literature in the fields of L2 reading and writing as well as professional translation, specifically as concerns the relationship between reading and
writing in L2 learners; paper-based writing for L2 learners; computer-based writing for L2 learners; offline machine translation for professional human translators; and electronic resources including online translators used in writing by L2 learners. Issues raised by the literature review are discussed, pointing to the motivation for further analysis through a quantitative study of online translation use among the learners in question, novice-high to intermediate-low L2 learners, both with and without training in the use of online translators. Chapter 3 provides the methodology used in the study, including an explanation of how the methods chosen specifically addressed the issue of the effect of online translators on L2 writing. A detailed report of the data obtained in the study is given in Chapter 4. Chapter 5 presents a discussion of the results. Finally, Chapter 6 includes limitations of the present study, suggestions for future research, as well as implications for the use of online translators for student writing both inside and outside the L2 classroom.
Chapter 2

Selected Review of the Literature

2.0 Introduction

The purpose of this chapter is to explore the literature of various fields related to online language translation and its diverse aspects. In particular, five distinct areas were identified as relevant to the study at hand. They are treated here in order of increasing pertinence to the use of online translators in L2 writing. The first area involves the relationship between reading and writing in the L1 and L2, from both a theoretical and a performance standpoint; the second concerns traditional, paper-based L2 writing; the third entails computer-based L2 writing and instructional programs available for use by L2 learners on the computer; the fourth involves a specific type of computer program, a machine translator as most often used in the fields of professional translation or translation research; the final area concerns specifically the use of electronic or online tools such as dictionaries and translators, as pertains to the L2 learner.

2.1 Reading and its relationship to writing

Research in both first and second language acquisition has looked at the possible links between reading and writing in terms of theory and performance. Historically, reading and writing have been treated as unrelated by many teachers and researchers. It is only fairly recently that the relation between the two has been explored. If there is indeed a correlation between reading and writing abilities, particularly if one’s level of reading comprehension either affects or predicts one’s level in writing, then looking at the
reading proficiency of L2 learners may have some implications for the use of online translators in L2 writing.

Studies clearly linking reading and writing do not abound in the L1 or L2 literature. According to Stotsky (1983), in one of the earliest comprehensive looks at the issue, “very little research” had looked at the influence of writing or writing instruction on reading, or at the effect of reading or reading instruction on writing (p. 627). Those studies that had been done tended to look at correlations between reading and writing, rather than exploring the influence one might have on the other. These studies have generally found a relationship between reading and writing ability.

2.1.1 Research investigating L1 reading and writing

Pearson and Tierney’s Reading/Writing Model (Pearson and Tierney 1984) advances the notion that reading skills are intricately related to writing skills. Authors write with the intention that their audience will create meaning as they read. Reading involves four interactive roles: planner, composer, editor, and monitor. In essence, readers actually construct their own text as they read, in a sense “writing” for themselves the text they expect to see using their prior knowledge (planner), ability to fill in coherent details (composer), judgment of the accuracy of their interpretations (editor), and control over which of these three processes should be used at any given time (monitor). Pearson and Tierney argue that the same processes that occur when reading are also used when writing for an audience other than oneself, pointing to a close relationship between the two.
Felland (1980) found that in a study of 456 L1 English students at randomly-selected high schools, those judged to be better writers read more books on average (based on a self-report questionnaire) than other writers. Loban (1963), in studying elementary L1 English students, summed up his findings by saying that “those who read well also write well; those who read poorly also write poorly” (p. 75). However, a closer look at the data indicates that the relationship is not as clear-cut as this statement might lead one to believe. For example, 30% of the fourth-grade readers who were judged lowest in reading ability nonetheless fell into the group of good or superior writers. The reasons for this lack of correspondence could range from a difference in difficulty in types of texts read and written (Reichelt, 2001 among others suggests that narrative texts are more difficult than descriptive texts) to differing skill sets needed for reading (e.g. textual interpretation) and writing (e.g. use of transitional devices). Other studies, such as Shanahan (1984), found that the assumed relationship between reading and writing may not be as strong as some suspect.

Shanahan and Lomax (1988) considered three alternative theoretical models of the relationship between L1 reading and writing. Under the reading-to-writing model, it is theorized that as learners develop linguistically, knowledge moves from reading to writing but not from writing to reading. The writing-to-reading model predicts the opposite; writing knowledge is transferred to reading as a learner develops, and not vice versa. A final model, the interactive model, allows for the movement of knowledge in both directions; learners’ reading knowledge can contribute to their developing writing knowledge as well as the reverse. For this last model, the influence is believed to go across discourse levels; for example, knowledge at a given level of discourse in reading
will contribute to the next higher level of discourse in writing. An instance of this would be that learners could use their spelling knowledge of a given language (a knowledge considered by the authors to be related to writing) to understand unknown words encountered while reading, thus improving reading comprehension. In turn, vocabulary knowledge (considered to be related to reading) could improve learners’ writing by allowing for increased diversity in word choice as a writer.

In order to test the three models, the authors collected data from 256 second-grade and 251 fifth-grade native speakers of English through four different tests aimed at measuring phonetic analysis, reading comprehension, vocabulary, and spelling ability as well as a story-writing task analyzed for T-unit length, vocabulary diversity, and various grammatical features. Results of a linear structural relationship (LISREL) analysis using latent and indicator variables showed that of the three models, the interactive model best explained the data for both beginning and proficient readers. Knowledge of reading-related variables (e.g. word analysis) predicted knowledge of writing-related variables (e.g. spelling) at the next higher level of discourse, and vice versa. While the overall results indicated that reading and writing are interrelated, with knowledge in one relating to knowledge in the other, it is important to note that the study did not analyze the relationship between individual variables in depth since many individual results were not statistically significant.

DeVries (1970) conducted a study among fifth-grade students to investigate whether increasing reading while decreasing writing could benefit student performance on compositions. All students were given a pre- and posttest, consisting of an expository essay to measure their progress after nine weeks. Students who were assigned fewer
compositions and more readings actually performed better in terms of written proficiency than students who had fewer readings but more compositions to write.

### 2.1.2 Research investigating L2 reading and writing

While studies into the relationship of L1 reading and writing may provide a general framework for a discussion of L2 reading and writing, it is unclear whether the literature about the development of reading and writing in the L1 learner can be applied to L2 learners. Higgs (1979) looked at the varying, and sometimes irreconcilable, theories and methodologies on second language acquisition by discussing the basics of language learning in both one’s native language as a child and a second language as an adult. Since experts are still unable to define language in a rigorous manner, nor say for certain how children acquire their first language, it is difficult to claim one theory or methodology of second language learning is superior or mutually exclusive to another.

Looking at the roots of language learning, that of a first language by children, Higgs observes that the apparent recording of lexical items in children occurs naturally, with children defining object names they learn (e.g. “cow”) with what the object does (e.g. “give milk”). Utterances are given and understood literally by children based on these recorded lexical items, and it is not until later that simple identification develops into relationships between objects and eventually to figurative meaning. Higgs argues that adult learners of a second language revert to a similar stage of rudimentary language competence, naturally associating a given lexical item in a foreign language with one exact equivalent in their native tongue. To explain this, Higgs puts forward the Lexical
Analog Hypothesis (LAH), the naive idea among second language learners that a foreign language is the same as the native language except with different, but analog words.

According to Higgs, teachers and materials often tend to reinforce LAH early on by stressing similarities between the L1 and L2 languages, allowing students to be drawn to this false hypothesis that they already naturally gravitate towards. To combat this the author suggests referring to ideas as opposed to lexical items as soon as possible: for example, instead of asking a student how to say “hi” in the foreign language, asking them how friends greet each other in the target language. This strategy will help learners understand that language acquisition is not a simple relexification of the language they already know, but that all languages have their distinct ways (grammar) of expressing meaning. For Higgs, in order to ensure success among language learners, it must be shown that languages are different ways of conveying meaning and that one language or grammar is not more logical than the other. Higgs suggests the use of examples in the native language to help break the learner’s perception that language is simply a logical string of lexical items and to show that every language consists of an independent and arbitrary system of grammar which is used to express meaning.

In a similar vein to Higgs (1979) and his Lexical Analog Hypothesis, Bland, Noblitt, Armington, and Gay (1990) identify a beginning learner’s attempt at direct translation of words from L1 to L2 as an indication of what they call the Naive Lexical Hypothesis. Students of a foreign language initially attempt to find a one-to-one correspondence between words in their first and second languages. As this hypothesis is proven wrong over time, learners modify it gradually until they become aware at more advanced stages of the independence of the two language systems and begin to
manipulate the target language in order to create meaning instead of constantly referring to the L1. A consideration of these two hypotheses is particularly important for our purposes since some learners may expect individual words or grammatical structures to translate directly into the L2.

In terms of the relationship between reading and writing between the L1 and L2, Janopoulo (1996) studied 79 incoming foreign graduate students who wrote a composition topic in English as part of their orientation. Participants were also sent a self-report questionnaire to determine their reading habits in both their L1 and their L2. His results found that those who often read for pleasure in English (their L2) tended to be better writers than those who did less or no pleasure reading at all in English. However, no significant correlation between L1 pleasure reading and L2 writing was found.

These findings may in part corroborate a claim by Krashen (1993) that L1 reading and writing skills do not necessarily transfer to skills in one’s L2. Krashen (1984) posits that a correlation between L1 reading and writing skills does not necessarily transfer to the L2. In view of this, Krashen (1993) advances a “reading hypothesis.” According to this hypothesis, free reading has a positive effect on a variety of abilities, including not only reading comprehension but also writing. The reverse however is not true: writing does not (or at least not necessarily) influence reading ability or even writing. “According to much common wisdom, we learn to write by actually writing. The reading hypothesis asserts that this is not true, at least as far as style is concerned” (p. 73).

Four articles in the literature in particular would appear to confirm this. Elley (1991) reports on three studies done in Singapore among nearly 3,000 L2 learners of English, aged six to nine years old. Those participants who followed a program of shared
book experience and free reading performed better on several measures, including reading comprehension and writing, than those who did not follow the program.

Bayliss and Raymond (2004) looked at Canadian L1 and L2 speakers of French studying law at the university. A variety of factors were examined to discover any relationship to student GPA. The variance of several scores in relation to each other was also examined, including L1 and L2 reading comprehension scores and composition scores. In terms of variance, no strong relationship was found between reading comprehension and writing for L1 speakers of French; however for the L2 students the shared variance between reading comprehension and composition scores was nearly 45%. According to the authors, “any predictive power for writing ability in the case of anglophone students is accounted for by the Reading Comprehension score alone” (p. 44).

Only partially aligning with Krashen (1993) is a study by Carson, Carrell, Silberstein, Kroll, and Kuehn (2001) which explored the relationship between L1 reading and L1 writing, L1 writing and L2 writing, L1 reading and L2 reading, and L2 reading and L1 writing among Chinese and Japanese learners of English. Significant correlations were found for all of these relationships except between L1 writing and L2 writing among Chinese speakers learning English. This study suggests that both L1 and L2 reading skills are related to L2 writing, while presenting a mixed message concerning how L1 and L2 writing are linked, indicating that more investigation would be needed to determine whether the participants’ L1 or some other factor accounts for the diverging results.

Other studies however seem to contradict outright Krashen’s reading hypothesis. Hedgcock and Atkinson (1993) compared two studies (Atkinson and Hedgcock, 1990 and
Hedgcock, 1989) investigating various factors with respect to L1 and L2 reading and writing. In particular, the authors examined the reading habits of L1 and L2 university-level speakers of English, as determined through a self-report questionnaire, and their expository writing proficiency, as judged through a holistically-scored essay. The results of the L1 study showed that there was a statistically-significant relationship between all 14 reading habit variables studied and writing performance in L1 participants. A stepwise regression analysis specifically showed a significant relationship between a higher frequency of fiction and textbook reading and higher scores on the essay. The results of the L2 study however showed no significant relationship between writing scores and textbook or fiction reading in the L1 or the L2. The authors conclude that either exposure to L1 and L2 texts has little or no impact on L2 writing proficiency, or that their methodology simply did not detect it. The main apparent difference between the two studies was the fact that the L1 study examined scores on a final-exam essay, whereas the L2 study looked at results of a placement-exam essay. It is uncertain whether other, extralinguistic factors such as the varying stresses experienced while taking a placement or final exam or writers’ varying L1s (not mentioned in the study) may have played a role in the results.

As can be seen from the above discussion, there is no consensus currently in the literature about the relationship of L2 writing to L1 or L2 reading. Certain studies do however indicate a relationship might exist between reading and writing, based on L1-only research as well as that studying L1 and L2 reading’s possible relationship with L2 writing. This relationship may be bidirectional (as Shanahan and Lomax, 1988 and others would indicate), unidirectional (as Krashen, 1993 and others would suggest), or may not
exist (as Hedgcock and Atkinson, 1993 would seem to indicate). Since no clear answer exists at present, a look at the relationship between reading proficiency and writing as part of the study presented in this dissertation was deemed helpful towards ensuring that whatever additional effect reading proficiency might have on writers in the various conditions was taken into account.

2.2 Paper-based L2 writing

A look at articles specifically concerning writing provides some needed background relating to L2 writing in general before investigating the effects of online tools on compositions in particular. As pointed out by Blake (1998), the fact that most L2 learners are adults who are already proficient in reading and writing in at least one other language (their native language[s] learned as a child), may either aid or hinder students attempting to perform these tasks in a new language since transfer of lexical, grammatical, or other features may occur. In addition, one’s L1 is generally learned primarily through aural and oral means well before reading and writing are initiated, whereas typical L2 learners do not yet have a good level mastery of the spoken language prior to commencing reading and writing.

Zamel (1982) looked into what L2 learners actually experience while writing in a foreign language. She interviewed eight ESL writers with different L1s. All of the participants had finished taking ESL writing courses. The author does not provide statistics or much detail about the procedure itself, but mentions three findings which will prove important to the current discussion.
First, students reported rarely or never using outlines; instead they reported looking back periodically, and pausing to reflect on, what had already been written in order to better develop and plan. Second, Zamel found that linguistic problems (lexical, syntactic, morphological) were not considered by any of the eight participants to be important obstacles. Participants devised ways to avoid letting the problems stop the initial writing process, such as leaving a word blank, circling words about which they weren’t sure, or putting the word down in the native, not the target, language. Lastly, and perhaps most importantly for the current discussion, the best writer (based on the author’s judgment) admitted to writing compositions in her native language first and then translating them later into English, the target language. This approach contradicts common practice in L2 instruction, in which learners are generally encouraged to compose in the L2. When queried, the participant said that she did this so as not to interrupt her thought processes while creating the text.

While it must be stressed that the methodology of Zamel’s study was not presented, the fact that at least one of the eight writers chose to write the text in her native tongue and then translate her work into English is a reminder that even before the advent of online translators (the study having been done in 1982), students could and did write compositions first in their L1, though the literature does not appear to investigate the frequency of such behavior. The self-described priority of the writer was not the form (creating the L2 form) but rather the meaning (formulating the thoughts first that would later lead to an L2 form).

Kern (1994) conducted a study in which think-aloud protocols were used during a reading task of 51 English-speaking students learning French. Students in the lower
ability group used more translation from their native English to French than groups with higher proficiency, but at the end of the semester-long study they had decreased their use of translation by 37%. Kern hypothesizes that learners can generate meaning for, and store in memory, parts of L2 that may go beyond their cognitive limits by translating them into their L1. As their language ability increases, less mental effort is needed to comprehend and process the language itself.

In another study, Cohen and Brooks-Carson (2001) investigated how writing an essay directly in the L2 compares to writing in the L1 or dominant language and then translating it into the L2; what students consider to be drawbacks and advantages of each mode; and how strategies might be different with those writing in a third language. The authors studied 39 intermediate-level French learners, of which 25 were native English speakers, 10 were Spanish-English bilinguals, and 4 had other native languages. In contrast with the finding of Zamel (1982), the direct method — writing directly in the target language — yielded better results than the translated method among English native speakers, while as expected there was no statistical difference between the two modes for Spanish-English bilinguals.

Looking at specific components, there were significantly better results for expression (sense of the language), organizational structure, clarity, and smoothness of connectors on the direct writing task than on the translation task. On average, students reported “sometimes” having to simplify their text when in translation mode to make the L1 fit better into the L2 translation. The vocabulary of those participants who did not simplify text was rated higher, indicating that simplification may be detrimental to overall writing. 64% of English-language participants also indicated using word-for-word
translation during the translation task, with a similar amount (two-thirds) reporting having difficulty finding equivalents for translation. Cohen and Brooks-Carson argue that thinking of concepts first in the L1 and then translating them mentally into the L2 may reduce the strain on working memory.

Luoma and Tarnanen (2003), while investigating L2 learners of Finnish, developed a conceptual structure (Figure 1.1) based on Skeehan (1998)’s oral performance model, to describe what they view as the performance priorities of writers.

![Diagram](Performance_priorities_in_writing)

Figure 2.1 Luoma and Tarnanen’s conceptualization of writing (adapted from Luoma and Tarnanen, 2003 p. 458)

The priorities—defined by Luoma and Tarnanen as meaning, form, and expression—are presented in the order of importance that the authors theorize writers place on each, starting on the left and proceeding to the right. The first priority when writing is “meaning,” consistent with Zamel (1982). According to Luoma and Tarnanen, meaning includes comprehensibility, readability, and fluency. “Form,” which involves both accuracy and linguistic complexity, is considered second. The authors believe “expression,” involving more stylistic concerns, comes last when a writer is completing a task. Luoma and Tarnanen’s hierarchy is a useful model since the effects of online translators might bear on both the meaning and form of compositions. If indeed meaning
is given higher priority than accuracy, as the authors argue and as indicated by their model, we would expect students to place a higher premium on the content of their compositions (developing what they are trying to say) as opposed to the form (developing how they say it, which may be largely delegated to the translator).

It is important to point out however that Luoma and Tarnanen do not make a distinction between L1 and L2 writing in their model. There is some comparison of the two in the literature, but the focus of most research appears to be placed on exploring one or the other. Scott (1996)’s book, *Rethinking Foreign Language Writing*, provides detailed background information on foreign language and ESL writing. Although the majority of the book refers to past and current thought on paper-based writing in FL and ESL, just two and a half pages (of 178) are devoted to the section discussing the connection between L1 and L2 writing. In that section, Scott cites several studies (Edelsky, 1982; Zamel, 1983; Friedlander, 1990; Valdes, Haro, and Arriarza, 1992) as supporting the belief that L1 writing skills transfer to the L2; however Edelsky (1982)’s study involved bilingual elementary school children and not L2 learners, Zamel (1983) looked only at “deficient” L1 and L2 writers (p. 6), and Friedlander (1990) and Valdes et al. (1992) indicate that the L1 supports writing in the L2 in planning and generating ideas.

Uzawa (1996) states that it is generally accepted that unskilled L1 and L2 writers tend to focus more on spelling and grammar — which would fall under “form” in Luoma and Tarnanen’s model — leaving less time for planning meaning and audience. Support for the model may also be found in Raimes (1987), as reported by Scott (1996). In a study of ESL students using think-aloud protocols, Raimes (1987) found that providing a
specified audience (such as a teacher or a fellow classmate) in a written task’s instructions had no apparent effect on L2 classroom writing. If we are to accept Luoma and Tarnanen’s model, in which register falls under “expression” and thus behind meaning and form in performance priorities, the lack of attention given to audience by Raimes’ participants is not surprising.

The L2 writing literature also explores the effects of dictionary usage. Hurman and Tall (2002) looked at the effects of paper dictionary use on quantitative and qualitative measures of scores on L2 writing, and speak to how frequently such tools might be used on a given task. They found that paper dictionary usage for compositions on exams correlates significantly with higher scores. Dictionaries were not generally used by participants in their study to check for grammar or spelling mistakes, but instead for creating content and for comprehension of instructions. Even more intriguing is the fact the authors noted “widespread ignorance of the differences between verbs and nouns” and “a number of phrases being taken directly from the dictionary” (p. 21).

For the experiment, two parallel writing exams were devised: one to be written with a dictionary and the other without. Each student wrote one exam in both of these conditions. The topics are not revealed in the article, but the format included various forms of writing, such as an essay, letter-writing, postcard-writing and form-filling (p. 22). A variety of factors were considered, including student level and dictionary usage. Student level was indicated based on which tier of the British General Certificate of Secondary Education (GCSE) they were taking—Higher Tier, the test difficulty level reserved for students deemed by their professors and themselves to be able to achieve higher marks (grades from A* to D) on the test, and Foundation, for students considered
to be less strong and more likely to receive lower marks (grades from C to G) in a
given subject. Tier level was factored in to see if more advanced students used
dictionaries more efficiently than less advanced students. Students’ performance was also
measured based on whether they used a dictionary on the exam: if so, how many
words/expressions were looked up, how frequently it was consulted, what type of
dictionary they used, which words it was used for. Students were to circle all words for
which they consulted a dictionary. The reliability of this procedure can be questioned,
since there is no assurance that students circled all words they actually consulted.

The results of the Hurman and Tall’s study show that students at both the lower
and higher level had 9% higher scores when they had access to a dictionary than when
they didn’t. In fact, they consulted the dictionary approximately once every two minutes
when it was available to them, based on self-report. The Foundation Tier, at a lower
level, tended to use the dictionary more to look up words in the instructions to understand
the question better (comprehension), whereas the Higher Tier used the dictionary to help
in their production and used it nearly twice as much as much as the lower-level students. It is
important to note that Higher Tier students were only slightly more accurate when they
consulted the dictionary (75.2% correct versus 74.0%), and that both High and
Foundation Tier students showed better performance when able to use the dictionary than
not. However, the lack of description by Hurman and Tall concerning the scoring method
does not allow further analysis of what made the higher scorers’ writing better or worse.

Hurman and Tall (2002) emphasized the importance of training students in
dictionary use. Common errors across the board included simply copying the infinitival
form without conjugating it for use, miscopying from the dictionary, and not making the
distinction between word categories, leading to things such as “Je gauche” for “I left” (instead of “Je suis parti(e)”) (p. 24). Some drawbacks to the study are the lack of explicit information on participants’ backgrounds and levels, and the fact that all test-takers apparently did not write about the same topic or use the same format in their responses, both of which possibly put internal and external validity in question.

Nevertheless, the study points towards two suggestions important to the present discussion: consulting aids can be helpful in written expression; if aids are used, students must be trained in the use of such tools. Bishop (2001) reported that 97% of students do not have any skills when it comes to dictionary usage, monolingual or bilingual, partly due to the fact that they don’t generally use the dictionary in their native language. He created a training course and looked at quality and accuracy of L2 composition writing. Those who attended the training course received scores roughly 10% higher than the control group (who did not attend the course) on rewritten papers, after both groups got specific feedback on grammatical errors and level of quality. It is not stated whether or not scores on the first drafts differed between the two groups in the study. The consideration that training appears to help in the use of traditional tools such as dictionaries is a crucial component of the current study investigating the role of a new tool, online translators.

2.3 Computer-based writing

This section deals with the research available looking at the effects of computer usage on foreign language or ESL writing. Blake (1998) provides an overview of the evolution of computer usage in language acquisition. Computer programs dedicated to
assisting language learners have been around since at least the 1960’s, with mainframes such as PLATO (Programmed Logic for Automated Teaching Operations) at the University of Illinois at Urbana-Champaign offering drills and similar practice exercises. Computer-assisted instructions (CAI) in the 1960’s and 1970’s went a step further, including branching routines that adapted to student performance by displaying subsequent exercises based on a student’s correct or incorrect response on a given item. In the 1980’s, thanks to advances in technology and a greater emphasis on learner-centered instruction, software was developed that allowed students more control over what exercises they would complete, and included multimedia (text, sound, graphics) elements that students could manipulate. With these changes came the new acronym CALL: Computer-Assisted Language Learning (Garrett 1987). Computers became an interactive tool that could be used individually by students on their own or together in the classroom with teacher-led instruction.

Iwai (1999) discusses the possible effects of technology use in L2 classrooms and presents a study which shows that an experimental CALL class may have helped students’ awareness of problems in content but not necessarily of errors in writing, such as grammar, spelling, or lexical choice. In his introduction, the author underlines the importance of “ascertain[ing] whether the modern tools for writing guarantee beneficial effects, and if so, why and for what aspects of writing” (p. 23). Importantly, Iwai cites a review by Bangert-Drowns (1993) which states that “instruction must explicitly identify and practice the skills that one expects to gain from the tool in order for those gains to occur” (p. 26). This is of particular importance in regards to the current discussion.
Iwai adapts Grabe and Kaplan’s (1996) writing model to the CALL L2 environment as a framework for the study, stating that writing consists of two major components: context (external reasons for writing, including situation and performance) and verbal working memory (the internal assembly and processing of real-world knowledge in language, through means such as goal-setting and metacognitive strategies, i.e. using one’s knowledge to plan language use). In Iwai’s study, 24 sophomore EFL students in Japan were enrolled in a class specifically created to measure CALL effects on writing before and after having attended fifteen 90-minute sessions. The class involved four main activities: informal e-mail writing among classmates, formal writing (subject to peer evaluation and subsequent self-revision), monitoring (the instructor would observe and point out anonymous good and bad writing examples to the class), and pairing (discussion of drafts with a partner via headsets), all but the last of these being done on computers.

The pre- and post-tests in Iwai (1999) involved the same task: participants read two paragraphs written in English by Japanese native speakers and made what they deemed to be appropriate corrections. The paragraphs were the same for both the pretest and the posttest. The study found that before the class, students paid most attention to “surface aspects” of the paragraph such as lexical and grammatical problems, with little attention to content. After CALL however, while the number of surface corrections remained statistically similar, correction of content problems increased dramatically: nearly four times as many content corrections were made on paragraph 1 and nearly three times as many were made on paragraph 2. Students were also asked to assess their own attention to certain aspects of a text when approaching its correction. During the pretest,
91.7% of students said they placed the highest priority on spelling and punctuation errors when looking at paragraph 1. After the CALL instruction however, this number dropped to 37.5%.

The author argues these facts put together indicate that CALL raises learners’ awareness since participants reported focusing less on purely linguistic aspects of the paragraphs and more on correcting content. An important limitation to Iwai’s study however is the lack of a control group, which may have affected the internal validity of the experiment. Since participants were exposed to the same paragraphs in the posttest as in the pretest, the issue of testing effect was not accounted for. Students, having already been exposed to both texts previously, may have had to focus less on the more salient feature of form the second time, allowing more attention to be directed towards the less salient content issues of the texts. Examples of content corrections are also not provided, making it difficult to determine what Iwai considered to be content-based correction.

Although arguably beyond the scope of Iwai’s study, it would have been informative if the experimenter had measured student production before and after the class as opposed to simply tracking error recognition and correction, since it can be assumed that students more often write compositions than correct or peer-edit them. Dreyer and Nel (2003) found significant gains in reading comprehension scores on the TOEFL (Test of English as a Foreign Language) after CALL instruction. Had a similar measure been performed before and after Iwai’s study on students’ written production in the L2, as opposed to just measuring their error recognition and the priorities they gave to various errors, it might be clearer what effect CALL had on the L2 learners involved in the study. Nevertheless, Iwai’s study shows at least that computer usage seems to raise
students’ awareness about the content of what is being said in written language, and about ways to improve it. It would be helpful to know whether the instructor focused particularly on content issues during instruction: perhaps if lexical and grammatical issues had been emphasized, improvement would have been shown in these areas.

Consideration of this study also brings up an important issue: if students do in fact attend more to content when engaged in various CALL activities (some of which will be dealt with below), it is possible that students’ compositions are richer in content when online translators are being used. The trade-off however is a lack of change in awareness of grammaticality, if the findings of Iwai’s study can in fact be applied more generally. Students may therefore be trying to produce content that is beyond their linguistic capabilities, and for which they are unable to detect possible translator errors. Students may need to be “trained”, especially at lower and intermediate levels, so as to understand this give-and-take between content and form.

Heift (2003)’s German Tutor might be a step towards creating the type of understanding of grammaticality lacking among students such as those in Iwai’s study. German Tutor is a drill-like system which provides direct corrective feedback to students based on a cumulative evaluation of students’ ability level, calculated according to the number of errors made in 12 different categories (e.g. direct object usage, word order, verb compliments). It is possible with German Tutor to be considered an advanced learner in one category, intermediate in another, and beginner in yet another based on past performance.

The feedback for errors in German Tutor is based directly on one’s level. The program gives students slash completion exercises (for example, “(Der Artikel)/
Zeit/laufen”). Students need to type in a grammatically-correct sentence with the given elements. If the student is classified as a beginner in definite articles, and they enter the incorrect definite article “der”, the program will note the error and will tell the user explicitly: “DER is not correct here. ZEIT is feminine.” (523) An intermediate learner making the same error however would receive the error message: “There is an agreement error between the determiner DER and the noun ZEIT.” (523) It would be up to the intermediate student to determine in what the error consisted, and also to know what was meant by the metalinguistic term “agreement.” Unlike a dictionary (and, arguably, an online translator), the program is leading the student to the correct answer without providing it directly.

To test the effects of German Tutor and see at what level students placed in various categories, 33 students of introductory German classes participated in three one-hour sessions, supplying 1352 tokens (i.e. sentences entered by students into the Tutor). There were some categories that all participants generally did equally well or poorly in, while in other categories there was a clear distinction between advanced and beginner students. For example, all students had roughly equal problems with word order, meaning that nearly everyone ranked as a beginner in this category, while a clear distinction could be made between beginners and advanced students in accuracy on indirect object choice (where beginners made 64 errors while only 1 error was made by advanced learners). In all, 30.3% of participants received “remedial” feedback for at least one category, indicating variability in proficiency according to the form in question. While the program and the study are interesting, inferential statistics are not provided, preventing us from
knowing whether the provided percentages and counts (showing the raw number of errors made in each category by different user levels) are statistically significant.

Another problem in Heift (2003) concerns methodology: all students were given corrective feedback by *German Tutor*, meaning there was no control group. Additionally, the cumulative effect of the program is only discussed briefly. Four students’ performances were randomly analyzed. Their scores (whose basis is not discussed) all went from 15 points to between 20 and 25. This possibly indicates a positive effect of feedback from a computer on learner grammar, but a follow-up analysis with more of the students’ performances compared would need to be conducted to assure reliability.

A somewhat related computer tool is *Système D*. In Scott (1990), the author takes a first look at this program, which aims to help students in a different way. *Système D* is a program by Noblitt, Sola and Pet which offers aid to English-speaking French L2 students as they write on the computer. Unlike *German Tutor*, it is not a tutor or corrector, but contains different modules that students can consult freely: a 4,400-word bidirectional dictionary, a verb conjugator which displays complete conjugations of verbs in all the major French tenses, a grammar index explaining and giving examples of key grammar points for beginning and intermediate students, a vocabulary index organizing words according to themes, and a phrase index providing common expressions organized by pragmatic function (such as greeting, congratulating, etc.).

For the study, an unspecified number of elementary to advanced learners were given a one-hour introduction to *Système D* at the beginning of the Spring semester and required to use it throughout the semester in their writing. Scott found that the majority of participants used the dictionary; some students occasionally used the verb conjugator; but
no student reported using the grammar, vocabulary or phrase indices. Scott characterized this as “quickly revert[ing] to typical student behavior: writing a composition as quickly as possible using the dictionary to satisfy all their linguistic needs” (p. 60). This would be consistent with the above discussion (Luoma and Tarnanen, 2003 among others) of emphasis being placed first on meaning (content) before form (grammar). Difference in usage between beginning, intermediate, and advance learners is not discussed except in a brief footnote: advanced users over time had less confidence in Système D’s dictionary as a reliable resource due to its limited lexicon.

Bland et al. (1990) examines evidence gathered through CALL to explore how students use language during composition writing. The authors identify stages in student development based on the types of queries they made during a study on student use of Système-D, the bilingual word processing and language reference tool mentioned above. In this study, the keystrokes of ten volunteer participants were tracked in Système-D during French composition tasks. These logs were then compared to the final products to help find out how participants had used the information looked up in Système-D. The authors identified three main classes of queries in the data. Class I, Token Matching, involves taking a word as-is (including any morphological endings) in English and attempting to find a French equivalent. Two examples cited were participants trying to find the French equivalent for the words “plants” and “would” instead of looking for the base word “plant” or considering the idea of “would” as a grammatical rather than lexical concept. These examples point to the fact that each word was unanalyzed and simply inputted as a “chunk” (p. 442) for which an equivalent “chunk” was sought in French.
For Class II, Type Matching, English base words are used to find a French equivalent, or grammar concepts were looked up in order to find a French equivalent of an English structure. Two examples of this are searching for “play” when the participant wanted the plural form, and looking up “negation” in the grammar index. Both strategies indicate going beyond a simple word-to-word correspondence and an attempt to match word types and grammar concepts in English with French, instead of assuming that each English word had an exact equivalent in French as with Bland et al.’s Naive Lexical Hypothesis mentioned above.

Class III, Relexification, involves participants’ use of circumlocution in the L1 or L2 to try to construct meaning, either by finding synonymous or near-synonymous terms or by using different structures in the L1 and L2. For example, a participant querying “tyrant” and then moving on to “ruler” when the first query did not work, or coming up with “en désordre” for the word “messy.” According to the authors, 24 percent of all queries were in French, with students sometimes querying the result they received in English to check whether or not what they found in French was correct.

Bland et al. state that assessing the types of queries made by students in Système-D can help the instructor determine how to approach teaching the L2. If a student is making Class I queries, time needs to be spent on emphasizing the differences between L1 and L2 to overcome the notion that there is a one-to-one correspondence between words in the two languages. When students are making Class II and III queries, strategies for dictionary reading and lessons on circumlocution might prove more helpful. The authors note that language acquisition does not proceed in a linear fashion, and that learners may make Class I and Class III queries in the same session. Rather, it is a
gradual process towards more advanced conceptualizations of language that must be fostered by instructors, perhaps through programs like *Système-D* in order to point students who make more basic queries in the right direction, for example as student looking for the French equivalent of “never” would find a link to an explanation of how French negation works. Bland et al. conclude that more research is needed to determine how effectively to deal pedagogically with the various stages of language learning. Although a new version of *Système D* was released in 2004, no articles were found describing or studying the new software in the literature.

Knight (1994) found that intermediate-level Spanish participants who had access to a Spanish-English computer dictionary while reading magazine articles on a computer outperformed students who did not have access to one in terms of reading comprehension and retention of new vocabulary on both immediate and delayed posttests. While showing that online dictionaries can be useful to L2 learning, the study focused on the effect of usage of the computer dictionary versus inferring from context, not paper-based dictionary use. For this reason, it is difficult to say whether it was the access to a dictionary, the fact that it was computer-based, or a combination thereof, that led to the results.

Based on the above studies, it is difficult to draw definitive conclusions about the possible usefulness of computer tools on L2 learning or production. While according to Iwai (1999) CALL can increase student attention to content, it is unclear how much of a role instruction, computer tools, or a combination of the two had in increasing student awareness of issues related to content. The lack of a control group in Heift (2003) and lack of details concerning the participants in Scott (1990) add to uncertainty over what
effect, if any, computer-based language tools may have on L2 writing. Blake’s (1998) call for increased research into the best implementation and assessment of CALL still holds true.

2.4 Machine translation, or non Web-based MT

The goal of this section is to look at the literature describing and analyzing machine translation (MT). The term broadly applies to translations performed or provided by computers at any stage of the process. It is estimated that more words are now translated via various forms of MT (both offline and online) than by human translators, and the demand continues to grow (LISA 2009). For the purposes of our discussion, a distinction will be made between two types of computer translation. The term “machine translation” (MT) will be used to refer to those computer programs or systems which were generally developed for research or to be licensed for a fee by professionals, which reside locally on a computer or network, and which are designed for translation between two or more languages. The term “online translation” (OT), also known in the literature as Web-based MT, will be used to designate tools that were developed to be freely available to the general public via a non-paying web-based form to perform translation between two or more languages. Section 2.4 discusses MT, while Section 2.5 will explore OT.

Machine translation programs are generally aimed at those researching or working in the field of professional translation, but there are obvious implications for the current discussion as they convert text from one language to another. Not all texts for MT are created equal, and not all translations are done for the same purpose. Melby (1997), based
on Kay’s landmark 1980 article “The Proper Place of Men and Machines in Language Translation,” proposes four categories of machine translation according to text type and desired outcome of the translation.

- Controlled domain-specific text and high-quality output
- Controlled domain-specific text and indicative output
- Dynamic general text and indicative output
- Dynamic general text and high-quality output

Controlled domains, according to Melby, are ones in which “well-understood formal and tractable language” is used (p. 30), whereas dynamic domains are ones presenting a wider variety of words and terms that are harder to manage due to being less formalized. An example of controlled-domain documents might be DVD instruction manuals, in which there is a limited set of specialized vocabulary that is generally used in a consistent manner. A dynamic domain has a very broad, often unpredictable or changing set of vocabulary that can vary widely from text to text. A short story might be an example of a dynamic general text.

The first category listed above, controlled domain-specific text and high-quality output, is the most often desired in the business and professional worlds, where often a specific set of specialized vocabulary is used in a fairly predictable way. It is difficult however, at least at the current stage of machine translation, to obtain an input of high-enough quality from machine translators that requires little or no intervention on the part of humans before presenting to the end user. One possible counter-example to this might
be the RALI: MÉTÉO system used by Environment Canada to translate weather reports from English to French, which has seen some success due perhaps in part to the high amount of repetitions inherent in the weather reporting system used (Langlais 2006).

The second category, controlled domain-specific text and indicative output, represents the most common type of output in professional translation. Even with controlled domain texts, terminology overlap from different organization or companies can be as little as 20% (Bennett and Gerber, 2003), so a high-quality machine translation cannot always be expected. The computer instead creates a time-saving rough draft of the translation that a human translator can revise as a basis for a polished translation presented to the final user (such a case is described in Sherry, 2003 concerning an ethics code that is first translated by a machine and whose details are then polished by organization members afterwards).

Dynamic general text domains are more common and involve documents that use a wide variety of vocabulary that can change significantly from text to text. The third category, for example, corresponds to a situation commonly encountered by L2 learners or the general public who wish to get a gist of a dynamic domain text such as a short story, general newspaper article, or other varying texts in the L1 to the L2. Its results produce a rough translation of the original text that would satisfy casual users but would presumably need revision by humans before being ready to be presented to an audience (see discussion of Leffa, 1994 below, however).

The fourth category is from a translation standpoint (but not necessarily pedagogically) the ideal situation: a machine translator that can provide high-quality translations of an array of general topics with little or no human intervention. This type of
translator would be able, for example, to render short stories about a wide variety of subjects into the target language with little help from humans. Bennett and Gerber (2003) call this the “holy grail” of the MT world (p. 181). The paradox is that category for, the most desired type of translations, is the arguably most difficult to achieve because of the relative variability and unpredictability of the words and structures used. Currently, machines are unable to do translations of the fourth type without human intervention. Melby states that “MT is currently appropriate in the first three cases; yet the fourth case, which currently requires human translation, accounts for the majority of all translation that is done, probably eighty percent.” (p. 31)

Carl, Pease, Iomdin, and Streiter (2001) approach translation types from a slightly different angle, presenting five qualities of translation, which in order from lowest to highest quality include: indicative, informative, literal, reliable, and user-oriented. Indicative or informative translations do not necessarily translate meaning or form accurately but are meant to provide, respectively, a rough idea or general information about the topic covered in the text. Carl et al. consider a literal translation superior to an informative or indicative translation, explaining that although the target language is not native-like, literal translation provides a “translation for each unit of the source text in a correct grammatical form.” Reliable translations go one step further, presenting both meaning and form in a more idiomatic way, while user-oriented translations will not be perfect, but are expected to be sufficiently accurate and idiomatic throughout to fully satisfy all the purposes of a given user.

According to Leffa (1994), the progression of MT technology has been slow, despite improving from only word-to-word translations in the 1950’s and 1960’s to
beginning to deal in the 1980’s with anaphoric references (such as “her” in “The woman sold her car”) in part thanks to the theory of reference frames (Minsky 1975). Machine translators are still unable to accurately perform some translation tasks consistently and remain error-prone in grammar and vocabulary due to the inability to deal with context. A variety of approaches have been developed to improve MT. Attieh (1992) and Hutchins (2003), among others, discuss in depth the origins and historical aspects of machine translation; for our purposes, a brief overview is given below.

In the 1950’s, machine translation was developed nearly exclusively for and by the military. The principal use originally was for what Attieh (1992) calls TAO — *traduction assistée par ordinateur* — or in English, Machine-Assisted Translation (MAT). The computer was not expected to give a final product, but rather to assist the military in quickly being able to get a general understanding of Russian communications. The machine would provide a rough translation to determine the subject or main idea of intercepted communications, and a human would decide whether or not a full human translation of the document was warranted.

According to Attieh, the first translation conference was held at MIT in 1952. In 1954 came a machine translator developed by Georgetown and IBM containing a simple list of words in the source language and a single corresponding translation for each. Since both computer technology and formal linguistic theory were still being developed — Hutchins (2003) notes that the work of pioneers like Noam Chomsky and Zellig Harris had just begun — most successful early MT models were based on simple dictionary or lexical correspondences as opposed to later rule-based translation approaches.
An initial enthusiasm and optimism for fully-automatic high quality translation systems soon gave way to disappointment, particularly after the well-known ALPAC (Automatic Language Processing Advisory Committee) report which found MT to be “slower, less accurate, and twice as expensive as human translation” (Hutchins, 2003 p. 7). This led to drastically cut budgets at universities and government institutions for MT research. Technological advances and several agencies’ interest in MT, notably the US Air Force’s hiring of SYSTRAN (later the creator of Babel Fish) to devise a Russian-English translation system, helped to gradually bring MT again to the forefront. Development of lexical and grammatical formalism led to renewed academic research in Grenoble, France and the University of Kyoto, among others, resulting in further development of rule-based and then interlingua or frame-based approaches. The 1980’s saw the first widespread uses of commercially-sold translation systems (Trados, TranslationManager, and Transit among others) used by professional translators at first over local intranet connections or via what would be later called the Internet, and eventually on locally-installed PCs. Online translation tools would appear in the late 1980’s in France and on the Internet in 1997.

There are several types of machine translation approaches that have been used through the years. Most often, different approaches have been combined to various degrees within a given system in an attempt to improve translation performance, but generally models are based on one approach with other component(s) playing a supporting or confirming role once an initial translation is done.

One type of machine translation uses a frame-based model (Agirre, Arregi, Artola, Ilarraza, Sarasola, and Soroa, 2001), sometimes referred to as an interlingua
method (Hutchins 2003, Koehn 2010). In such an approach, instead of translating words directly from one language to another, words are first converted into frames which represent abstract lexical concepts (Agirre et al., 2001). The machine translator attempts to match frames from the source language with frames from the target language, forming hypotheses about the appropriate target form. The human translator can confirm or disconfirm the output while the system compiles interactions to refine the “dictionary” of possible frames to use for translation (Agirre et al., 2001). A similar approach is called transfer-based MT (Kit and Wong, 2008), in which the processing of text is divided into three tasks: analysis, transfer, and synthesis. During analysis, surface forms are transformed into abstract representations. These representations are then transferred, or mapped, to corresponding ones from the target language. Lastly, during the synthesis phase, the abstract representations in the target are synthesized together into the surface forms which constitute the output.

One category of MT identified by Cribb (2000) is example-based translation. A database is kept matching up pairs of phrases and sentences in two languages, which are accessed when needed. Carl et al. (2000) discuss the differences between Example-Based Machine Translation (EBMT) and Rule-Based Machine Translation (RBMT). The former, EBMT, generally relies on translations performed by the machine previously that are stored in memory and can be used for future translation, allowing for adaptability of the translator over time. According to Carl et al., this type of translation typically does very well with sentences previously encountered and with well-defined subject areas with less variation in texts, but poorly with broader or multiple subject areas or completely novel sentences.
The latter, RBMT, generally relies on preset formulas for translation vocabulary, syntax, and other features from the source to the target language and does not store any translations entered. A possible limitation of this is that, unlike a system processing naturally-occurring texts which present language as it is actually used, linguist-inputted rules may not — and possibly cannot hope to — fully capture all a language’s features, suffering from oversimplification and/or overspecification of rules (Bennett and Gerber, 2003). For Carl et al., RBMT performs better with new sentences in a broader coverage of contexts, but translation results are not retained and thus no information is learned from previous attempts. Carl et al. argue that on their own, EBMT and RBMT sometimes reach the higher end of the quality spectrum, but tend to give indicative or informative translations of texts overall. They propose that if the two approaches are combined, the quality could eventually reach higher levels since the strengths of the adaptability of EBMT could be combined with the greater coverage of RBMT.

Another category of MT, statistical machine translation, relies on calculations of probability to figure out how likely an item in one language will correspond to that in the other. Google states that its popular Google Translate OT tool is based on a statistical translation model (Genzel 2010). Langlais (2002) describes an experiment with one such system which involved 14 professional French-native translators using a statistical translation system called TransType that offered suggested translations to users for translating a text from English to French based on statistical analyses of a training corpus. This system also incorporated translation memory so that once a translation of a word or phrase had been made by a user, it was taken into account for subsequent occurrences. TransType proposed a translation deemed acceptable by the professional translators 42%
of the time. Two limitations of the study, however, were that the system’s lexicon was “handicrafted” by the researchers — which may not fully replicate normal usage of professional translation software — and there was actually a reported drop of the participants’ productivity by 17%, meaning that it took longer to use TransType than to translate without it.

A related approach is corpus-based translation, in which the computer is fed a large number of samples with bilingual corpora: texts that have already been translated from the source language to the target language which the program aligns sentence-by-sentence, phrase-by-phrase, or even word-by-word to seek text strings that are likely to correspond between the two languages. If a translation program finds, for example, that “la maison” occurs frequently in the same phrases or sentences as “the house,” the system will consider it probable that these two strings are equivalent. The more bilingual corpora are treated by the machine, the more likely the system will find correctly corresponding translations. For corpus-based translation, “there’s no data like more data” (Al-Onaizan et al. 2003). A limitation however with corpus-based translation and statistical models, as compared to rule-based translators, is that the computer may not propose suitable translations simply because they have not been previously encountered by the system. “Just because something has not been seen in the training text does not mean that it is impossible” (Koehn 2010).

Attieh (1992) distinguishes two types of machine translation: traduction automatique or TA (literally, automatic translation, usually referred to as machine translation or MT in English), referring to computer translation from one language to another without human intervention, and the above-mentioned TAO — Traduction
Assistée par Ordinateur, called MAT (Machine-Assisted Translation) in English — in which the human operator uses the machine translator as a tool to choose the appropriate translation. Attieh holds that only the latter can produce satisfactory results. Some limitations of machine translation presented by Attieh include polysemy, synonomy, anaphora, homonymy, metaphory, and ambiguity. Some such problems are highlighted by the (in)famous translation of “The spirit is willing but the flesh is weak” to Russian and back to English being rendered as “The vodka is strong but the meat is rotten.” (p. 5) The author advances the idea, attributed to Yehoshua Bar-Hillel, that producing a high quality machine translation is impossible since for understanding and producing language, world knowledge (e.g. context, culture, experience), is needed in addition to word knowledge.

Word-by-word, syntactic, and interlingua (“langue pivot”) translations have proven not to be as useful as hoped, according to Attieh, because the meaning of words is often situationally or culturally contextual and thus cannot be easily translated. The previously-discussed Canadian weather translation service MÉTÉO however is alleged to be 97% accurate, which the author attributes to the restricted, formulaic domain in which it is used (weather reports) and the fact that syntactical, morphological, lexical, and semantic analyses are automatically performed on the translations, which helps catch potential errors. Attieh argues that machine translation will likely be pursued in the interest of time, cost, and productivity, but in line with Langlais (2002), cautions that for some translators at least, the post-translation work needed after using a machine translation may be more bothersome than doing the translation without the machine.
Leffa (1994) describes machine translators as “unable to produce error-free translations, but capable of helping people extract meaning from many otherwise unintelligible foreign language texts.” (p. 392) Appearing at a time when personal computers were becoming more widespread, Leffa’s article touches on a subject largely ignored elsewhere: the comprehensibility of machine-translated text among native speakers. The program used in Leffa’s study, a unidirectional English-Portuguese translator, was designed to look at textual ambiguities through a double-pass system, where individual words or set phrases are first translated from a dictionary of approximately 3,000 words and assigned probable syntactic and semantic values. Programmer-entered rules (such as article agreement in Portuguese) help the program to anticipate the functions and meanings of words and phrases based on preceding and subsequent terms in the sentence. A second pass is then conducted, building on the tentative hypotheses established in the first pass, to analyze and give a translation to the words and phrases within the sentence.

In Leffa’s study, two science articles (one about the Milky Way and the second about iron) of about 300 words were chosen and were both translated separately into Portuguese by the machine translator described, as well as by a professional human translator. A multiple-choice test was devised to test the comprehension of the two articles in 88 native Portuguese eleventh-graders, with each test participant receiving both texts: one of the texts as translated by a human translator and the other as translated by the computer. Half of the participants received the computer-generated Milky Way text and the human-translated iron text, while the other half received the human-translated Milky Way text and the computer-generated iron text.
The results of Leffa’s study show that there was no statistical difference ($p > 0.05$) in comprehension test results between readers of the machine-translated texts (who scored on average 2.84 on the texts as translated by professionals and 2.76 on the texts as translated by computers). Leffa hypothesizes that this may be because of the “similarity between the translation process used by the computer MT program and the comprehension process presumably used by the reader: both employ analysis followed by synthesis.” (p. 398) He believes that just as the computer deconstructs texts via its dictionary and pre-programmed rules, readers deconstruct the text via their knowledge based on the subject and previous experiences to make input comprehensible.

It is important however to point out, as Leffa does, that the software had one advantage in the experiment: it was fed all the relevant vocabulary from the two articles. For this reason it approaches more a “controlled domain-specific” environment (as outlined in Melby 1997) as opposed to the typical environment for the present discussion, online translators for general texts. It might be further argued that since specific vocabulary from the articles was entered into the system, a preliminary word translation was already performed for the computer by programmers. The results of this study nonetheless bring up two important issues which will need to be explored: the extent to which machine or online translations can produce comprehensible input in spite of the errors produced, and some of the specific goals of L2 writing in relation to translators (for example, meaningful output versus grammatically-correct output).

Sabieh (2002) does see potential benefit to what she deems a sensible use of translators. Sabieh addressed the International Translation Conference on the Challenges of Translation and Interpretation in the Third Millennium, discussing current issues
important to professional (human) translators. She discusses among other things the role and increasing use of MT and OT, both by beginning and advanced students as well as professional translators. According to Sabieh, many in the field of translation, especially students, see MT as efficient and as a serious candidate for replacing human translators some day. While not attempting to prove or disprove how typical this assertion is concerning the feasibility or desirability of machine translators replacing humans, she suggests that machine translation should “not be considered a tool itself”, but instead a “powerful partner” (p. 2). Sabieh argues that computers may force students to approach language on a cognitive level, examining language instead of simply using forms to convey ideas.

2.5 Online translation and other writing tools

This section concerns electronic tools that purport to convert words or phrases from one language into another for L2 learners or users. One such resource available for those writing in a second language is electronic dictionaries. McAlpine (2003) takes on an interesting view concerning L2 dictionary use. “The role of the dictionary, whether it is used electronically, on-line or in print form, is to help learners expand their vocabulary and increase their awareness of common grammatical errors” (p. 72). Other studies mentioned by McAlpine (Luppescu and Day, 1993; Knight, 1994; Hulstijn, Hollander, and Greidanus, 1996; Fraser, 1999) also point to the beneficial effects of paper and electronic dictionary usage in the retention of vocabulary. A shortcoming of paper dictionaries however is their limited space. The author points out that online dictionaries can potentially have a virtually limitless amount of entries and hyperlinked cross-
references. Despite this argument in favor of electronic dictionaries, McAlpine claims that less-developed incarnations of electronic dictionaries often contain fewer entries and examples of usage than paper dictionaries can have a dangerous consequence: they propagate a “naïve view” (p. 74) that there are one-for-one equivalents between one’s native and foreign language.

The project undertaken by McAlpine is an online monolingual dictionary that he describes as having production as its main focus, in that links to words that are most commonly associated with a given vocabulary word or concept might help learners to look beyond word-for-word translations, see language in context and allow them to reproduce it in context. For example, an ESL learner trying to find the word which expresses the idea “take a child away” might look up “take” and eventually be lead through cross-referenced links to the word “kidnap.”

The dictionary’s creator himself indicates a problem with this system: the amount of time and money it would cost to compile and crosslink an entire dictionary made in this way is prohibitive (a comparable paper investigation by Fillmore and Atkins (1994) of the meanings and contexts for the word “risk” alone took up 44 pages). One thing McAlpine does not mention is that while the online dictionary may be thorough and give good examples, some language learners will likely not be willing to take the time to read through material found in numerous hyperlinks to find the right term they need. Furthermore, such a dictionary would have to be limited to intermediate and advanced learners, since a beginning learner coming across “kidnap” may have no idea if this is the exact word s/he is looking for, since related sentences may not make the meaning clear enough (one can take a child to numerous places without kidnapping them, for example).
While certainly being a potentially richer source of information than paper dictionaries and translators, it remains to be seen how learners would react to the project.

While no study encountered in the literature directly compares dictionary and online translator use, the two involve utilizing an outside resource to aid the writer in finding forms (lexical and in some cases grammatical) to be used in writing. The line between translator and dictionary is sometimes blurred. Burton (2003) suggested online translators may best be used as a sort of mini multilingual dictionary for students to check out or test vocabulary words. Sometimes the tools themselves contain unclear use of terminology, such as is the case with http://www.wordreference.com. The main function of this site is to give single-word definitions or “translations” as they are called on the site; however, numerous examples are cited for many words. In the English-Spanish dictionary, under the entry “happy” for example there are no less than 16 examples of expressions and sentences given in bold, such as “I’m so happy about the results estoy muy satisfecha con los resultados.” The use of the word “translation” for such searches may in part be to the way online translators are used: according to Yang and Lange (2003) over half of phrases input by users to SYSTRAN’s Babel Fish, discussed later, involve one- or two-word translations.

Myers (2000) examined what was referred to as a hand-held translator. A study was done over a 12-month period of two Chinese ESL students in a technical program in Canada. Both had been in Canada for two years prior to the study. The dictionary they used was ‘Instant Dict’ GD 100 P, which is bilingual and can “speak” the translated word aloud. No mention is made of the size of the translator’s vocabulary. To evaluate their use of the translator, a comprehension task and a writing task were created. Each word or
phrase they had the machine translate was noted, and the author later noted whether the translated item was correct or incorrect in the context it occurred in the text. The students worked together on the tasks. The participant with the lower TOEFL score (not specified) asked more questions of the other participant when collaborating. According to Myers, through use of the translator, the participants “quickly” learned the meanings of common prefixes and suffixes since the translator did not always provide translations of words with affixes, making the participants think of how to enter a word to get a workable translation. It should be noted that a word entered for translation on the machine is normally shown to the user in a typical context. It is not clear from the description if one can enter an entire phrase into the translator to get meanings to expressions or whether only individual words can be used.

One result of this study was that the translator was used to go from English to Chinese both for reading comprehension (which could be expected) and for written production (which seems counterintuitive). Myers hypothesizes that in many such cases, the participants already had an English word in mind which they had encountered previously and only wanted to verify their recollection of its meaning or usage. In addition, Myers says that English spelling improved over time among the participants, although no numbers are given for either percentage of correct usages or for spelling. This is a major drawback to this study: the author collected data on accuracy for over a year, but gave no numerical indication of the usefulness of the translator in the article. The author does specify that the translator would only be useful to intermediate or advanced students, since beginning users would not be able to tell if the translation provided would work well in a given context.
This potential inability to determine the contextual appropriateness of a translation provided by an electronic translator (such as that described by Myers) is also a potential problem with online translators, the subject at hand. The first OT system was Babel Fish by SYSTRAN, launched on December 9, 1997 on the AltaVista search website (Yang and Lange 1998). What was to become Babel Fish actually began nearly a decade earlier in 1988 as a service on Minitel, the French precursor to the Internet broadly used at post offices, businesses, and homes until the widespread adoption of the Internet slowly displaced it. Due to the cost, limited capacity, and small coverage (limited to France) of Minitel, the project was transferred to the Internet and quickly became a popular service: in less than a year, SYSTRAN reports that usage had exceeded 500,000 translations every day and by 2000, approximately 1.3 million translations a day (Yang and Lange 2003). A number of similar free translation services have appeared online, including among many others Reverso, LogoVista, Free Translation, and Google Translate. According to Koehn (2010), total translations by sites such as Yahoo’s Babel Fish and Google Translate exceeded 500 million words a day in 2007.

Despite their continued evolution, there are still flaws in machine translators, online or not. Luton (2003), in an article entitled “If the computer did my homework, how come I didn’t get an ‘A’,” discusses the evolution of electronic translators to the advent of free web translators in the 1990s. Luton points out several pitfalls of online translators such as Babel Fish. They don’t always recognize idioms or context: Luton gives the example of a basic phrase, “Where’s the bathroom?” being translated incorrectly into French (Babel Fish uses “salle de bains”, which refers to a washing room that has no toilet). (p. 768) Translators also do not recognize sometimes when not to
translate proper names (“brun de John” for “John Brown”), and even fairly common colloquial terms are mistranslated such as “cool”—“C’est frais” is how Babel Fish translates “That’s cool.” In some cases, SYSTRAN was aware of limitations of its free online translator for specific phrases. Yang and Lange (1998) mention specifically that a common requested translation was “What's up?”; by the time of the present study in 2008 and as of this writing, the translation of this phrase was still not correctly rendered by Babel Fish, at least from English to French.

Christian Raby, from SYSTRAN Software, describes some of the workings behind Babel Fish (Raby 1998).

SYSTRAN’s methodology is a sentence by sentence approach, concentrating on individual words and their dictionary data, and on the parse of the sentence unit, followed by the translation of the parsed sentence. Three major groups describe the SYSTRAN architecture: Dictionary, Systems Software and Linguistic Software. Each of these consists of a great number of modules which all work together to create a fully automatic MT (Machine Translation) system. (p. 499)

In the Dictionary group are stem dictionaries and expression dictionaries, including idioms and rules for disambiguation; the Systems Software group include various modules that parse syntactic and semantic relationships in a deterministic manner — each decision by one module is passed on to the next for analysis, with a filter to flag any errors detected for further processing if needed; and the Linguistic Software group which
contains algorithms for building target language constructs, such as negations and adverb placement.

Klein (1998) gives some examples of mistranslations he found by using Babel Fish. “Miss distance,” a measure of how much a missile misses its target, is translated as “señorita distancia.” Babel Fish incorrectly identified “miss” as a term of address, as opposed to an attributive preceding the noun “distance.” According to Klein, machine translation historically has been interactive, allowing users to choose from several choices while translating, or allowing modifiable dictionaries. Price however for such systems can be prohibitive. OT allows the user no control over dictionaries or databases, but has the advantage of being available for free. Klein does not completely dismiss Babel Fish, however. He claims that if taken under the correct context — not as a replacement for human or traditional machine translation, but as a way of “international information transfer” — Babel Fish could be useful for transmitting general “suggestions” to users about the ideas of a text in a language other than their own (p. 9).

Allen (2000) also discusses the uses and possible value of online translators. Allen, a creator of custom localized versions of software, mentions the fact that online translators are usually free test versions of paying translation software. He argues that translators like Babel Fish are offered for free on search portals and other web sites in order to make Internet users aware of the existence of paid translation services or software and to “create a market” among a wider Internet audience for machine translation programs (p. 45). The machine translation company Softissimo in France reported that they had over 10,000 submissions into their free online translator every month in 2000. According to Allen, online translators are often limited to French, Italian,
German, English, and Spanish (which are known to some in the field as FIGES) because these are languages for which the largest potential market is available, justifying the investment of time and money involved in making a free online portal available to the general public.

The level of quality and accuracy of online translators is still up for debate, with conflicting results from limited studies in the literature to date. Ablanedo, Aiken, and Vanjani (2007) paint a fairly positive picture for OT. The authors had 10 English-language text samples translated by an expert human translator, an intermediate-level Spanish translator, as well as Babel Fish. The results of the translations, as judged by two objective reviewers, found that all the translations were usable and understandable, whether human- or machine-produced. Those by the most fluent human translator were judged to be 100% accurate, the intermediate-level translator was judged to have 80% accuracy, while the online translator achieved only 70% accuracy but was able to complete the translations 195 times faster than the humans.

Cribb (2000) considers the possibility that machine translators, including online ones such as Babel Fish, might someday replace English as the principal means for communicating among people who do not share a common native language. Cribb (2000) mentions Moore’s law, in which computer technology is predicted to double in speed and capacity every 18 months, and compares it to the relatively slow and deliberate advances in L2 and foreign language methods. To test out online translators, Cribb entered several passages into Babel Fish. The first — a letter written in English, translated into French, and then translated back into English — produced several lexical errors, but translated the majority of semantic context coherently. Another text, an article in French from
Luxembourg, was translated into English with very few errors. Cribb posits that the system produces high-quality results for denotative meanings, but less well for more connotative meanings. While admitting that many texts (particularly translated between German and English) do not produce such good results, Cribb concludes that “despite the current weaknesses of MT output, current capabilities warrant its serious consideration as a bridge across the global language gap” (565).

Yates (2006) and Kit and Wong (2008) both analyze Babel Fish’s accuracy in translating legal documents. While clearly stating that Babel Fish is intended for general, non-subject-specific translations, Yates’ study, involving translating German and Spanish legal documents into English, is motivated by what she views as the widespread use of machine translators among law librarians. Five sentences were taken from the German civil code and five from the Mexican civil code, written in 1896 and 1928 respectively, while five additional sentences were taken from recent law-related press releases from each country’s respective foreign ministry. Each sentence was entered into Babel Fish and coded for lexical and structural errors. Errors were rated on a three-point scale, from minor to severe. Any sentence containing one or more severe errors was considered a failed translation, while those containing only minor or moderate errors were considered successful translations. Since Yates argues that legal language is inherently difficult to translate due to incompatible legal systems that vary by nation, the threshold for successful translation was one that “conveys the same general meaning as a professional translation of the text [...] but not necessarily in a strict legal sense” (p. 491).

Yates found that 15 out of the 20 translation attempts were failed, with sentences containing anywhere from one to seven severe errors. All sentences contained at least one
minor error, usually accompanied by moderate or severe errors as well. Overall, Babel Fish translated German sentences to English better than Spanish sentences. Most errors for Spanish translations were from the press releases, which Yates found surprising due to outdated vocabulary and structures found in the civil code excerpts. Unfortunately, Yates does not break down the errors by category (lexical versus structural), but only by source language and text type. Due to the number of severe errors, Yates concludes that at least as law is concerned, Babel Fish should only be used to understand broad, general ideas of a text (for example, to determine the subject of an article) and not for precise translation of concepts. The difference in error numbers between sentences translated indicates that more research is needed to determine Babel Fish’s performance with other languages and text types.

Similar to Yates, Kit and Wong (2008) investigated the translation of two legal texts, an excerpt from the United Nations Declaration of Human Rights and one from a European Union treaty, from 21 different languages to English using 6 different online translators: Babel Fish, Google Translate (GT), PROMT, SDL, SYSTRAN, and WorldLingo. The results showed that no one system outperformed the others on all languages: some translators were better at certain language pairs than others, with higher scores generally between European languages (GT’s high score on Arabic-to-English being an exception). The authors determined that systems based on SYSTRAN’s translation engine (Babel Fish, GT, and SYSTRAN), while showing variation among them, outperformed online translators based on other engines. Scoring was based on BLEU: Bilingual Evaluation Understudy, a system meant to automatize evaluations of machine translations and shown in some studies to have a high correlation to judgments
performed by humans. Babel Fish’s BLEU scores ranged from 0.2062 to 0.2807 and were more than double those found by Somers (2003), suggesting that text type and size can have a profound difference on online translator’s ability to perform — or at least the evaluation of such performance, the two not always being equal. Aiken and Wong (2006) found, for example, that in a study evaluating 20 Spanish-to-English translations of sentences selected randomly from a Spanish textbook, Babel Fish and WorldLingo were outperformed (55% grammatical and lexical accuracy) by another free OT system, SDL (75%), as judged by a formula including occurrence of missing or extraneous words, lexical choices, and total number of words versus number of correct words. Nonetheless, Babel Fish and WorldLingo were still perceived by the human evaluator to provide superior translations overall. With no generally accepted human or computer measure of translator accuracy, it is difficult to compare results from one study to another.

Aiken and Balan (2011) investigated the performance of one OT system, the above-mentioned Google Translate, by what is likely to be the broadest test across languages of online translation. GT is believed to have the biggest translation corpora in existence, with a few hundred billion words in English alone (Helft 2010). Aiken and Balan translated six sentences of various types between 2,550 language pairs: 51 languages which were matched to 50 others for translation. BLEU scores varied from a high of 100 (Spanish and Arabic), meaning all the translations should be accurate enough for a human to understand, to a low of zero for a number of language pairs (e.g. Finnish and Filipino), meaning none of the translations were found to be understandable between the two languages. While a large number of language pairs were tested, the limited and heterogenous nature of samples used in the study (translating only six sentences of
eclectic types ranging from “my hovercraft is full of eels” to a quote from the UN Declaration of Human Rights) makes it difficult to determine the applicability of the data besides underlining the variation of performance among different language pairs, a limitation mentioned by Allen (2000) and Yates (2006), among others.

Recognizing another proposed limitation of free online translators like Babel Fish, idiomatic speech, Gregory, Kellogg, and Mankin (2001) investigated the use of a go-between interface (called a “French Idiom Transducer” or FIT) to identify and render English equivalents for French idiomatic expressions that aren’t correctly translated by Babel Fish. The article explains the authors’ understanding of how Babel Fish processes input. A machine translator must have lexical, structural, and pragmatics information in both the source and target languages. The information in the input must be parsed, compared with the database, and matched with equivalent information in order to generate output with corresponding information in the target language. The authors claim that idiomatic expressions and phrasal verbs pose a particular problem since they cannot be compositionally translated from their individual parts, but rather must be identified as a whole. Doing so can be difficult due to the number of variations (conjugational, syntactic, and other) that might exist.

The authors set out to develop a “wrapper” interface for Babel Fish that could attempt to identify these expressions. This wrapper worked in two stages, first identifying possible idioms in the French input, matching it with idioms in the database with the expected (incorrect) translation from Babel Fish. In the second stage, the wrapper application looked at Babel Fish’s output. If the output from Babel Fish matched the anticipated output, then the FIT application would identify the text as an idiom and
replace that portion of the text to give the appropriate translation. In their study, a native speaker of French wrote a paragraph using some of the idioms and the paragraph was entered into the FIT interface to see if the interface (working with Babel Fish) would parse and translate them correctly. Five out of the six idioms were translated correctly, while one was not identified by FIT. The authors conclude that it would be possible to develop an application to improve Babel Fish’s translations by identifying and correctly rendering idioms and phrasal verbs missed by Babel Fish, but that due to the complexity and high number of such idioms, it would be difficult to include all of these in a database or have an application productively identify and translate new ones, which may explain why SYSTRAN, the maker of Babel Fish, does not appear to have developed a robust OT system to translate idioms that have at least in some cases been translated incorrectly for over a decade.

The possible use of OT that concerns us most is its possible application among L2 learners. SYSTRAN explained early on that “machine translation was never meant to teach language” while noting nonetheless that “students are using it to do their foreign language homework” (Yang and Lange 1998). An issue to be considered is whether or not L2 students realize the approximative nature of OT, which would fall into category three of Melby’s paradigm for types of MT translation (dynamic general text and indicative output) or if student expectations run closer to that of the fourth category (dynamic general text and high-quality output). This is perhaps why automatic translators have come as a disappointment for instructors and students alike.

One can apply a point made by Kay (1980) about computational linguists and their failed attempts to design effective translators for the L2 users who utilize these tools
as well; results can be unsatisfying and possibly even be more time-consuming “when we attempt to use computers to do something we do not really understand.” (p. 5) L2 learners who do not understand the workings or vocabulary of a language and yet try to use today’s translators are bound to end up with a final product not up to their teacher’s or their own expectations. It is important to consider as well the many pedagogical implications of students using a tool which may not be beneficial to their progress.

Luton (2003) starts from the point of view that translators are inherently harmful to speakers writing in a foreign language. As discussed earlier (based on Hurman and Tall 2002) it has been suggested that students in general pay more attention to meaning than to linguistic form and may not be aware of grammatical or lexical errors made by the translator, which might therefore have negative effects on the quality of work produced by students using online translators. Luton considers using an online translator for writing an assignment as cheating and suggests dictionaries be allowed, but not translators. The author underlines the fact, however, that it can be difficult to tell just when a student has used an online translator, especially for beginners and intermediate students.

Luton proposes however three “red flags” that can indicate when a student has used a translator (since these mistakes may also occur independently of online translator usage, it might be more appropriate to call these “yellow flags” rather than “red flags”):

- mistranslated idioms (“I get upset” becomes for example “J’obtiens le renversement”, or literally “I obtain the turning-over”)
- interspersed English words, especially misspelled ones (since translators do not translate misspelled or unknown terms)
McCarthy (2004) explores the implications OT has on second language learners. While the focus of the article is mostly on the issue of using online translators specifically for translation courses, there are a number of points made that can be considered in the present discussion. McCarthy mentions several ways in which translators like Babel Fish can be beneficial in the classroom. Students can be encouraged to consider the differences between languages and how humans and machines deal with language, as well as gain first-hand experience in how important context and cultural background are to expressing oneself in the target language. Similar goals for training translators are suggested elsewhere, with Somers (2003) suggesting learning about weaknesses and strengths of MT should be part of both foreign language and translator curricula, so these uses are worth considering.

In his translation courses, McCarthy (2004) presents three main demonstrations concerning online translators: gisting, in which students can see how well online translators succeed or fail in expressing the general idea of a passage; translation traps, in which students see how common expressions in one language cannot also be translated word-for-word and using the same syntax as the target language (e.g. the first ten pages → les dix premières pages, I am hungry → j’ai faim); and “ping-pong translation” (p. 8), in which students see the incomprehensibility of a text translated from English to French to English to French, a sort of double “backward translation” as used by Richmond (1994) among others.
While discussing instructional disadvantages of student use of online translation, McCarthy recommends mentioning to the class issues such as plagiarism/cheating, a lack of intellectual progress from not puzzling out the language on one’s own, and wasted time for the instructor grading and assessing whether or not mistakes were made by a human or a machine. McCarthy asked class participants to offer strategies for instructors to deal with online translator-assisted work handed in. Among the 12 suggestions cited, of note for the purposes of the current discussion are grading work produced or aided by an online translator on its merits (which normally carries with it the assumption that the work produced with the help of the online translator will be worse), imposing full academic sanctions on anyone caught using an online translator (which assumes that such work would be readily distinguishable from that produced without a translator’s aid), having regular lessons on online translators’ deficiencies (which it is argued would reduce time spent on other, more important topics), and submitting a completed first draft of all tasks in-class to avoid access to online translators (which would disadvantage students who work more slowly and inhibit further development of assignments in the final draft). McCarthy concludes that there is no one-size-fits-all solution to the issue of online translation, but that teachers should weigh the issues and solutions suggested to develop their policies concerning online translation.

While McCarthy focuses on translation students, Stapleton (2005) and Williams (2006) discuss the use of online translators among more general L2 student populations from the point of view of ethics, accuracy, and pedagogy. Stapleton conducted a study among 43 Japanese undergraduate students who were taking English as a foreign language to see what factors went into students’ decisions to use various online resources.
Although the study’s main focus was on the participants’ choice of online websites as sources of information for academic writing, one of the self-report questions at the end of the study asked students if they had chosen to use an online translator to assist in writing their composition. Most participants did not report doing so, but six of them admitted in a follow-up interview that either they personally, or other students, use OT to translate either individual words (generally the “stronger” students, according to the author) or longer strings of text (“weaker” students, based on their grades in the course). Because the questionnaire was anonymous, no analysis of the relationship between the scores of the participants and their self-reported use of online translators was reported. Stapleton however criticizes OT usage, encouraging teachings to stress the “unethical and damaging nature of translations and electronic ‘lifting’.” (p. 187)

Williams (2006), on the other hand, argues that encouraging teachers to present the topic of OT in class will allow both instructors and students to understand better what online translators do, how language works, and how online translators can be evaluated as being beneficial or detrimental to writing. Four main French structures (prepositions, adjectives, nous, and verbs/verb phrases) were analyzed by Williams to investigate the accuracy of English to French translations by three different online translators: AltaVista Babel Fish, Google Language Tools, and Free Translation. Williams found that with geographical prepositions, all three tools were generally accurate with translating “to/at/in” but had many errors translating the preposition “from” with place names. For adjectives, online translators generally produced correct number and gender agreement for adjectives adjacent to the nouns they modify, while remote adjectives lacked such agreement. Adjective placement was generally correct, even for adjectives that change
their meaning depending on their position (e.g. *ancien*, meaning “former” or “ancient”). Nouns after the verb *jouer* (“to play”) were also investigated, with online translators faring better with sports than musical instruments both in terms of preposition usage and vocabulary. As concerns verbs, the translators did not generally do well with translating “to have just” (*venir de*) or the idea of “ago” (*il y a*), but did well with the two verbs corresponding to the verb “to know” (*connaître* and *savoir*) when dealing with non-pronominal, post-verbal objects. Verbs with particles (e.g. “to wake up”) also have mixed results, and the position of the particle can affect whether or not the translator correctly renders the meaning into French (e.g. “to wake up” being translated as *se réveiller* versus *réveiller vers le haut*). No one online translator consistently outshone the others, with Babel Fish and GT providing the same correct or incorrect translations as each other in all cases mentioned in the article, while Free Translation would occasionally offer a different translation than the others, sometimes failing where the others had succeeded, or vice versa.

For Williams, “free, online versions of most software produce inaccurate, unacceptable translations” (p. 567). In view of this, demonstrating the sorts of errors produced by OT can be a useful language-learning experience, emphasizing to students the fact that speaking or writing in another language is not just “a matter of plugging words into a formula” (p. 372). Williams argues that a lesson in online translation can also be used for the broader goal of electronic literacy. A sample electronic literacy lesson is provided by Williams focusing on things such as noticing important information based on placement and font size, as well as the presence of links to translation tools. Instructors can have students go to search portals looking for online translators and then
have them discover cautionary notes on some sites indicating the rough or inexact nature of online translations (the equivalent of teaching consumers to look for the fine print).

Such a lesson can also serve the goals of language pedagogy, helping students understand through examples of correct or erroneous translations of English words and phrases into French concepts such as polysemy and lexical or structural ambiguity, while breaking the common misconception held by some novice language learners that there exists a one-to-one correspondence between languages. Having students use online translators for such an assignment could help them realize that, according to Williams, online translators are not intended to help second language learners produce comprehensible text in a target language, but rather to allow people unfamiliar with a foreign language to understand some basic information that would otherwise be inaccessible to them.

While Williams (2006) introduces some pertinent observations and discusses several facets of the online translation issue, important details are omitted concerning the investigation into online translator accuracy that could provide crucial opportunities for further analysis and discussion. A few interesting examples are provided, but the reasoning behind the choice of vocabulary and structural features Williams tested are not explained, nor are concrete figures given indicating how many tokens were used or the success or failure rate of online translators, other than Williams’ claim that the three translators studied produced the same translation in over 90% of cases. Williams nonetheless provides a useful basis for discussion and further investigation into the issues surrounding online translation usage among foreign language learners.
Some beginning research has been done into OT, and several authors give recommendations about what role, if any, they believe online translators should be given in the L2 classroom. No study however has been found specifically investigating compositions written by L2 students while using an online translator. Given the free availability of online translators to any student with access to the Internet via computer, smartphone, tablet, or other means, the question of the effect of OT usage among students warrants further investigation and is the main goal of the present study.

2.6 Conclusion

Online translators, although available for use by the general public for well over a decade and presumably in use by L2 students for roughly as long, are just now starting to be investigated more in the literature. When online translators have been discussed, they have generally been accorded only a passing glance, leaving many questions unanswered. While each individual article or study has its drawbacks, together they provide a basis for a more in-depth look at the effects of online translators on the writing of L2 learners.

Studies looking into the nature of L2 writing and reading generally indicate that there is some relationship between the two. While some results are in line with the reading hypothesis of Krashen (1993), suggesting that reading influences writing, others follow Shanahan and Lox (1988) and the interactive model, whereby reading and writing influence each other. Still other studies (e.g. Hedgcock and Atkinson, 1993) show no relation between L2 reading and writing. Because there is no consensus on this matter, the current study examined the L2 reading level of participants in an attempt to account for any possible effects on their L2 writing.
In Zamel (1982) we saw that none of the students questioned report placing importance on linguistic considerations such as morphology, syntax, or even lexicon, but instead focused on meaning. The student deemed to be the best writer admitted to writing in the L1 and subsequently translating her thoughts into the L2. It is quite possible that students aware of the presence of online translators may follow the same strategy, typing a text in the L1 and placing it into the online translator without much consideration as to the linguistic features involved, either thinking the computer has taken care of this for them, or simply not considering the matter. The theoretical conceptualization of writing presented in Luoma and Tarnanen (2003) also places form in a secondary position behind meaning in terms of writing priorities.

The effect of other writing aids has been documented in the literature. Even though students had problems with form, they still achieved significantly higher global scores on their compositions when allowed to use a paper dictionary, in the study outlined by Hurman and Tall (2002). This would indicate that student writing improved overall with an outside aid even if errors in form were still present. The study in this dissertation evaluated students on six different criteria to get a clearer picture of what areas, if any, are affected by online translator use.

The literature suggests that other computer usage can have an effect on students. The focus of students’ peer editing, as described in Iwai (1999), was changed after a series of CALL training sessions, with students self-reporting an increased emphasis on correcting content. Training may not always have the desired outcome however, as Scott (1990) indicates. Despite a one-hour introduction to the various functions of Système D, no students reported using the grammar, vocabulary, or phrase indices provided them in
the program. Whether due to the relatively short amount of time or the type of instruction before students used the software, training did not have the effect wished for: students relied almost exclusively on the electronic dictionary and not other writing aids. As concerns translators, Melby (1997) has suggested that those who use them may not be fully aware of the different types of text to translate (ranging from controlled domain-specific to dynamic general texts) and thus do not know what sort of results they can expect from the translator. The idea of the Naive Lexical Hypothesis (Bland et al., 1990) combined with a perceived penchant among beginning L2 students for accepting what is printed or displayed as being correct (Somers 2003) means that without proper guidance, students may accept online translator output at face value instead of as one of many resources to consider in writing. In the present study to gauge the impact of training, one group is instructed in possible strengths and pitfalls of translator usage. Results from student compositions as well as a self-assessment are used to assess what effect, if any, training has on those who use online translators in their writing.

Another aspect to consider is the readability of texts provided by online translators. A customized machine translator for Leffä (1994) produced translations that were understood equally well (no statistically-significant difference) by native speakers as translations produced by professional human translators. Leffä proposes that this is because computers translating a text may go about doing so in a way similar to that used by human readers when deciphering texts, first analyzing and then synthesizing. Luton (2003) suggests that it might be quite difficult to determine, particularly in beginners and intermediate students, whether or not an OT has been used. The study at hand is scored by instructors who were not told which texts were produced by students using online
translators and which texts were written without the aid of translators, providing insight as to whether or not the use of translators can be detected by teachers correcting compositions. An inability of scorers to tell the difference would go towards disproving the idée reçue that translators produce easily recognizable errors, while an ability to detect the difference combined with lower scores on the compositions would point to a negative impact resulting from translation usage in L2 writing. If Sabieh (2002) is correct, one might see cases in which texts created with the help of online translators are better than those created without such assistance, which would go towards supporting the idea that those using translators may be able to approach writing on a more cognitive level, with a better final product.

As discussed above (Hurman and Tall, 2002; Bishop, 2001), students unfamiliar with dictionary usage, which represent the vast majority of students, may choose the wrong word from a dictionary. There is little talk however of banning dictionaries from all work to be completed for L2 classes. It is possible that just like paper dictionaries and other writing and reading tools, both on the computer and off, online translators need to be fully understood by teachers and students in order to be used in a productive, useful manner.

It has been suggested (for example Burton, 2003) that a potential use of online translation is to check vocabulary words students are unsure of. While not about an online translator but instead an electronic hand-held dictionary, Myer (2000) analyzed student usage and found that L2 learners often looked up words in the target language to check accuracy, meaning, or usage when writing, as opposed to looking up words in the L1 to then use in the L2. Since students’ computer usage is queried via self-report in the
current study, the various ways in which L2 writers use translators can discussed in an attempt to glean information about their effectiveness.

The general attitude of many instructors (as represented by Luton, 2003) is that online translators by their very nature are harmful to L2 students; their use is also considered cheating by many. Online translators were not initially designed with L2 writers in mind, so it is important to use caution when considering or using tools outside of their initial intended use (Somers, 2003). Occasional anecdotal evidence is presented showing examples of common errors produced by online translators (mistranslated idioms, L1 words populating the L2 text, proper nouns literally translated into the L2), but no study has attempted to systematically analyze the impact online translators have on L2 writing.

Some instructors and departments are prohibiting the use of these tools without clear evidence in the literature that would merit such a position, possibly in part as a matter of principle: they may consider that work that has been done through the help of a translator is not a student’s own work, and therefore would be considered cheating even if such work may turn out to be beneficial to L2 acquisition. On the other hand, there is no compelling proof presented that online translation does have a beneficial effect on writing in an L2. The goal of the present study, based on the above mentioned articles and studies touching on related areas, is to take the first steps towards identifying what positive or negative effects, if any, L2 students’ use of online translators has on various aspects of written composition, and whether or not training can help improve student use of online translators and their subsequent scores on their compositions.
Chapter 3

Methodology

3.0 Introduction

This chapter discusses the methodology used for the present study, including participants, design, materials, procedures, and rating. The goal of this research was to investigate what effect, if any, online translator usage would have on L2 writing in French. Based on the review of the literature, the following specific research questions were developed and investigated:

• **Research question 1a:** Does the use of online translators by L2 writers result in quantifiably different global scores on compositions written in the L2?

• **Research question 1b:** Of students who use an online translator, do those who have been trained in the use of online translators achieve higher scores on L2 compositions than those who have not received such training?

• **Research question 2:** Does online translator usage affect perceived performance on any specific features of L2 learner writing, including comprehensibility, vocabulary, syntax, grammar, spelling, and content?

• **Research question 3:** Can instructors detect with statistical significance the difference between a text written with the aid of an online translator and one written without such aid?

• **Research question 4:** Do student self-reports over the positive or negative effects of online translator usage correspond to observed effects of online translator usage on L2 writing?
Since no previous experiment formally testing online translator usage among L2 students was found to inform the methodology of this study, tasks were designed in part based on prompts used by Scott (1996) in her investigation of the Système-D software described in the review of literature. The following sections detail the methods used for data collection and analysis.

3.1 Participants

The study involved a total of 34 participants from the University of Illinois at Urbana-Champaign (UIUC). Of these, the results of 32 participants were considered. All participants signed a consent form (Appendix A), approved by UIUC’s Institutional Review Board. After the study was completed and an analysis of a background questionnaire administered to all participants was performed, two participants were excluded from analysis. The background questionnaire solicited information on a variety of possibly interfering variables such as native language, time spent abroad, future plans for taking French, and exposure to other languages (Appendix B). The results of one participant were not considered since the participant reported having a native language other than English (Spanish). A second participant did not attend the final week of the study, a situation which resulted an incomplete set of data since the participant did not take the posttest. No other participants were found to have a background that would suggest a need for elimination from the study, leaving a final n of 32.

All participants were enrolled in FR 103, FR 104, or FR 106 (respectively third-semester French, fourth-semester French, and a combined second- and third-semester course) at the time of completing the study and reported having no plans of taking further
courses in the Department of French. This last prerequisite was put in place in an attempt to ensure that any positive or negative effects associated with the various treatments would not affect participants’ future academic performance in French.

Participants were recruited by the researcher through class visits, flyers, and emails sent to class sections. It was stressed to students that participation was voluntary. The instructors of these courses were also asked to encourage student participation through reminders. In order to help encourage participation, participants were offered up to 20 extra credit points for each session they attended. Since the study took place in part during class time, alternate assignments unrelated to the experiment were offered to students not participating so as not to disadvantage those who chose not to participate or who were ineligible to do so due to plans to pursue further coursework in French.

3.2 Experimental design

The following experimental groups were established for this study:

- **Group A**: no translator access and no translator training (control group)
- **Group B**: translator access and no translator training (experimental group 1)
- **Group C**: translator access with translator training (experimental group 2)

This set-up controlled for translator access (TA) as well as for translator training (TT) among participants using a translator in an attempt to make sure variables possibly affecting the results could be analyzed and taken into account. A group with no translator access (-TA) but access to training (+TT) was not established due to the limited \( n \)-size of
the participant pool and the expectation that this condition would yield no significant difference in relation to the control group, since the such participants would have no subsequent access to an online translator for the study.

Participants were randomly assigned to one of the three groups. In all, there were 10 participants in Group A (-TA, -TT), and 11 each for Group B (+TA, -TT) and Group C (+TA, +TT). Participants were told via email after the written pretest and before the first experimental task which group they would be in. This was done in part so as not to discourage some students from taking part in the study, since many potential participants had expressed an interest in using a translator and may have been less likely to participate had they known in advance they would be assigned to the control group. Recruitees had been told that they would be assigned either to a group that would be allowed to use translators or one that would not.

3.3 Materials and procedure

As stated above, students from three courses (FR 103, 104, and 106) participated in the study, for the most part during regular instructional time. Each course had one day devoted to the study every week for a total of five weeks, but due to limited lab space each course had sessions for the study on different days: for example, all FR 103 sessions were held each week on Mondays, while all FR 104 sessions were held on Tuesdays. This arrangement allowed for all tasks to be spaced out evenly for each student, regardless of the course he or she was enrolled in.

In order to finish all tasks in the allotted five weeks, most sessions were held during class time, while some sessions were held after classes ended (in the late afternoon
or early evening). Students’ instructors proctored the study for daytime session, while instructors and other teaching assistants proctored afternoon and evening sessions. All proctors were given detailed instructions to follow, were required to attend an informational session prior to the study’s start, and were trained as requested by the Institutional Review Board since they were in charge of distributing, administering, and collecting materials. The following materials and procedures were used:

- A self-report background questionnaire (Appendix B) was given the first day of the study to collect basic biographical information (including age, nationality, native language, time abroad) and detect previous classroom Internet usage, including online translators, in French or any other L2s. The questionnaire also collected data to attempt to gauge motivation to learn French as a function of field of study, future plans if any to use the language, reading habits of students in their L1 and L2, and a self-assessment of motivation towards class. A combination of open-ended and Likert-type questions was used in this questionnaire. Besides the one case mentioned in the previous section, no data from this questionnaire were deemed to warrant a participant’s exclusion from the study.

- A first pretest (Appendix C) for proficiency in reading comprehension was also administered the first day of the study. This pretest took the form of an externally-designed exam — an excerpt from the practice version of the College-Level Examination Program (CLEP) for French — used for the purposes of evaluating student reading level. Participants were given 30 minutes to complete the pretest. The CLEP was chosen since it is a standardized test which includes a reading comprehension component divided into three subsections: discrete sentences, short cloze passages, and reading passages and
authentic stimulus materials. Since different L2 studies in the literature have used one or more of these different task types, it was believed that choosing an instrument that measures these three areas would allow for better comparison with a wider variety of studies in the literature than using one of these task types alone.

The practice, paper CLEP was chosen over the computerized version for logistical reasons since only the former could be given via paper and pencil to study participants on-site during normal class time, which accommodated both proctor and participant schedules to ensure a higher participation rate. The CLEP was chosen over the UIUC departmental proficiency exam in French since the later is required of some incoming students and it was possible that some participants had taken this exam while others had not.

The results of the CLEP Pretest were not used to categorize students into groups a priori, but rather as a means to further interpret data on the four compositions received from the experimental and control groups to check for any possible correlation between reading level and writing ability, both with or without online translator usage.

• A second pretest (Appendix D), this time for written proficiency, was administered the same day as the first pretest, but in the evening for 30 minutes. While not determining in which group students were placed for this study, participant scores from the Written Pretest were analyzed in an attempt to see what role initial written proficiency might play in the effectiveness of online translator usage. The pretest was adapted from the tasks used in Scott (1990, 1996), in which students were asked to write two short paragraphs on a given topic (describing and introducing themselves).
In the pretest designed for this study, no reference to *Système D* modules was included since they were not used in this study. No translators were allowed for the pretest in order to first measure all students’ ability in writing in the L2 without access to online translation. Since Scott’s tasks were aimed at novice learners, the pretest was written with topics that correspond more appropriately to the intermediate-low level. These prompts were written after consulting the 2001 ACTFL written proficiency guidelines as well as subject matter and vocabulary covered in the third edition of *Vis-à-Vis*, the textbook currently in use at the time of the study. Dr. Alice Omaggio Hadley, who collaborated on the original ACTFL guidelines and co-authored *Vis-à-Vis*, was a member of the preliminary examination committee that reviewed the writing prompt for the Written Pretest as well as those for the three remaining writing tasks.

Participants were each provided a sheet (Appendix J) that listed key combinations for typing French accents on Windows PC. Upon completion of the composition, participants returned the prompt and accent sheet to the proctor and were instructed to submit their compositions via email attachment to the researcher.

- Each participant attended one of two *instructional sessions*, depending on the condition to which they were assigned.

Participants in the +TT condition (Group C) participated in a **50-minute translator training session** (Appendix K). The training was in part based on the recommendations of Burton (2000) and the discussion of Luton (2003), highlighting for students some potential strengths and shortcomings inherent in online translators. Babel Fish had originally been chosen for the experimental tasks because it was the OT system most discussed in the literature. However, due to Babel Fish being changed to a new
service (from AltaVista to Yahoo!) shortly before the start of the current study, a decision was made to switch to Free Translation (http://freetranslation.com) since Babel Fish was – and as of writing still is – not correctly processing some inputted text containing apostrophes.

Through a combination of teacher-led and individual activities, students were presented information about the difference between dictionaries and online translators; viewed and inputted into an online translator a series of sentences in French and English, some in Babel Fish and others in Free Translation; were asked to judge the correctness of the translations and note out any perceived errors; and were shown a list of potential strengths and weaknesses of online translators (Appendix L). Some information about translator weaknesses, such as changes in font style or quotation marks, were intentionally excluded to see if such inconsistencies would appear in translations and be noticed by raters.

Participants who did not receive translator training instead attended a 50-minute cultural session about the professions of translation and interpretation (Appendix M). Students read aloud paragraphs presenting general information describing these fields, followed by questions answered individually by participants and a group discussion by proctors with participants. The Cultural Lesson acted as a control to the Training Session and was included to ensure that the total time participants from Groups A and B spent in the study was comparable to that of Group C.

- On the third and fourth weeks of the study, participants completed a writing task (Appendices E and F) which were assigned in order to gauge what effect, if any, translator usage and training had.
Directly before each writing task, a **20-minute pre-writing activity** (Appendices H and I) was included to simulate the writing process students were normally asked to follow when writing compositions for a grade. The goal behind such activities is to activate schemata related to the task writers are to complete afterwards. Once the pre-writing was completed, proctors collected these prior to participants starting the composition task. This was done so that participants did not refer back to any French text, either provided or written by the students themselves, appearing on the pre-writing sheets.

Groups were given identical composition prompts except that those from Group A (-TA, -TT) included instructions reminding those participants that they were not allowed to use an online translator, while participants from Groups B and C were given instructions stating, in part: “You are to use the translator at [http://freetranslation.com](http://freetranslation.com) to help you in writing your composition.” Participants were not provided the topic of these tasks prior to writing. The composition prompts for Tasks One and Two were also adapted from those used in Scott (1990, 1996) for use in *Système D*. Students were told to write three short paragraphs on a given topic. As with the Written Pretest, the level of prompt was adjusted to correspond to the intermediate-low level of proficiency. Additionally, since online translators do not contain additional modules for grammar, vocabulary, etc. mentioned in Scott’s tasks reference to *Système D*, no reference was made to these tools in the tasks’ directions for this study.

All compositions, including the two tasks and written pre- and post-tests, were written by participants on Windows PCs running Office 2007. After each task was completed, participants submitted it via email attachment to the researcher.
• Following the two written experimental tasks, a posttest for written proficiency (Appendix G), equivalent in form and procedure to that of the pretest but with a different prompt, was administered to all participants to determine any change in writing proficiency at the end of the study. As with the pretest, no participants were allowed to access the online translator. This posttest was included to see whether participants’ scores in any of the conditions were significantly different from both those in the initial Written Pretest as well as the two experimental tasks. An increase in scores among those who had used the translator during Tasks One and Two might point towards a possible positive effect of translator usage that could be measured even once such access were no longer available, while lower scores could give an indication that participants had become reliant on the translator. While these considerations were not primordial to the current study, it was felt they could inform future research avenues.

• Finally, a 20-minute self-report exit questionnaire (Appendix N) consisting mostly of open-ended, short-answer questions was given to participants to collect their feedback on their performance during the study as well as on the study itself. Specific attention was paid to the comments of those who underwent translator training to see whether or not they believed instruction on potential strengths and weaknesses was beneficial to their translator usage, as well as all those with or without training who used online translators to see whether or not they believed the translator helped them in their writing and in what ways and to what extent they utilized this tool.
3.4 Rating

Each composition was evaluated by six French-speaking raters, all of whom were native English speakers with experience teaching language courses in the Department of French at Illinois. The ratings of two native-speaking Francophone scorers were not included in the analysis due to the fact that one rater did not use the scale as directed (giving half points instead of integer scores on some papers) and the other was unable to complete rating. Other French-native scorers with experience as teaching assistants in the Department were not available to rate, so only the ratings of the English native speakers were retained for the purposes of analysis.

It is impossible to have raters with identical backgrounds, but an effort to gauge any important rater differences was made both at the qualitative and quantitative levels. For a qualitative look, each rater reported their background in French and instruction through an email questionnaire. Additionally, the acting Director of Basic Language Instruction was interviewed and asked to give impressions on each rater. The abilities of all participating raters in instruction and French proficiency were judged to be high or very high by the acting Director of Basic Language Instruction based on his professional experiences with them, either as their supervisor (as Director or Course Coordinator), as a colleague, or impressions given to him by those who had observed their teaching. All raters had at least 11 years experience speaking French, all had studied French for eight semesters as an undergraduate, and all had spent between one semester and three and a half years living in a French-speaking region (France, Belgium, and/or Quebec). While teaching experience varied from two to seven years at various institutions in the United
States or abroad, all raters underwent required orientation in the Department and had taught for at least one semester at the University of Illinois.

Additionally, in an attempt to increase interrater reliability for more meaningful results, each rater was required to attend a one-hour training session that presented an overview of OT, the procedures of the study, and the rating system (Appendix P) they were to use. A rubric was devised specifically for this study covering aspects that a close review of related literature had suggested may come into play, namely content (Luoma and Tarnana 2003), comprehensibility (Leffa 1994), syntax (Watters & Patel 2000), vocabulary (Scott 1996) and spelling (Iwai 1999), as well as remaining grammar. Each rater was provided with this rubric as well as a copy of the four prompts, with references to task name and translator use deleted so as not to affect perceptions of translator use.

Those rating the essays, in addition to providing component and global scores, also indicated whether or not they believed the student who wrote the paper had access to a translator while writing. Raters were instructed to score compositions based on the merit of each of the components and not based on whether they suspected online translation use. While it is maintained by some that it is easy to spot if an essay has been written with the use of an online translator, it is not believed that this common wisdom had been formally tested in the field: it is possible, for example, that some tasks not written with the help of an online translator could be falsely identified as translated (which in the current study will be referred to as a Type I error) or that some cases of translator usage could go undetected (a Type II error).

Compositions were printed and presented to raters for scoring as-is except for the following: (1) a participant number code was placed at the top right of each composition
so that scores could be matched up by the researcher to the participant (2) any proper names mentioned in the composition (either that of the participant or those of friends or acquaintances) were replaced by placeholder names to ensure confidentiality and anonymity. The fictitious names chosen had not been used in any of the original tasks but matched the gender of the original (so as not to add any agreement error not already present in the original) as well as matching the type of initial letter (consonant, vowel, or mute h) so as not to affect elision or liaison.

The total number of tasks completed (128 three-paragraph compositions for the Written Pretest, Tasks One and Two, and Posttest combined) made impossible the ideal of all essays being scored by each rater. Each composition was instead scored by two raters and a quantitative measure — interrater reliability — was performed to assess with what confidence scores could be considered for analysis. Results from the CLEP reading comprehension pre-test were also checked for possible correlation with the written proficiency results. Finally, the responses of the scorers concerning the assumed use or non-use of online translators in the compositions they evaluated were compared to actual translator usage in order to determine if this might be detectable among instructors.

3.5 Conclusion

Due to the absence of research dealing specifically with online translator usage in L2 learners, it was difficult to foresee any clear patterns for the current study. Looking at peripheral evidence and discussion in the literature, however, it was speculated that there would likely be a difference between the scores and the types of errors made by those who do use an online translator and those who do not, based on Hurman and Tall
(2002)’s study on dictionary usage and L2 writing, as well as discussion in the literature (e.g. Luton, 2003) about errors made by online translators. If any positive or negative effect were to be found on L2 writing scores among those who wrote with the aid of online translation, it was expected that those who had been trained in translator usage would receive higher scores on the tasks than those receiving no training, based in part on the results mentioned in Bishop (2001) concerning higher scores for those trained in dictionary usage. Whether instructors would be able to detect the difference between a text written with or without the aid of an online translator was another matter considered. Leffa (1994) found that native speakers did not have any more trouble comprehending texts written entirely by machine translators versus those written by professional human translators. It was unclear if this would apply to non-native instructors evaluating L2 student texts written with the aid of an online translator, but anecdotal evidence from teachers as well as discussion of “red flag” indicators in Luton (2003) would rather tend to point towards an ability, at least in some cases, to distinguish between a text written by a student using an online translator and one written without a translator.
Chapter 4

Results

4.0 Introduction

In this chapter, the data collected through the different tasks, tests, and questionnaires, as well as quantitative analyses of these data, are presented. The results for the CLEP reading pretest are given first, followed by an analysis of interrater reliability; a detailed look at rater subscores (for each of the six features rated) and global scores (the sum of the subscores) for the Written Pretest, Tasks One and Two, and the Posttest; a presentation of the accuracy of rater judgments on whether or not a translator was used for the compositions they scored; and finally, a quantitative look at select participant responses on the background and exit questionnaires. A detailed discussion of these findings, as well as a discussion of more qualitative aspects of data and comments obtained from participants and raters, will occur in Chapter 5.

4.1 CLEP Pretest results

All participants took a pencil-and-paper version of a sample CLEP reading proficiency exam (Appendix C). The scores were analyzed using the GLM procedure. As shown in Table 4.1a, a one-way ANOVA performed for the results of Groups A (-translator, -training), B (+translator, -training) and C (+translator, +training) showed no significant difference among the groups ($F(2, 29) = 0.06, p = 0.9381$). As indicated in Table 4.1b and the boxplot in Figure 4.1 below, the overall mean was 15.375 out of 35 possible questions. There was no difference in mean scores ($\bar{x} = 15.182$) between
experimental groups B and C while the mean score for Group A ($\bar{x} = 15.800$) was only slightly higher.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>2.6272727</td>
<td>1.3136364</td>
<td>0.06</td>
<td>0.9381</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>594.8727273</td>
<td>20.5128527</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>597.5000000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1a One-way ANOVA results for CLEP Pretest scores

<table>
<thead>
<tr>
<th>Level of Group</th>
<th>N</th>
<th>Score</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>10</td>
<td>15.8000000</td>
<td>3.48966729</td>
<td></td>
</tr>
<tr>
<td>Group B</td>
<td>11</td>
<td>15.1818182</td>
<td>4.46806853</td>
<td></td>
</tr>
<tr>
<td>Group C</td>
<td>11</td>
<td>15.1818182</td>
<td>5.34449589</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1b CLEP Pretest score means across groups

Figure 4.1 CLEP Pretest score distribution by group

It should be noted that the standard deviation between groups was relatively high (SD = 3.490 for group A, 4.468 for group B, and 5.345 for group C). This may be due in part to the small n-size for each group ($n = 10$ for group A and $n = 11$ each for Groups B and C). However, given the fact that the means were not found to be statistically significantly different among the three groups, we can say that the CLEP results support the assumption that the reading comprehension level of the three groups overall was
comparable at the start of the study. Another test designed to determine the similarity of
the groups prior to treatment, the Written Pretest, will be discussed shortly.

4.2 Interrater reliability

All remaining tasks (Written Pretest, Task One, Task Two, Posttest) relied on
rater scores. The 128 compositions collected from participants were each scored
separately by two of the raters, for a total of 256 ratings for global scores. Since it was
not feasible for all raters to score all compositions, each of the six raters was asked to
score 42 or 43 compositions. In this way, each composition could be rated twice in an
attempt to ensure the results were reliable.

Kendall’s tau-b and Spearman’s rho tests were run to determine the level of
interrater reliability between the two global ratings given to all compositions. The results
for these correlations (Table 4.2a) showed that there was a significant relationship
between the two raters’ scores across all composition tasks, \( r_s = .700, p \text{ (two-tailed)} < 0.01 \)
and \( \tau = .556, p \text{ (two-tailed)} < 0.01 \). Based on these tests, there was a high correlation
between the ratings given to compositions, suggesting the scores obtained are reliable.

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Rater 1</th>
<th>Rater 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kendall's tau_b</td>
<td>AllRater1Tot</td>
<td>Correlation Coefficient</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>AllRater2Tot</td>
<td>Correlation Coefficient</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>128</td>
</tr>
<tr>
<td>Spearman's rho</td>
<td>AllRater1Tot</td>
<td>Correlation Coefficient</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>AllRater2Tot</td>
<td>Correlation Coefficient</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>128</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.01 level (2-tailed).

Table 4.2a Interrater reliability results for all compositions
To further confirm reliability of ratings, Kendall’s tau-b and Spearman’s rho tests were also performed for each composition task (Pretest, Task One, Task Two, Posttest). There was a significant relation between ratings for each of the compositions as well: the Pretest ($r_s = .747$ and $\tau = .584$), Task One ($r_s = .752$ and $\tau = .605$), Task Two ($r_s = .586$ and $\tau = .458$), and the Posttest ($r_s = .680$ and $\tau = .535$), with each of the tests run being two-tailed and with correlations significant at the $p < 0.01$ level (Tables 4.2b through 4.2e).

The scores for the compositions can be said to be reliable since there was a significant correlation between scores given independently by two raters.

**Correlations**

<table>
<thead>
<tr>
<th>Kendall's tau_b</th>
<th>PreRater1Tot</th>
<th>PreRater2Tot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlation Coefficient</td>
<td>PreRater1Tot</td>
</tr>
<tr>
<td>Kendall's tau_b</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>PreRater1Tot</td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>391.875</td>
</tr>
<tr>
<td>PreRater2Tot</td>
<td>Correlation Coefficient</td>
<td>12.641</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>.730</td>
</tr>
<tr>
<td>Spearman's rho</td>
<td>PreRater1Tot</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correlation Coefficient</td>
<td>374.938</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>12.095</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>21.725</td>
</tr>
<tr>
<td>PreRater2Tot</td>
<td>Correlation Coefficient</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>.730</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).**

**Table 4.2b Interrater reliability results for Pretest**

**Correlations**

<table>
<thead>
<tr>
<th>Kendall's tau_b</th>
<th>T1Rater1Tot</th>
<th>T1Rater2Tot</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlation Coefficient</td>
<td>T1Rater1Tot</td>
</tr>
<tr>
<td>Kendall's tau_b</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>T1Rater1Tot</td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>405.969</td>
</tr>
<tr>
<td>T1Rater2Tot</td>
<td>Correlation Coefficient</td>
<td>13.096</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>.628</td>
</tr>
<tr>
<td>Spearman's rho</td>
<td>T1Rater1Tot</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correlation Coefficient</td>
<td>292.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>9.419</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>17.161</td>
</tr>
<tr>
<td>T1Rater2Tot</td>
<td>Correlation Coefficient</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>.628</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).**

**Table 4.2c Interrater reliability results for Task One**
Correlations

<table>
<thead>
<tr>
<th>Kendall's tau_b</th>
<th>T2Rater1Tot</th>
<th>T2Rater2Tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2Rater1Tot</td>
<td>Correlation Coefficient</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>300.875</td>
</tr>
<tr>
<td>T2Rater2Tot</td>
<td>Correlation Coefficient</td>
<td>9.706</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>.603</td>
</tr>
<tr>
<td>Spearman's rho</td>
<td>T2Rater1Tot</td>
<td>Correlation Coefficient</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>261.875</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>8.448</td>
</tr>
<tr>
<td></td>
<td>T2Rater2Tot</td>
<td>Correlation Coefficient</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>32</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

Table 4.2d Interrater reliability results for Task Two

Correlations

<table>
<thead>
<tr>
<th>Kendall's tau_b</th>
<th>PostRater1Tot</th>
<th>PostRater2Tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>PostRater1Tot</td>
<td>Correlation Coefficient</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>32</td>
</tr>
<tr>
<td>PostRater2Tot</td>
<td>Correlation Coefficient</td>
<td>.535</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>32</td>
</tr>
<tr>
<td>Spearman's rho</td>
<td>PostRater1Tot</td>
<td>Correlation Coefficient</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.680</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>PostRater2Tot</td>
<td>Correlation Coefficient</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>32</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

Table 4.2e Interrater reliability results for Postest

4.3 Written Pretest results

A second measure of the participants’ level, a three-paragraph composition task (Appendix D), was also conducted at the start of the study before treatment and experimental tasks. The Written Pretest was administered to check for differences among participants’ initial writing level prior to these; for this reason, no participants were allowed to use online translators. A one-way ANOVA was run on the means of the global scores of Groups A, B, and C for each composition. Since each composition was scored
by two raters, the average of the two scores was used for analysis. As shown in Table 4.3, no significant difference was found among groups ($F (2, 29) = 0.41, p = 0.6690$). Figure 4.2 shows that Group A ($\bar{x} = 22.250$ out of 30 possible points) scored over a point higher on average than Group B ($\bar{x} = 20.955$) or C ($\bar{x} = 20.864$); however, the difference was not enough to be significant at the $p < 0.05$ level. This indicates participant writing overall was statistically similar among groups prior to the experimental tasks.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>12.4069602</td>
<td>6.2034801</td>
<td>0.41</td>
<td>0.6690</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>441.3977273</td>
<td>15.2206113</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>453.8046875</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3 One-way ANOVA results for Written Pretest scores

![Figure 4.2 Written Pretest score distribution by group](image)

4.3.1 Written Pretest subscores

The global composition scores represent the sum of the six subscores (Overall Comprehensibility, Content, Spelling, Syntax, Remaining Grammar, and Vocabulary). One-way ANOVA tests were also run for each of the six categories to see if any difference would be found among the groups. Importantly, no significant differences
were found among the groups for any of the six categories at the $p < 0.05$ level: Overall Comprehensibility ($F (2, 29) = 0.02, p = 0.9762$), Content ($F (2, 29) = 2.40, p = 0.1085$), Spelling and Accents ($F (2, 29) = 0.67, p = 0.5206$), Syntax ($F (2, 29) = 0.76, p = 0.4756$), Remaining Grammar ($F (2, 29) = 0.20, p = 0.8235$), Vocabulary ($F (2, 29) = 1.36, p = 0.2714$). Although Group A had higher means than both Groups B and C in four out of the six categories, the difference never reached statistical significance. These results are presented in Tables 4.4a through 4.4f. The findings show that not only were groups not significantly different in their overall writing ability, but also not significantly different in any of the six components raters scored, suggesting similarity among groups prior to treatment and experimental tasks.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>0.04900568</td>
<td>0.02450284</td>
<td>0.02</td>
<td>0.9762</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>29.44318182</td>
<td>1.01528213</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>29.49218750</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4a One-way ANOVA results for Written Pretest: Overall Comprehensibility

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>1.84147727</td>
<td>0.92073864</td>
<td>2.40</td>
<td>0.1085</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>11.12727273</td>
<td>0.38369906</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>12.96875000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4b One-way ANOVA results for Written Pretest: Content

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>0.94090909</td>
<td>0.47045455</td>
<td>0.67</td>
<td>0.5206</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>20.43409091</td>
<td>0.70462382</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>21.37500000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4c One-way ANOVA results for Written Pretest: Spelling and Accents

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>0.66818182</td>
<td>0.33409091</td>
<td>0.76</td>
<td>0.4756</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>12.70681818</td>
<td>0.43816614</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>13.37500000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4d One-way ANOVA results for Written Pretest: Syntax

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>0.27272727</td>
<td>0.13636364</td>
<td>0.20</td>
<td>0.8235</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>20.22727273</td>
<td>0.69749216</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>20.50000000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4e One-way ANOVA results for Written Pretest: Remaining Grammar
<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>1.45965909</td>
<td>0.72982955</td>
<td>1.36</td>
<td>0.2714</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>15.50909091</td>
<td>0.53479624</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>16.96875000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4f One-way ANOVA results for Written Pretest: Vocabulary

### 4.4 Task One results

After participants attended either a training session concerning the potential strengths and weaknesses of online translation (Appendices K and L) or a cultural lesson about the field of translation (Appendix M), each participant wrote a three-paragraph composition. For Task One (Appendix E), Group A did not have access to an online translator, while Group B (which had no prior training in translator usage) and Group C (which had received such training) were given access to and prompted to use a translator during the writing process.

A one-way ANOVA was run for Groups A, B, and C on the means of the two raters’ scores for each composition. As shown in Table 4.5, no significant difference was found among the groups \( F(2, 29) = 2.88, p = 0.0725 \). There is a notable difference between the mean score of Group A (\( \bar{x} = 20.700 \) out of a possible 30) and those of Groups B (\( \bar{x} = 23.909 \)) and C (\( \bar{x} = 23.500 \)), as illustrated in Figure 4.3, and this result approached significance at the \( p < 0.05 \) level \( (p = 0.0725) \). It is possible that the lack of statistical significance may at least in part be due the relatively low \( n \)-sizes in each group. Given these results, however, we cannot confirm there was any pairwise difference among the control group and the two groups using an online translator.
4.4.1 Task One Content subscore

In addition to global composition scores, one-way ANOVA tests were performed for each of the six categories to see if any difference would be found among the groups. Significant differences were found this time, specifically for the components Content and Remaining Grammar. Table 4.6a shows the ANOVA results for Content ($F(2, 29) = 4.74, p = 0.0165$). Pairwise analysis was performed to see which means differed significantly. Bonferroni (Dunn) t Tests for the category Content show in particular that the difference between Group A (- translator - training) and Group B (+ translator – training) was significant (Table 4.6b). Tukey's Studentized Range (HSD) and Scheffe’s tests were also conducted and similarly showed a significant difference between these groups.
Table 4.6a ANOVA results for Task One: Content subscores

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>2.84789632</td>
<td>1.42393466</td>
<td>4.74</td>
<td>0.0165</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>8.70681818</td>
<td>0.30023511</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>11.55468750</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.6b Bonferroni (Dunn) t Test results for Task One: Content subscores

As illustrated in Figure 4.4, the mean of Group A ($\bar{x} = 4.150$ out of a possible 5) is different by over half a point from those of Group B ($\bar{x} = 4.864$) and Group C ($\bar{x} = 4.682$), but only a pairwise comparison of the first two groups reaches the level of statistical significance. As can be seen in the figure, it should be noted that there is one participant in Group C that scored over a point below the mean, lowering the group’s overall mean score. Given the small n-size in each group, however, it was not considered appropriate to exclude this score from the analysis, particularly in this case since the score falls within the midrange of Group A’s scores. Based on these results, it can be
concluded that Group B significantly outperformed the control group at the $p < 0.05$ level in Content scores on Task One.

### 4.4.2 Task One Remaining Grammar subscore

An analysis of Remaining Grammar also yielded a significant difference ($F(2, 29) = 3.59, p = 0.0403$). Table 4.7a shows the results for Remaining Grammar.

Bonferroni (Dunn) $t$ Test results indicate that the means between Group A (- translator - training) and Group B (+ translator – training) differed significantly (Table 4.7b). Tukey’s Studentized Range (HSD) and Scheffé’s tests confirm a significant pairwise difference between these groups.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>2.65681818</td>
<td>1.32840909</td>
<td>3.59</td>
<td>0.0403</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>10.71818182</td>
<td>0.36959248</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>13.37500000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.7a ANOVA results for Task One: Remaining Grammar subscore

<table>
<thead>
<tr>
<th>Group Comparison</th>
<th>Difference Between Means</th>
<th>Simultaneous 95% Confidence Limits</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Group B - Group C</td>
<td>0.1818</td>
<td>-0.4769</td>
<td>0.8405</td>
</tr>
<tr>
<td>Group B - Group A</td>
<td>0.6909</td>
<td>0.0160</td>
<td>1.3658 ***</td>
</tr>
<tr>
<td>Group C - Group B</td>
<td>-0.1818</td>
<td>-0.8405</td>
<td>0.4769</td>
</tr>
<tr>
<td>Group C - Group A</td>
<td>0.5091</td>
<td>-0.1658</td>
<td>1.1840</td>
</tr>
<tr>
<td>Group A - Group B</td>
<td>-0.6909</td>
<td>-1.3658</td>
<td>-0.0160 ***</td>
</tr>
<tr>
<td>Group A - Group C</td>
<td>-0.5091</td>
<td>-1.1840</td>
<td>0.1658</td>
</tr>
</tbody>
</table>

Table 4.7b Bonferroni (Dunn) $t$ Test results for Task One: Remaining Grammar subscore

As shown in Figure 4.5 below, the mean of Group A ($\bar{x} = 2.900$ out of 5) is once again different by over half a point from those of Group B ($\bar{x} = 3.591$) and Group C ($\bar{x} = 3.409$), but only a comparison of the first two showed significance difference. As can be seen in the figure, it should be noted that there is again a participant in Group C that scored well below the mean, which might in part explain why no statistically significant
difference between Groups A and C. In view of the results, Group B outperformed the control group in Remaining Grammar scores at the $p < 0.05$ level on Task One.

Figure 4.5 Boxplot of Task One: Remaining Grammar subscore distribution by group

### 4.4.3 Task One other subscores and summary

Among the other four subscores, no significant statistical pairwise difference was found for Overall Comprehensibility ($F(2, 29) = 3.37, p = 0.6937$), Spelling and Accents ($F(2, 29) = 3.24, p = 0.0536$), Syntax ($F(2, 29) = 3.03, p = 0.0638$), Vocabulary ($F(2, 29) = 0.14, p = 0.8662$), as shown in Tables 4.8a through 4.8d. While means for Spelling and Accents (Group A, $\bar{x} = 3.550$ versus Groups B and C, each $\bar{x} = 4.227$) and Syntax (Group A, $\bar{x} = 2.900$, Group B $\bar{x} = 3.590$, Group C, $\bar{x} = 3.409$) differed, these results both approached but missed significance at the $p < 0.05$ level, possibly due in part to $n$-size.

![Distribution of Score](image)

Table 4.8a ANOVA Results for Task One: Overall Comprehensibility subscore

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>0.61150568</td>
<td>0.30575284</td>
<td>0.37</td>
<td>0.6937</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>23.94318182</td>
<td>0.82562696</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>24.55468750</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


To summarize the results for Task One, we can say that while overall scores among the three groups were not different, the categories of Content and Remaining Grammar were significantly different \((p < 0.05)\) between the control group (Group A) and Group B, which had access to a translator but no prior training. There was no significant difference among the three groups for the remaining four subscores.

### 4.5 Task Two results

For the second experimental task (Appendix F), as with the first, participants in Group A (-translator -training), Group B (+translator -training) and Group C (+translator +training) wrote compositions, and a one-way ANOVA was run across the means of the two raters’ overall scores for each group. As shown in Table 4.9a, a statistically significantly high level of difference was found \((F(2, 29) = 7.05, p = 0.0032)\). As can be seen in Figure 4.6 and Table 4.9b, the means varied among the three groups: Group A, \(\bar{x} = 20.650\) out of 30; Group B, \(\bar{x} = 23.681\); and Group C, \(\bar{x} = 25.364\), with Group C’s lower quartile being higher than the upper quartile of Group A.
Table 4.9a ANOVA results for Task Two scores

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>118.6681818</td>
<td>59.3340909</td>
<td>7.05</td>
<td>0.0032</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>244.2068182</td>
<td>8.4209248</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>362.875000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.6 Boxplot of Task Two score distribution by group

Bonferroni (Dunn) \( t \) Tests for overall scores show more specifically that a pairwise comparison between Group A (-TA -TT) and Group C (+TA +TT) reached the level of statistical difference (Table 4.8b). Tukey’s Studentized Range (HSD) and Scheffe’s tests confirmed this result. While the difference between the means for the comparison of Groups A and B was also high, it did not reach the level of statistical significance at the \( p < 0.05 \) level. These results show that Group C, which was instructed to use an online translator during writing and which had been previously trained in translator usage, outperformed at a significant level (\( p < 0.01 \)) the control group with no such access or training.

<table>
<thead>
<tr>
<th>Group Comparison</th>
<th>Difference Between Means</th>
<th>Simultaneous 95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group C - Group B</td>
<td>0.2273</td>
<td>-0.5939 - 1.0484</td>
</tr>
<tr>
<td>Group C - Group A</td>
<td>0.9045</td>
<td>0.0631 - 1.7460</td>
</tr>
<tr>
<td>Group B - Group C</td>
<td>-0.2273</td>
<td>-1.0484 - 0.5939</td>
</tr>
<tr>
<td>Group B - Group A</td>
<td>0.6773</td>
<td>-0.1642 - 1.5187</td>
</tr>
<tr>
<td>Group A - Group C</td>
<td>-0.9045</td>
<td>-1.7460 - 0.0631</td>
</tr>
<tr>
<td>Group A - Group B</td>
<td>-0.6773</td>
<td>-1.5187 - 0.1642</td>
</tr>
</tbody>
</table>

Table 4.9b Bonferroni (Dunn) \( t \) Test results for Task Two scores
4.5.1 Task Two Overall Comprehensibility subscore

In addition to global scores, one-way ANOVA tests were again performed on each of the categories to see if any difference would be found among the groups in the subscores. Significant differences were found for four of the six categories: all but Syntax and Vocabulary.

Table 4.10a shows the ANOVA results for Overall Comprehensibility ($F(2, 29) = 3.77, p = 0.0350$). Bonferroni (Dunn) $t$ Tests show a pairwise difference between Group A and Group C (Table 4.10b). Results for Tukey's Studentized Range (HSD) and Scheffe’s tests also showed a significant difference between these groups.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>4.58465909</td>
<td>2.29232955</td>
<td>3.77</td>
<td>0.0350</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>17.63409091</td>
<td>0.60807210</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>22.21875000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.10a ANOVA results for Task Two: Overall Comprehensibility subscores

<table>
<thead>
<tr>
<th>Group Comparison</th>
<th>Difference Between Means</th>
<th>Simultaneous 95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group C - Group B</td>
<td>0.2273</td>
<td>-0.6176</td>
</tr>
<tr>
<td>Group C - Group A</td>
<td>0.9045</td>
<td>0.0388</td>
</tr>
<tr>
<td>Group B - Group C</td>
<td>-0.2273</td>
<td>-1.0721</td>
</tr>
<tr>
<td>Group B - Group A</td>
<td>0.6773</td>
<td>-0.1885</td>
</tr>
<tr>
<td>Group A - Group C</td>
<td>-0.9045</td>
<td>-1.7703</td>
</tr>
<tr>
<td>Group A - Group B</td>
<td>-0.6773</td>
<td>-1.5430</td>
</tr>
</tbody>
</table>

Table 4.10b Bonferroni (Dunn) $t$ Test results for Task Two: Overall Comprehensibilty

As shown in Figure 4.7, the mean of Group A ($\bar{x} = 3.0500$ out of a possible 5) differs by nearly a full point from that of Group C ($\bar{x} = 3.9545$). The lower quartile of Group C is equal to the upper quartile of Group A (3.5). While the mean Group B ($\bar{x} = 3.7273$) differs slightly from that of Group C, and over a half point as compared to Group A, neither of these pairwise comparisons yields a statistically significant difference.
These results show that Group C (+TA +TT) significantly outperformed Group A (-TA -TT) at the $p < 0.05$ level in Overall Comprehensibility scores for Task Two.

Figure 4.7 Boxplot of Task Two: Overall Comprehensibility score distribution by group

### 4.5.2 Task Two Content subscore

There was also a significant difference for Content ($F(2, 29) = 5.23$, $p = 0.0115$).

Table 4.11a shows the ANOVA results for Content, which are statistically significant. Bonferroni (Dunn) $t$ Tests indicate that the difference between Groups A and C in particular were significant (Table 4.11b). Tukey's Studentized Range (HSD) and Scheffe’s tests also confirm the means of these groups differed.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>4.51832386</td>
<td>2.25916193</td>
<td>5.23</td>
<td>0.0115</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>12.53636364</td>
<td>0.43228840</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>17.05468750</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.11a ANOVA results for Task Two: Content subscores

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Difference Between Means</th>
<th>Simultaneous 95% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group C - Group B</td>
<td>0.3182</td>
<td>-0.3942 - 1.0305</td>
</tr>
<tr>
<td>Group C - Group A</td>
<td>0.9182</td>
<td>0.1882 - 1.6481</td>
</tr>
<tr>
<td>Group B - Group C</td>
<td>-0.3182</td>
<td>-1.0305 - 0.3942</td>
</tr>
<tr>
<td>Group B - Group A</td>
<td>0.6000</td>
<td>-0.1299 - 1.3299</td>
</tr>
<tr>
<td>Group A - Group C</td>
<td>-0.9182</td>
<td>-1.6481 - 0.1882</td>
</tr>
<tr>
<td>Group A - Group B</td>
<td>-0.6000</td>
<td>-1.3299 - 0.1299</td>
</tr>
</tbody>
</table>

Table 4.11b Bonferroni (Dunn) $t$ Test results for Task Two: Content subscores
Figure 4.8 shows a boxplot for this category. The mean of Group A ($\bar{x} = 3.900$ out of 5) was again the lowest of the three, being significantly different from Group C’s mean of 4.818 but not from Group B’s mean of 4.500. In the figure, Group C’s box does not display an upper and lower quartile since all ratings were either a 4 or 5; by contrast, only one composition received a score of 5 for group A in Content. These results show that the control group had a lower mean Content score on Task Two than Group C (+TA +TT), with this result significant at the $p < 0.05$ level.

![Figure 4.8 Boxplot of Task Two: Content subscore distribution by group](image)

4.5.3 Task Two Spelling and Accents subscore

One-way ANOVA results for Spelling and Accents ($F(2, 29) = 7.60, p = 0.0022$) showed a significant difference, nearly reaching the $p < 0.001$ level, as shown in Table 4.12a. A pairwise analysis confirmed a difference again between Groups A and C, but also this time between Groups A and B, at the $p < 0.05$ level as determined by the results of Bonferroni (Dunn) $t$ (Table 4.12b), Tukey's Studentized Range (HSD), and Scheffe’s tests.
Table 4.12a ANOVA results for Task Two: Spelling and Accents subscores

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>8.76491477</td>
<td>4.38245739</td>
<td>7.60</td>
<td>0.0022</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>16.72727273</td>
<td>0.57680251</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>25.49218750</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparisons significant at the 0.05 level are indicated by ***.

<table>
<thead>
<tr>
<th>Group Comparison</th>
<th>Difference Between Means</th>
<th>Simultaneous 95% Confidence Limits</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Group C - Group B</td>
<td>0.4091</td>
<td>-0.4138 1.2319</td>
<td></td>
</tr>
<tr>
<td>Group C - Group A</td>
<td>1.2727</td>
<td>0.4296 2.1159</td>
<td>***</td>
</tr>
<tr>
<td>Group B - Group C</td>
<td>-0.4091</td>
<td>-1.2319 0.4138</td>
<td></td>
</tr>
<tr>
<td>Group B - Group A</td>
<td>0.8636</td>
<td>0.0205 1.7068</td>
<td>***</td>
</tr>
<tr>
<td>Group A - Group C</td>
<td>-1.2727</td>
<td>-2.1159 -0.4296</td>
<td>***</td>
</tr>
<tr>
<td>Group A - Group B</td>
<td>-0.8636</td>
<td>-1.7068 -0.0205</td>
<td>***</td>
</tr>
</tbody>
</table>

Table 4.12b Bonferroni (Dunn) t Test results for Task Two: Spelling and Accents

Figure 4.9, shows that Group A’s mean ($\bar{x} = 2.600$ out of 5) was lower than those of both Group B ($\bar{x} = 3.364$) and Group C ($\bar{x} = 3.636$), with Group A’s upper quartile being equal to the median of Group B’s subscores and the lower quartile of Group C. In view of these results, both experimental groups significantly outperformed the control group for Spelling and Accents.

Figure 4.9 Boxplot of Task Two: Spelling and Accents distribution by group

4.5.4 Task Two Remaining Grammar subscore

For this category, a one-way ANOVA ($F(2, 29) = 6.94, p = 0.0034$) was run, showing a statistically-significant difference between means (Table 4.13a). Once again, a
pairwise analysis through Bonferroni (Dunn) $t$ Tests, as showed in Table 4.13b, found a significant difference ($p < 0.05$) between Groups A and B, as well as Groups A and C, with Tukey's Studentized Range (HSD) and Scheffe’s tests confirming the results.

Table 4.13a ANOVA results for Task Two: Remaining Grammar subscores

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>$F$ Value</th>
<th>Pr &gt; $F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>5.97784091</td>
<td>2.98892045</td>
<td>6.94</td>
<td>0.0034</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>12.49090909</td>
<td>0.43072100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>18.46875000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.13b Bonferroni (Dunn) $t$ Test results for Task Two: Remaining Grammar

Means for Groups B ($\bar{x} = 3.364$) and C ($\bar{x} = 3.636$) were again much higher than those of Group A ($\bar{x} = 2.600$). As shown in Figure 4.10, Group A’s upper quartile of 3.0 was equal to the lower quartile of both Group B and Group C (which for Group C also
represented the lowest score of the group). As was the case with Spelling and Accents, the control group did not perform as well in Remaining Grammar as the two groups who were allowed to use an online translator.

### 4.5.5 Task Two other subscores and summary

For the two remaining categories, there was no statistically-important difference found among the groups. One-way ANOVAs were run for Syntax \( F(2, 29) = 0.31, p = 0.7375 \) and Vocabulary \( F(2, 29) = 1.13, p = 0.3371 \), but neither of these approached significance (Tables 4.14a and 4.14b).

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>0.95511364</td>
<td>0.47755682</td>
<td>1.13</td>
<td>0.3371</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>12.26363636</td>
<td>0.42288401</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>13.21875000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.14a ANOVA results for Task Two: Syntax subscores

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>0.21818182</td>
<td>0.10909091</td>
<td>0.31</td>
<td>0.7375</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>10.28181818</td>
<td>0.35454545</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>10.50000000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.14b ANOVA results for Task Two: Vocabulary subscores

It is interesting to note that out of all the scores for both Tasks One and Two, both overall and for categories, Group A had a larger mean score \( (\bar{x} = 3.700) \) only for Vocabulary (Table 4.14c), having a very slightly higher result than Group B \( (\bar{x} = 3.682) \) but not at a level approaching statistical significance. In every other measure, Group A had lower scores than both Groups B and C, and as noted above, in many cases the difference reached a statistically-significant level. Additionally, Group C had a higher
mean in every component and overall as compared to Group B, but this difference never reached significance.

| Group   | Score LSMEAN | Standard Error | Pr > |t|  | LSMEAN Number |
|---------|--------------|----------------|------|---|----------------|
| Group A | 3.70000000   | 0.18829377     | <.0001 |   | 1               |
| Group B | 3.68181818   | 0.17953107     | <.0001 |   | 2               |
| Group C | 3.86363636   | 0.17953107     | <.0001 |   | 3               |

Table 4.14c Means for Task Two: Vocabulary subscores

In summary, Group C (+translator +training) had means that were statistically higher than the control group (-translator -training) on Task Two for global scores as well as for Overall Comprehensibility, Content, Spelling and Accents, and Remaining Grammar. The means for Group B (+translator -training) were higher than those of Group A for the latter two categories, but not overall. These findings indicate that each translator group either outperformed or did not do statistically significantly worse on Task Two than the group where participants did not use online translation.

4.6 Posttest results

A week after Task Two, participants were asked to write a final three-paragraph composition (Appendix G) in order to check for differences in writing after performing the experimental tasks. As with the Written Pretest, no groups were allowed to use online translators for the Posttest, this time to check for any effect using OT may have had on participants’ subsequent unassisted writing. A one-way ANOVA was run on the mean global scores for each composition. As shown in Table 4.15, no statistically-significant difference was found among the groups \( F(2, 29) = 0.27, p = 0.7668 \). Figure 4.11 displays the means for Groups A \( (\bar{x} = 21.600 \text{ out of a possible 30}) \), B \( (\bar{x} = 21.682) \) and C \( (\bar{x} = 22.727) \). The means for Groups A and B were nearly the same, while Group C’s
mean was over a point higher, but not different enough to reach statistical significance, perhaps in part due to a high standard deviation (SD = 5.120) compared to the other groups (Group A, SD = 2.706; Group B, SD = 3.649). These results indicate that the groups did not perform differently on the Posttest at a statistically-relevant level.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>8.5240057</td>
<td>4.2620028</td>
<td>0.27</td>
<td>0.7668</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>461.2181818</td>
<td>15.9040752</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>469.7421875</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.15 ANOVA results for Posttest scores

Figure 4.11 Boxplot of Posttest score distribution by group

4.6.1 Posttest subscores

In order to see if there were differences among the groups in any of the categories, one-way ANOVA tests were also run for each. As with the global scores, there were no significant differences found among the groups for any of the six component scores at the $p < 0.05$ level: Overall Comprehensibility ($F (2, 29) = 0.51, p = 0.6062$), Content ($F (2, 29) = 0.08, p = 0.6062$), Spelling and Accents ($F (2, 29) = 0.03, p = 0.9746$), Syntax ($F (2, 29) = 0.66, p = 0.5232$), Remaining Grammar ($F (2, 29) = 0.46, p = 0.6345$),
Vocabulary \( (F (2, 29) = 0.20, p = 0.8224) \). These results are presented in Tables 4.16a through 4.16f.

As none of these results approaches statistical significance, just as the overall scores for Posttest did not, we can conclude that performance among the three groups did not differ statistically significantly on the Posttest in terms of their overall scores or any of the categories, suggesting that use of an online translator by Groups B and C, as well as prior
training in OT for the latter group, did not adversely affect subsequent writing scores as compared to the control group once a translator was again not permitted.

### 4.7 Rater judgements on translator usage

In addition to assigning points, raters were asked to decide whether they believed each composition scored was written with the help of an online translator or without such aid. Raters were not told whether a given composition prompt was used for the Written Pretest or Posttest (on which no students were allowed to use a translator) or an experimental task (on which some students were); nor were raters told exactly how many total students had been allowed translator usage. This information was withheld since it was believed such knowledge might have influenced, knowingly or not, rater judgments.

Tests were performed to assess raters’ accuracy in identifying OT usage. As seen in Table 4.17a, there was a significant association between the raters’ judgment and actual translator usage, \( \chi^2 (1) = 28.818 \) and Likelihood ratio = 28.131, both with \( p < 0.001 \). While the results are very highly significant, it is still interesting to note that there were a number of times that raters chose incorrectly.

<table>
<thead>
<tr>
<th>Chi-Square Tests</th>
<th>Value</th>
<th>df</th>
<th>Asymp. Sig. (2-sided)</th>
<th>Exact Sig. (2-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>28.818a</td>
<td>1</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Continuity Correction(^b)</td>
<td>27.309</td>
<td>1</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>28.131</td>
<td>1</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Fisher's Exact Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear-by-Linear Association</td>
<td>28.705c</td>
<td>1</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>256</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 27.16.
b. Computed only for a 2x2 table
c. The standardized statistic is 5.358.

Table 4.17a Relationship between rater judgment of translator use and actual use overall
As seen in the crosstabulation below (Table 4.17b), raters correctly identified compositions as not being written with translator access 135 times, and as being written with translator access 46 times. This works out to be 181 correct judgements out of 256, a 70.70% success rate. Of the misses, there were 33 Type I errors (suspecting translator usage where there was none) and 42 Type II errors (not identifying translator usage even though one had been authorized).

### Table 4.17b Crosstabulation of rater judgement of translator use and actual usage overall

<table>
<thead>
<tr>
<th></th>
<th>CorrectAll</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 No</td>
<td>1 Yes</td>
<td>Total</td>
</tr>
<tr>
<td>RatersAll</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 No</td>
<td>135</td>
<td>42</td>
<td>177</td>
</tr>
<tr>
<td>Expected Count</td>
<td>116.2</td>
<td>60.8</td>
<td>177.0</td>
</tr>
<tr>
<td>% within RatersAll</td>
<td>76.3%</td>
<td>23.7%</td>
<td>100.0%</td>
</tr>
<tr>
<td>% within CorrectAll</td>
<td>80.4%</td>
<td>47.7%</td>
<td>69.1%</td>
</tr>
<tr>
<td>% of Total</td>
<td>52.7%</td>
<td>16.4%</td>
<td>69.1%</td>
</tr>
<tr>
<td>Std. Residual</td>
<td>1.7</td>
<td>-2.4</td>
<td></td>
</tr>
<tr>
<td>1 Yes</td>
<td>33</td>
<td>46</td>
<td>79</td>
</tr>
<tr>
<td>Expected Count</td>
<td>51.8</td>
<td>27.2</td>
<td>79.0</td>
</tr>
<tr>
<td>% within RatersAll</td>
<td>41.8%</td>
<td>58.2%</td>
<td>100.0%</td>
</tr>
<tr>
<td>% within CorrectAll</td>
<td>19.6%</td>
<td>52.3%</td>
<td>30.9%</td>
</tr>
<tr>
<td>% of Total</td>
<td>12.9%</td>
<td>18.0%</td>
<td>30.9%</td>
</tr>
<tr>
<td>Std. Residual</td>
<td>-2.6</td>
<td>3.6</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>168</td>
<td>88</td>
<td>256</td>
</tr>
<tr>
<td>Expected Count</td>
<td>168.0</td>
<td>88.0</td>
<td>256.0</td>
</tr>
<tr>
<td>% within RatersAll</td>
<td>65.6%</td>
<td>34.4%</td>
<td>100.0%</td>
</tr>
<tr>
<td>% within CorrectAll</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>% of Total</td>
<td>65.6%</td>
<td>34.4%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

### 4.7.1 Rater judgements for each composition

Additional tests were run to see whether rater judgments corresponded significantly to actual translator usage on each writing task. Since all compositions for the Pretest and Posttest were written without translator access authorized, it is not possible to run chi-square tests or Likelihood Ratios (since actual usage is a constant and would
involve division by zero). There were 17 Type I errors (false positives) out of 64 ratings for the Written Pretest — a 73.44% success rate — and 11 Type I errors out of 64 ratings for the Posttest (82.83% success).

For Task One, tests were run to determine the relationship between judged translator usage and actual usage (Table 4.18a). A highly significant association was found: \( \chi^2 (1) = 14.207 \) and Likelihood ratio = 20.098, \( p < 0.001 \).

<table>
<thead>
<tr>
<th>Chi-Square Tests</th>
<th>Value</th>
<th>df</th>
<th>Asymp. Sig. (2-sided)</th>
<th>Exact Sig. (2-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>14.207a</td>
<td>1</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Continuity Correctionb</td>
<td>12.125</td>
<td>1</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>20.098</td>
<td>1</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Fisher's Exact Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear-by-Linear Association</td>
<td>13.985c</td>
<td>1</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 6.56.
b. Computed only for a 2x2 table
c. The standardized statistic is 3.740.

Table 4.18a Relationship between rater judgment and actual translator use for Task One

Interestingly, as seen in Table 4.18b below, there were no Type I errors for Task One. Perhaps more importantly, there were 23 Type II errors (false negatives) out of 44 judgments. This result shows that there was only a 47.73% success rate in identifying compositions that had been written with the aid of a translator, with raters choosing correctly on compositions that had been written with OT just under half the time.

As above with the Written Pretest and Posttest, it is not possible to obtain chi-square or Likelihood Ratio statistics on the relationship between ratings and actual usage for just those compositions where a translator was used since the variable would be a constant. Standard residual for false negatives in crosstabulation (Table 4.18b) was only -
2.6, which does not rise to the level of concern, typically considered to be above [3]. While raters overall were able to correctly identify translator usage at a statistically-significant level, a closer look at the raw data shows that, although not statistically proven, raters appeared for Task One to correctly identify compositions where no translator had been used more often than compositions where one had been used.

<table>
<thead>
<tr>
<th>t1RatersAll</th>
<th>0 No</th>
<th>Count</th>
<th>20</th>
<th>23</th>
<th>Total</th>
<th>43</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expected Count</td>
<td>13.4</td>
<td>29.6</td>
<td>43.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>% within t1RatersAll</td>
<td>46.5%</td>
<td>53.5%</td>
<td>100.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>% within t1CorrectAll</td>
<td>100.0%</td>
<td>52.3%</td>
<td>67.2%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>% of Total</td>
<td>31.3%</td>
<td>35.9%</td>
<td>67.2%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Std. Residual</td>
<td>1.8</td>
<td>-1.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1 Yes</th>
<th>Count</th>
<th>0</th>
<th>21</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expected Count</td>
<td>6.6</td>
<td>14.4</td>
<td>21.0</td>
</tr>
<tr>
<td></td>
<td>% within t1RatersAll</td>
<td>.0%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td></td>
<td>% within t1CorrectAll</td>
<td>.0%</td>
<td>47.7%</td>
<td>32.8%</td>
</tr>
<tr>
<td></td>
<td>% of Total</td>
<td>.0%</td>
<td>32.8%</td>
<td>32.8%</td>
</tr>
<tr>
<td></td>
<td>Std. Residual</td>
<td>-2.6</td>
<td>1.7</td>
<td></td>
</tr>
</tbody>
</table>

Total Count | 20 | 44 | 64 |

|       | Expected Count | 20.0 | 44.0 | 64.0 |
|       | % within t1RatersAll | 31.3% | 68.8% | 100.0% |
|       | % within t1CorrectAll | 100.0% | 100.0% | 100.0% |
|       | % of Total | 31.3% | 68.8% | 100.0% |

Table 4.18b Crosstabulation of rater judgment and actual translator usage for Task One

For Task Two, there was again a significant association found between the judgment of raters and the actual use of online translation by participants, but unlike for the previously-mentioned compositions, these results were not very highly significant, but only significant at the $p < 0.05$ level, with $\chi^2 (1) = 6.398$ and Likelihood ratio $= 6.633$, $p = .015$ (Table 4.18c). These results indicate that raters were again able to identify translator usage correctly for Task Two with statistical significance.
Chi-Square Tests

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>df</th>
<th>Asymp. Sig. (2-sided)</th>
<th>Exact Sig. (2-sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Chi-Square</td>
<td>6.398a</td>
<td>1</td>
<td>.011</td>
<td>.015</td>
</tr>
<tr>
<td>Continuity Correctionb</td>
<td>5.106</td>
<td>1</td>
<td>.024</td>
<td></td>
</tr>
<tr>
<td>Likelihood Ratio</td>
<td>6.633</td>
<td>1</td>
<td>.010</td>
<td>.015</td>
</tr>
<tr>
<td>Fisher's Exact Test</td>
<td></td>
<td></td>
<td></td>
<td>.015</td>
</tr>
<tr>
<td>Linear-by-Linear Association</td>
<td>6.298c</td>
<td>1</td>
<td>.012</td>
<td>.015</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 9.69.
b. Computed only for a 2x2 table
c. The standardized statistic is 2.510.

Table 4.18c Relationship between rater judgment and actual translator use for Task Two

As seen in the crosstabulation below (Table 4.18d), raters correctly identified compositions 41 times out of 64, for only a 64.01% success rate. There were only 5 Type I errors but 18 Type II errors (not finding translator usage even though such use had been allowed). Based on these results, for Task Two we can say overall that raters were still able however to identify translator use to a statistically-significant degree.

Table 4.18d Crosstabulation of rater judgment and actual translator usage for Task Two
To summarize, instructors scoring compositions were able to judge if a translator was used to a statistically-significant level overall and for each composition. A closer look at the data shows that raters misjudged a fairly large number of cases where students had been allowed to use the translator, but were nearly perfect in correctly identifying compositions for which no translator had been used. In addition to making a determination as to whether or not online translation was used, raters were asked to explain the rationale behind their choices. A detailed discussion of their comments will be forthcoming in Chapter 5.

4.8 Participant self-reports

On the first day of the study, participants filled out a background questionnaire (Appendix B) asking a variety of questions about their experiences with French and computers. After completing the Posttest, students were also asked to comment on several aspects of their participation in the study as part of the Exit Questionnaire (Appendix N). While all participants were asked to complete this questionnaire, the focus of the questions was placed mostly on gauging translator use and impressions among participants in Groups B and C. There will be a more detailed look at specific participant comments in the Discussion chapter, but a cursory presentation of the main results of these two questionnaires is presented below.

4.8.1 Background Questionnaire

All participants were required to fill out a background questionnaire prior to taking the CLEP Pretest. As mentioned previously, the lengthy instrument (Appendix B)
asked participants to provide a wide range of information of various types (categorical, ordinal, and interval) for a number of items that should prove useful for further investigation. The current discussion will focus on the language background and previous online translator usage of participants.

The 32 students in the study all reported having English as their native language. All but one person was born in the US; one respondent reported being born in India but had lived in the US since the age of 5. Due to extended exposure to American English, it was not deemed necessary to exclude this participant from the study. Participants were also asked to provide information on other languages they spoke, had learned, or had studied. Unfortunately, although being instructed on the questionnaire to include French among the languages they spoke, most participants did not list French, their length of experience in it, and/or their self-reported fluency. Because the majority of respondents did not complete this section fully, it was decided not to use responses for further analysis but instead to rely CLEP and Written Pretest scores to compare ability levels in French. As mentioned in sections 4.1 and 4.3 above, all three groups were found to be statistically similar through both of these measures.

Participants were also asked about their computer usage in English, French, and other languages. Several questions focused the use of online translators. It was believed that if one group had significantly more experience with online translators than the others, this might be a confounding variable. In an attempt to control for this eventuality, participants were asked to rate on a 5-point Likert-type scale, ranging from 1 (Never) to 5 (Always), how often they had used a computer in the following five contexts:
• to translate text to or from English during the last semester they were enrolled
• to translate text to or from French during the last semester they were enrolled
• to translate text to or from French in class for the last semester they had French
• to translate text to or from French outside class the last semester they had French
• to translate to or from other languages besides English or French during their
last semester enrolled.

The results of each of these items will be discussed below to see if any difference existed among groups that might cause concern.

4.8.1.1 Online translation to or from English

As can be seen in Table 4.19a, the results of a one-way ANOVA did not find any significant difference across groups for online translation in English ($F(2, 29) = 0.04, p = 0.9612$), with means falling between 2.1818 for Group C and 2.3000 for Group A (Table 4.19b). A rating of 2 on the item corresponds to “Rarely,” indicating that at least according to self-reports, most participants across groups had not often availed themselves of online translation for English.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>0.08181818</td>
<td>0.04090909</td>
<td>0.04</td>
<td>0.9612</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>29.91818182</td>
<td>1.03166144</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>30.00000000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.19a One-way ANOVA results for prior translator use for English

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>SCORE</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>10</td>
<td>2.30000000</td>
<td>1.15950181</td>
<td></td>
</tr>
<tr>
<td>Group B</td>
<td>11</td>
<td>2.27272727</td>
<td>0.78624539</td>
<td></td>
</tr>
<tr>
<td>Group C</td>
<td>11</td>
<td>2.18181818</td>
<td>1.07871978</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.19b Means for prior translator use for English across groups

111
4.8.1.2 Online translation to or from French

A one-way ANOVA was run to see if there were differences among groups for their reported prior use of online translators for French. As with English, the difference across groups for OT in French did not reach the level of significance ($F (2, 29) = 0.06, p = 0.9459$), indicating that groups reported a similar amount of usage (Table 4.20a). The means were 2.4000 for Group A and 2.5455 for Groups B and C (Table 4.20b), showing that students reported using computers to translate text to or from French somewhere between “rarely” (2) and “sometimes” (3).

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>0.14545455</td>
<td>0.07272727</td>
<td>0.06</td>
<td>0.9459</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>37.85454545</td>
<td>1.30532915</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>38.00000000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.20a One-way ANOVA results for prior translator use for French

<table>
<thead>
<tr>
<th>N</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Group A</td>
<td>10</td>
</tr>
<tr>
<td>Group B</td>
<td>11</td>
</tr>
<tr>
<td>Group C</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 4.20b Means for prior translator use for French across groups

4.8.1.3 Online translation to or from French during classtime

After soliciting information concerning general usage of translators for French, the Background Questionnaire asked participants questions to determine where they used the translator: in or outside of class. According to Departmental policy at the University of Illinois at Urbana-Champaign at the time of the study, the use of online translators was not allowed for either purpose in the basic language sequence. The results of a one-way ANOVA (Table 4.21a) once again did not show a significant difference among the groups ($F (2, 29) = 0.13, p = 0.8799$). Participants reported using French in class slightly
less than overall (between $\bar{x} = 1.7000$ and $\bar{x} = 1.9091$), falling on the lower end on the Likert-type scale between “Never” and “Rarely” (Table 4.21b).

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>0.30056818</td>
<td>0.15028409</td>
<td>0.13</td>
<td>0.8799</td>
</tr>
<tr>
<td>Error</td>
<td>29</td>
<td>33.91818182</td>
<td>1.16959248</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>31</td>
<td>34.21875000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.21a One-way ANOVA results for prior translator use for French in class

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>SCORE Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>10</td>
<td>1.70000000</td>
<td>0.94868330</td>
</tr>
<tr>
<td>Group B</td>
<td>11</td>
<td>1.90909091</td>
<td>1.04446594</td>
</tr>
<tr>
<td>Group C</td>
<td>11</td>
<td>1.90909091</td>
<td>1.22102788</td>
</tr>
</tbody>
</table>

Table 4.21b Means for prior translator use for French during class

### 4.8.1.4 Online translation to or from French outside of class

As a final question concerning prior OT in French, respondents were asked how often, if at all, they had translated text outside of classtime either to or from French. Results of the one-way ANOVA run (Table 4.22a) showed no statistically-significant difference among the groups ($F(2, 29) = 0.58, p = 0.5652$). Although not reaching the level of significance, Group B’s mean ($\bar{x} = 2.6364$) was higher than those of Group A ($\bar{x} = 2.1250$) and C ($\bar{x} = 2.2727$), as seen in Table 4.22b. It should be mentioned that two participants from Group A did not reply to this question (even though all participants from the three groups had replied to the previous questions about translator use). Since the previous questions (answered by all participants) did not show any significant difference among groups for translator usage, including the second question (asking more broadly about translating text to and from French without specifying location or occasion), it is believed that this omission is not cause for concern. Based on the
responses of participants who did answer, overall translation use for French outside of class for these groups could be described as occurring “rarely.”

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2</td>
<td>1.36439394</td>
<td>0.68219697</td>
<td>0.58</td>
<td>0.5652</td>
</tr>
<tr>
<td>Error</td>
<td>27</td>
<td>31.60227273</td>
<td>1.17045455</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>29</td>
<td>32.96666667</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.22a One-way ANOVA results for prior translator use for French outside class

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Group A</td>
<td>8</td>
<td>2.12500000</td>
</tr>
<tr>
<td>Group B</td>
<td>11</td>
<td>2.63636364</td>
</tr>
<tr>
<td>Group C</td>
<td>11</td>
<td>2.27272727</td>
</tr>
</tbody>
</table>

Table 4.22b Means for prior translator use for French outside class

4.8.1.5 Prior online translation to or from other languages

For this final question about previous online translator use, only three participants reported using a translator for a language other than French or English. All three participants answered “1”, indicating that they had never translated text to or from another language. Because of the small number of respondents, and the fact that each who answered indicated he or she had not used a computer for this purpose, such use did not likely play a notable role.

Looking overall at reported translation prior to the study, the results obtained showed no statistical difference among the groups in terms of translator use to or from English, to or from French overall, as well as to or from French in or outside of class. For this reason, the groups can be assumed to be similar prior to the start of the study.
4.8.2 Translator use during the study

The 22 participants in Groups B and C, who were authorized to use a translator on Tasks One and Two, were asked in the Exit Questionnaire, “Did you actually use the translator for the writing tasks?” (Appendix N). All 22 participants reported in clear terms that they had used the translator, either by specifically replying “yes” to this question, or by describing the ways in which they used it. Based on participant self-reports, all participants who were instructed to use the online translator actually did so, as was intended.

4.9 Conclusion

The above sections have detailed a number of results, some anticipated and others unexpected, which were obtained from the present study. These findings have shed some light on the participants’ performance, raters’ perceptions, and the effect that online translator usage has on each of these. The results outlined above, in addition to a consideration of qualitative aspects of tasks completed by participants and comments given by raters, will serve as the basis for a detailed discussion in the next chapter.
Chapter 5
Discussion

5.0 Introduction

The study outlined in this document was designed to address a series of hypotheses related to the possible effects of online translation use on L2 writing in French. This chapter will discuss at length the results of the study; by looking one-by-one at each of the research questions in view of the findings detailed in the previous chapter, while offering the researcher’s interpretations of these results; then, by presenting and examining illustrative samples taken from participants’ writing as well as student and rater comments that may shed some further light on the results; and lastly, by summarizing the outcomes of the current research.

5.1 Effect of translator use on overall scores

Research Question 1 was divided into two parts, the first of which asked, “Does the use of online translators by L2 writers result in quantifiably different global scores on compositions written in the L2?” and the second part asking, “Of students who use an online translator, do those who have been trained in the use of online translators achieve higher scores on L2 compositions than those who have not received such training?” Given the results, it is logical to treat these two subquestions concurrently. The data collected offer a somewhat mixed picture that nonetheless appears to exclude statistically-significant negative effects of translator usage on overall composition scores as well as component subscores in this study, and may even speak to a possible positive
effect of OT on student writing in French, at least among those who had received training in translator usage.

At the start of the study, no significant difference was found among the means of global scores among groups in pretests both of reading comprehension and writing, indicating the groups were similar. On the first experimental task where one set of participants (Group A) was not allowed to use a translator while the other participants (Groups B and C) were, the translator groups scored slightly higher than the control group, approaching but not quite reaching the level of significance ($p = 0.0725$). On the second such task, however, the mean of global scores in Group C (whose participants had been trained in translator usage) was higher than those of the control group at a significant level ($p = 0.0032$). Group B, whose participants had received no such training, had a global overall mean falling between those of the other two groups, averaging over three points higher but not significantly different from Group A (-translator -training), and over a point and half lower than Group C (+translator +training), but not statistically different from this group, either. In the Posttest, where again no groups were allowed to use OT, no statistical difference was found among mean scores, with groups’ means within roughly a point of each other out of 30 possible points.

One possible interpretation for these results is that online translation did in fact affect students’ written production. While neither experimental group scored significantly higher overall than the control group on the first composition for which they were allowed to use OT, Group C may have been able to score statistically higher than the control group on the second such task thanks to prior experience with the online translator, through training and the previous task where they were allowed to use the
translator. It is possible this preparation allowed participants in Group C to become more skilled at effectively translating text and incorporating it into their compositions, thus allowing them to outscore their counterparts in the control group who didn’t have the benefit of an online translator when writing. This interpretation would be consistent with Somers (2003) and Williams (2006), who suggest that exposing students to online translation may be beneficial towards helping students understand how beneficial or detrimental using a translator in a certain way might be, as well as Bishop’s (2001) finding that students trained in dictionary use scored higher when using a dictionary to write in their L2 than those who received no such training.

Group B, which did not have training in online translation, did not benefit sufficiently from translator use to have meaningfully higher global scores than the control group on Task Two, but apparently enough to allow the group’s overall mean not to differ significantly from the group that had received translator training. It is also important to note that Group B’s mean scores did not suffer as compared to the control group – they did not score statistically worse than Group A and in fact in raw means outperformed them – indicating that using the translator did not harm their overall scores on the writing tasks. This interpretation would go against common wisdom in the field, including Luton (2003) who only discusses negative aspects of translator use for L2 writing, and McCarthy (2004) who advises discussing online translators with students but whose suggestions (such as moving composition assignments in-class to reduce translator usage) generally assume that using OT has detrimental effects on the final product. At least in terms of scores, this was not found to be the case in this study.
However, despite Group C’s statistically better performance on Task Two using an online translator after training as compared to Group A’s participants who did not use one, we cannot definitively accept and must therefore reject both parts of the first research question as stated. The translator groups combined did not significantly outperform the control group on either experimental task, and Group C’s mean was not statistically different from that of Group B. Further research would be needed to determine the importance of Group C’s significantly higher result on Task Two. Table 5.1 summarises these findings.

<table>
<thead>
<tr>
<th>Finding</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>No significant difference among groups</td>
</tr>
<tr>
<td>Task One</td>
<td>No significant difference among groups</td>
</tr>
<tr>
<td>Task Two</td>
<td>Group C (+translator +training) mean scores significantly higher than Group A (p &lt; 0.01)</td>
</tr>
<tr>
<td>Posttest</td>
<td>No significant difference among groups</td>
</tr>
<tr>
<td>Overall</td>
<td>On global scores, translator groups not proven overall to outperform control group; Group C not shown to outperform Group B</td>
</tr>
</tbody>
</table>

Table 5.1 Summary of findings for global scores

5.2 Effect of translator use on categories

While global scores speak to the overall performance of participants on the compositions, a second research question was put forward in an effort to get a more precise picture of what specific effects OT might have on student performance: “Does online translator usage affect perceived performance on any specific features of L2 learner writing, including comprehensibility, vocabulary, syntax, grammar, spelling, and content?” For all of these categories, scorers were asked to give each composition a rating between 1 and 5, with values being tailored to the specific feature but representing
a continuum where 1 represented the lowest score and 5 the highest (Appendix P). A look at each component follows.

5.2.1 Effect of translator use on Overall Comprehensibility

One feature that raters were asked to assess was the overall comprehensibility of each composition. In the scoring criteria given to raters, this component was described as the ability of the participant’s writing to “be understood by a native French speaker with little or no knowledge of English” (Appendix P). The scale for Overall Comprehensibility went from 1, indicating a composition could not be understood by a native Francophone, to 5, describing a paper that would be understood without any problems.

For the Written Pretest, Task One, and the Posttest, no difference was found among the three groups as regards comprehensibility. On Task Two, however, the mean score for Group C (+translator +training) was significantly higher than those of the control group. In other words, participants who had been trained in translator usage and used such a tool to write produced compositions that were judged to be more understandable than those of the control group. This finding is somewhat surprising. Based on a review of the literature, it was not expected that translator use might render student writing easier to understand. At best, it was expected that there would be no noticeable effect on comprehensibility, based on the study done by Leffa (1994) where high-school students performed equally well on a test measuring their comprehension of a science text translated by a profession human translator and the same text produced by a machine translator, or the results of Ablanedo et al. (2007) that found that Babel Fish translations obtained were all understandable, even if less accurate.
A possible explanation for better comprehensibility scores in Group C is the fact that participants in this group had an advantage over those in Group A, who only had their own knowledge of French to make themselves understood, while Group C could rely on their own abilities in French as well as text obtained with the aid of a translator to construct comprehensible output. This interpretation would be somewhat in line with Sabieh’s (2002) idea that machine translation might be a “powerful partner” that allows students to think about language use.

Group B’s mean comprehensibility score was higher, albeit to an insignificant level statistically, than the control group. One explanation as to why its participants’ performance was not significantly better is that unlike Group C, they had no prior training in identifying output that could possibly inhibit a native speaker’s understanding. To adapt an example from Hurman and Tall’s (2002) discussion on dictionaries, an online translator might give both a participant in Group B and one from Group C the word “gauche” as a translation for the verb “left,” but Group C’s student may have recalled from training that isolated words out of context are a potential pitfall for online translators, whereas a student from Group B may not have been looking out for such errors. While lexical in nature, this coupled with other errors could have an effect on whether a native speaker (as opposed to a language teacher versant in the students’ L1 who might catch the error) would understand.

While the significantly better performance of Group C on Overall Comprehensibility is noteworthy, since it only occurred on the second of the two experimental tasks, we cannot say with certainty that the treatment had a positive effect. It does appear clear however that participant use of an online translator did not have a
negative effect on how well raters judged that their compositions would be understood by native speakers, an unexpected finding that goes counter to common wisdom on the incomprehensibility of compositions that have been written with the use of a translator.

5.2.2 Effect of translator use on Content

In rating the content of a given composition, scorers were asked to consider “whether or not the ideas used sufficiently creative and appropriate to the subtopics outlined in the task” (Appendix P). The scale for these scores went from 1, not addressing any subtopics (the prompts for each paragraph on a given task) or the overall theme effectively and creatively, to 5, doing so for the theme and all three subtopics.

No significant difference was found between groups on either the Written Pretest or Posttest. On Task One, the difference in means between Groups A and B was significant, with the latter outperforming the former on Content. Group C also had higher scores than did the control group, but not enough to reach the level of statistical significance. On the second experimental task, however, the situation was reversed: it was Group C that had statistically better Content scores than the control group, while Group B had higher, but not statistically different, scores than Group A.

It was expected that students using the online translator might produce better content after consideration of the previous literature. Based on the model put forward by Luoma and Tarmannen (2003), observations by Uzawa (1996), and the CALL study done by Iwai (1999), it was thought that if an online translator could be used to handle form or surface features of their text, students would have more time to concentrate on developing content. The fact that a translator group scored higher on Content for both
Tasks One and Two would go towards supporting this reasoning. It is more difficult to explain why both groups did not outscore the control group on both experimental tasks, or why it wasn’t the same group that did so both times. The results for Content however do follow what we have seen thus far in that Group C significantly outperformed the control group on Task Two, suggesting perhaps that with training and practice with online translators, the participants in Group C were able to achieve higher scores than their counterparts without such preparation and access. Group B did score higher on average than the control group by over a half point both times, even if only Task One’s result reached the level of significance.

Since the findings are mixed, it is difficult to conclude definitively whether either treatment group (translator with no training, or translator with training) had an advantage over the control group. As before, however, neither group that used online translation performed worse than the control group, so we can state that the scores of students availing themselves of a translator were not negatively affected.

5.2.3 Effect of translator use on Spelling and Accents

A third area scorers were asked to focus on was Spelling and Accents. Those rating tasks were asked to give a score based on the following question: “Are the letters and diactrical marks in words written as would be expected in standard or colloquial written French?” (Appendix P). A clarification was given to raters (“Note: This does not include conjugation or other grammatical usage”) in attempt to pinpoint orthographical errors specifically as opposed to errors made on morphological or other grammatical
features. Raters could score the composition on a scale of 1, (almost) all words containing errors, and 5, (nearly) all words written correctly.

The Written Pretest and Posttest found no significant difference among the three groups. For Task One, the difference closely approached significance ($p = 0.0536$) between the control group and Groups B and C, with the mean for the latter two over 0.6 points higher than that of Group A. For the second experimental tasks, however, both experimental groups significantly outperformed Group A.

Although there was no clear guidance from the literature about what might be the possible effects of translator usage on orthographical features, the results for Task Two are not entirely unexpected. Uzawa (1996) discussed the focus that L2 learners often place on spelling and grammar when learning a language. Hurman and Tall (2002) found however that dictionaries were not used by their participants to check for spelling mistakes, and Iwai (1999) found in fact that CALL instruction negatively affected the importance L2 learners placed on spelling, but the latter study was for participants reviewing other students’ papers, which they had already read previously, as opposed to composing their own text.

Conversely, Myers (2000) found that two students using a pocket electronic dictionary for translation saw their L2 spelling improve, although without qualitative data or a control group, it is difficult to know whether or not this improvement could be attributed to consulting the electronic resource. While online translators have been reported in the literature to make various lexical and grammatical mistakes, no examination of errors of an orthographical or diacritical nature was encountered. Anecdotally, the researcher has not noticed such errors being produced by OT. Since L2
students on the other hand are known to make mistakes or errors with orthography, it would follow that text composed with the aid of a translator could reduce the presence of these.

Although the results of the study do not allow a definitive determination, the fact that the means for translator Groups B and C were significantly higher than Group A’s for Spelling and Accents for Task Two, and very closely approached significance for Task One, argues against a negative effect and potentially for a positive effect of OT use for these features.

5.2.4 Effect of translator use on Syntax

Raters were also asked to make judgments concerned a given composition’s syntax. For the purposes of the scoring for this study, syntax was narrowly defined as whether or not “the word order used [was] appropriate to standard or colloquial written French” (Appendix P). As detailed in the rubric, scores could range from 1, inappropriate word order throughout the composition, to 5, appropriate word order throughout.

No statistical difference was found among groups for the Written Pretest and Posttest. As with Spelling and Accents, although the means of Group B and C differed by over a half point as compared to that of the control group for Task One, this result did not quite reach the level of statistical significance \((p = 0.0638)\). For Task Two however, unlike all the previous categories, the difference between the means of the three groups on Syntax did not meet or even approach significance.

These results are somewhat, but not entirely, surprising. Watters and Patel (2000) found that online translators often made errors in word order. McCarthy (2004) also
demonstrated errors with syntax in his lessons on translator usage specifically between French and English. One might assume that translated text would cause syntactic problems that would lower the scores of the experimental groups. While the translator training session (Appendix K) in which Group C participated was designed to give an example of an error in syntax (inputting “I really don’t like your house” into Babel Fish had produced “*Je vraiment n’aime pas votre maison” in French), Group B did not attend this training. Without further guidance on this from the literature, one might suppose that students and online translators make errors in word order that are either similar in number and/or type; such a claim would merit further investigation. Based on the findings from the current study, syntax on student compositions was not significantly affected positively or negatively by participants’ usage of OT.

5.2.5 Effect of translator use on Remaining Grammar

In addition to the above features, raters were told to give a score for the broadly-defined component, “Remaining Grammar.” A note about what fell into this component told raters that it included “tense/mood, subject-verb agreement, number/gender agreements, articles, negation, etc.” (Appendix P). The motivation behind including this category was two-fold: first, to try to catch various linguistic features that might affect the overall impression the rater had of the composition and which did not fall under the other components, but which would have been too cumbersome to have raters evaluate individually; second, the composition rubrics used by the Department at the time of the study included a general “Grammar” category, and it was believed that in general instructional settings teachers often explicitly or implicitly base scores on a broad
evaluation of grammar. The rubric given to raters to evaluate Remaining Grammar was on a scale of 1, inappropriate grammar throughout, to 5, appropriate throughout.

Results show that the groups were similar in performance for Remaining Grammar in both the Written Pretest and Posttest. Significant differences were observed however for both experimental tasks. In Task One, Group B (which had access to a translator but no prior training in its use) had higher mean scores than Group A (having no translator access or training). While Group C (which had translator access and prior training) did achieve higher scores than the control group, this difference was not significant. In the second experimental task, however, both translator groups statistically outperformed Group A on Remaining Grammar to a significant level ($p < 0.01$).

The literature was divided on this issue. Some articles would have suggested that writing compositions with the aid of a translator would either have no effect or could negatively affect grammar in translator-aided texts as compared to those normally written by students without using OT. In a study of *Système D*, Scott (1996) stated that no students self-reported using the tool’s grammar index for assistance in their writing. Leffa (1994) specifically described machine translation as producing output prone to such errors. Heift (2003) reported an improvement of scores for those who used *German Tutor*, but unlike Free Translation, this tool consisted of exercises designed specifically to teach students proper grammar. One of the types of queries identified in Bland et al.’s (1990) early study on *Système D* saw students trying to obtain flexional forms for given words or explanations on grammatical concepts (such as negation), showing students may sometimes use electronic tools to aid with grammar. Williams (2006) found that
placement and agreement of adjectives specifically in French was generally correct for three online translators.

The results of the current study for grammar indicate that students translating text online for composition writing can outperform those who do not avail themselves of a translator. Students in Group B had higher means than the control group in both tasks for which they used a translator, while Group C’s near-significant results in Task One and significant results in Task Two appear to add support to the notion that OT not only didn’t harm, but likely improved instructor perception of grammar usage on student writing.

5.2.6 Effect of translator use on Vocabulary

Lastly, raters gave each composition a Vocabulary subscore. Those scoring the compositions were asked to judge whether word choice was “accurate and effective in standard or colloquial written French.” In an attempt to avoid any overlap between other categories, raters were further told that this category “does not include spelling, accents, or grammatical usage” (Appendix P). Raters were asked to assess vocabulary use on a scale of 1 to 5, from (almost) all of the words not being accurate and effective, to (nearly) all of them being so.

As with all the other component scores, group means did not significantly differ for the Written Pretest and Posttest, indicating the groups were similar in performance when not allowed to use the translator. For Vocabulary, there was also no statistical difference among groups for Tasks One or Two. It is perhaps noteworthy, however, that although not approaching statistical significance, this category was the only time that
Group A had a higher mean overall or on any category for either experimental task, even though it had higher (but not significantly so) means on several components for the Written Pretest and Posttest.

The fact that participants using an online translator did not do statistically worse on vocabulary than the control group may be somewhat surprising, but not entirely so. Yates (2006) deemed 15 out of 20 translation attempts by Babel Fish on law texts to be failed due to lexical and structural errors, but no breakdown of the types of errors was given to know how many were related to vocabulary. Cribb (2000) found after translating several passages that Babel Fish made lexical errors, while nonetheless producing understandable content. In particular, he found that the results were best for denotative meanings as compared to connotative meanings. If this is indeed the case, student performance with translators might depend on the topic or specific items they are translating. The prompts for the current study included some topics that could elicit either type of vocabulary, with each topic asking participants both to describe specific activities they did (e.g. on campus during the semester, or in preparation for their major or future career) as well as less concrete items (such as explaining why vacations can be interesting or boring, or convincing a French exchange student why they should choose (or not) the participants’ school). Because of this, the vocabulary for the tasks did not necessarily fall neatly into the categorization done by Melby (1997) of text types for machine translators, since some controlled domain-specific text may have been included in addition to dynamic general.

Burton (2003) mentioned a possible use of translators as being a dictionary for students to check or test out vocabulary words. Zamel (1982) found that students who
self-reported simplifying vocabulary received lower ratings on pen-and-paper compositions than those who did not. One possibility is that students using a translator did not simplify their vocabulary as much as those without such aid; lexical errors produced by the translator on some output may have been partly compensated for by vocabulary covering a broader semantic range, resulting in rater judgments for the composition as a whole that did not significantly differ from those for the control group. Further research would be needed to confirm this.

Based on the results for the Vocabulary component, no statistical difference was found among control and translator groups, suggesting that overall judgments on vocabulary usage were not significantly affected either way for texts written with the help of online translation.

5.2.7 Conclusion

Table 5.2 summarizes the results for component scores across groups and tasks. For the two experimental tasks, statistically-significant differences were found in eight total relationships between an experimental group and the control group, but never between the two experimental groups. Based on these results, we can conclude that the control group did not significantly outperform students using translators on any component. In fact, one or both translator groups had higher mean scores than the control group on two out of six measures for first experimental task, and four out of six measures on Task Two.

For Remaining Grammar, Group B (+translator -training) scored significantly higher than Group A (-translator -training) on both experimental tasks. In terms of
Content, the control group was outscored by one translator group (Group B) on the first task and the other (Group C, +translator +training) on the second. On Task Two, Groups B and C both had higher mean scores than the control group for Spelling and Accents, while Group C alone outscored the control group on Overall Comprehensibility. There was no significant difference between the control and translator groups on Syntax and Vocabulary, although for the former both Groups B and C approached significance on Task One.

<table>
<thead>
<tr>
<th>Overall Comprehensibility</th>
<th>Finding</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group C (+translator +training) had a higher mean than Group A (-translator -training) on Task Two</td>
<td>Group B’s (+translator -training) mean higher than A’s on Task Two, but not significantly so</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Content</th>
<th>Finding</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group B had a higher mean than Group A on Task One; Group C outperformed Group A on Task Two</td>
<td>Group C’s mean was higher, but not significantly, than control on Task One, as was Group B’s on Task Two</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spelling and Accents</th>
<th>Finding</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both Groups B and C had significantly higher means than Group A on Task Two</td>
<td>Means for Groups B and C higher on Task One than Group A, approaching significance ($p = 0.0536$)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Finding</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group B had a higher mean than Group A on Task Two</td>
<td>Group B and C had higher means on Task One than Group A, approaching significance ($p = 0.0638$)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Remaining Grammar</th>
<th>Finding</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group B outperformed Group A on both Tasks One and Two. Group C did so on Task Two.</td>
<td>Group C scored higher than Group A on Task One, but not significantly so.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vocabulary</th>
<th>Finding</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>None of the correlations approached statistical difference</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall</th>
<th>Finding</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Translation use had an effect on some components. On Task One, at least one translator group out-performed control group on two of six measures. On Task Two, this was so for four of six components</td>
<td>For all categories but Syntax and Vocabulary, at least one translator group outscored control group significantly on at least one experimental task</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.2 Summary of findings for scores for six categories

Although no clear pattern emerges from these results, five findings point towards a possible positive effect of online translation on L2 writing for component scores:
• No statistically-significant differences were found in either the Written Pretest or Posttest, while a number of differences were found on the experimental tasks, in which some students were allowed to use a translator and others were not.

• On experimental tasks, eight out of 24 possible correlations showed a translator group with statistically higher means than the control group, with another four relationships narrowly missing statistical difference, meaning that half of the correlations either reached or approached significance.

• There were no relationships that showed the control group with significantly higher scores than a translator group. In the only case the control group had a higher mean than a translator group on a component score, this difference did not approach significance.

• Every time one of the online translator groups outperformed the control group significantly, the other one either did so as well, or had a higher but not statistically different score.

• The only category for which a translator group never performed better at a significant or near-significant level than the control group in at least one experimental task was Vocabulary (with the results for Syntax approaching significance on Task One).

A possible explanation for the differing results between Task One (which saw relatively few significant differences among groups) and Task Two (which saw translator groups, and in particular Group C, often outperforming the control group) may be the same as that offered for the global scores. With more practice, translator groups may have become more proficient at effective translator use. The fact that Group C received training in online translation may explain why their results were more often statistically
higher than Group B’s in relation to the control group. Further research would be needed to confirm these findings, particularly in cases where the difference narrowly missed statistical significance.

To return to the research question for this portion of the discussion, we can say that results indicate that in some cases, online translator usage does indeed affect perceived performance on features such as comprehensibility, grammar, spelling, and content; although no clear pattern emerged, in those cases where there was a significant effect, it was always positive for those who used online translation. OT usage did not however statistically significantly affect scores on vocabulary and syntax.

5.3 Rater judgments on translation use

The third research question called for an investigation into raters’ judgments about perceived translator use for each paper they scored. When considering all compositions (Written Pretest, Tasks One and Two, and Posttest) together, raters were in fact able to identify translator usage to a very highly significant level. The same was true when considering the Pretest, Posttest, and Task One separately. In each of these cases, there was a significant correlation between suspected translator usage and actual usage, suggesting that raters were able to identify correctly which participants had used a translator and which had not. For Task One, however, raters correctly judged compositions written without a translator as being such perfectly (100% of the time), but made a total of 23 Type II errors for translation-aided compositions, only correctly judging these 48% of the time. Although it must be stressed that this result was noticeable, the difference overall between judgments and actual translator usage was not
enough to reach a level of statistical significance: raters were right often enough on the compositions written without translator aid that ratings correlated with actual usage globally despite the presence of a number of Type II errors for Groups B and C.

On the second experimental task there was again a significant correlation between rater judgments and actual translator use, but less strongly so than for the other compositions. Raters correctly classified tasks based on perceived translator use just over 64% of the time in all, with five Type I errors (representing about a three-in-four success rate for compositions not written with the aid of a translator) and 19 Type II errors (showing that raters correctly detected translator usage a little over half of the time). It is interesting to note that the total number of errors for Tasks One and Two was nearly the same, with raters only making one more error overall on Task Two (24) than Task One (23), the correlation was less significant for Task Two in part because the errors were of both types, having a negative impact on the correlation. On the other hand, all errors for Task One, being in one direction, had less of a detrimental effect on the overall correlation model.

After a discussion of results for global and component scores, the fact that there were a number of Type II errors on identifying OT-aided writing might be expected since Group C outperformed the control group on the second experimental task. However, a further look into the rating data reveals an unexpected result. On Task Two, translator use went undetected more often among Group B (15 out of 22 times, or just over 68% of the time) than among Group C (only 4 out of 22 times, or a little over 18% of the time). This observation suggests that training in translator usage did not necessarily mean that participants were more skilled in avoiding detection when using the translator than their
counterparts who had not attended the lesson on OT use. The motivation behind the training was to show participants what an online translator was, to make sure they knew how it worked, and to alert participants to strengths and weaknesses of translators; the goal was not to assist them in using translation without being noticed. Even though raters more often than not judged correctly that these compositions were written with the aid of OT, they still gave Group C significantly higher global scores than the control group.

As discussed above, an analysis of the component scores showed that on four out of six measures, Group C outperformed the control group on Task Two. It is surprising that raters would correctly identify translator use and yet assign higher scores, given the overall negative impression that online translation has in the field and in the literature, including Stapleton (2005) and Williams (2006) (the latter referring to most online translators as giving “inaccurate, unacceptable translation”) and the fact that some teachers believe that they can easily spot translator usage when they see it (such as the “red flags” mentioned in Luton 2003). One possible explanation for this unexpected result is that raters in the study were instructed to judge each individual component based on its own merit regardless of whether or not translator use was suspected, whereas teachers in an instructional setting might assign a failing or reduced grade as a sanction for perceived translator use.

In all, there were 181 correct judgments out of a possible 256, resulting in a 70.70% success rate across tasks. While the overall model showed a high correlation between judgments of translator use and actual usage, raters were far from infallible in their choices. It should be noted that there were many more compositions for which translator access was not allowed than ones where students were permitted to use one: all
three groups for both the Written Pretest and Posttest were not allowed translator access, and Group A additionally did not have such access on either experimental task. For this reason, 208 out of the 256 ratings were for compositions where translator use was not allowed. This fact also helps explain in part why the overall correlation can be very highly significant across the four tasks while still having a relatively large number of Type II errors for those compositions that were written with the aid of a translator.

To summarize, foreign language instructors rating compositions were correctly able to determine whether or not an online translator was used to a statistically-significant extent overall and for all four compositions. Despite this finding, raters committed a number of Type II errors, showing that in some cases participants were able to write a composition with the help of a translator without the result being judged as such. In addition to providing judgments on translator use, raters were asked to provide comments justifying their decisions. A look at samples of student writing in relation to both rater comments and participant self-reports will be presented after a discussion of the latter.

5.4 Participant self-reports

The following section will deal with the last research question, “Do student self-reports over the positive or negative effects of online translator usage correspond to observed effects of online translator usage on L2 writing?” A presentation of participant comments for the cultural lesson (which served as a control for the translator training session) and translator training will lead to a discussion of the relationship between participants’ impressions and the observed positive or negative effects of online translator usage on their writing. Although some tabulations will be presented, the focus will be
more on a qualitative discussion of results for these data since they involve student opinion with comments that cannot always be easily categorized or quantified.

5.4.1 Self-reports on cultural lesson

Participants in Group A (-translator, -training) and Group B (+translator, -training) both attended a cultural lesson concerning the field of translation (Appendix M). The lesson was given by the participants’ regular instructor during regular class time, with those students from Group C (+translator, +training) being excused from class and expected to attend the translator training session later in the day. Participants attending the cultural lesson were presented with general information about the field of translation, such as the differences between interpretation and translation, types of translation, and career opportunities available for people interested in pursuing translation professionally. The lesson was taught entirely in English, and participants were asked to answer comprehension questions after each, with a whole-class discussion of answers. The motivation behind this lesson was to serve as a control to the translator training session that Group C (+translator +training) attended instead later the same day. In this way, the total amount of time spent in the study was equal for both those with and without translator training, and all participants attended an instructor-led lesson in which they completed exercises on their own and had an instructor later go over correct answers with them.

In Table 5.3, participant responses are tabulated for responses to Question 3 on the Exit Questionnaire concerning reactions to the cultural lesson. As might be expected, most students did not feel the lesson helped them with their subsequent writing tasks or
with their French. Perhaps for this reason, and due to the framing of the question, most comments were negative. One participant astutely remarked that the cultural lesson “did not improve my writing, but that wasn’t the focus.”

<table>
<thead>
<tr>
<th></th>
<th># of respondents Group A</th>
<th># of respondents Group B</th>
<th>Total count</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Informative</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Interesting</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Straight-forward</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Negative</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Didn’t improve writing / French</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Would have preferred lesson in French</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Better to review / prepare for writing instead</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Lesson only about what translators do</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>No info on how to translate / use online translators</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Didn’t retain / think about it afterwards</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Dictionary would have been more helpful</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>No help remembering correct formulas for tenses</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Don’t think much would have helped</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Information was redundant</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only taught how (online?) translators were used</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td>14</td>
<td>12</td>
<td>26</td>
</tr>
</tbody>
</table>

Table 5.3 Tabulation of participant comments on cultural lesson

It is obvious from the results that participants in both groups would have wanted language- or writing-focused information, with many students saying they wished the lesson had been in French or had allowed for review. It had been felt however that providing a lesson with writing practice in French might represent an unwanted confounding variable instead of acting as a control to the online translator training session participants in Group C attended.

Two students (one each from Groups A and B) commented that they wished the lesson had taught them to use online translators. Since they had been asked in the preceding question (discussed below) if they would have liked training, it is possible this...
was mentioned as a continuation of the comments to the preceding question, or to stress
the fact that they would have preferred to receive such training.

Several students mentioned that they felt they did not remember or think about the
cultural lesson afterwards, while one participant said that access to a dictionary would
have been more helpful for writing than the cultural lesson (which, again, would have
introduced an unwanted variable into the study). One participant complained that the
lesson gave no help for remembering “formulas for writing in various tenses.” This is
perhaps related to various strategies used by instructors to help students decide which
verb tense to use (si + imparfait … conditionnel, if-clauses, for example). It is unclear
why the participant gave this reply to a question about the cultural lesson; it was perhaps
something that s/he was generally concerned with for class, was struggling with in
French, or which came up during the participant’s writing process in the study.

One last comment was puzzling. A participant in Group B wrote that the lesson
“just taught me about translators and how they are used.” The cultural lesson did not
teach students how online translators are used, and attendance reports show the student
was not in the translator training session. It is possible the student meant that the lesson
explains how professional human translators are used in various fields (financial
translation, scientific translation, etc.), or that the participant had forgotten what the
cultural lesson was about, since others reported not retaining information from the
lesson.

Overall, students did not have many positive comments about the cultural lesson,
with some correctly suspecting that the lesson was not meant to help them improve their
writing in French (instead, serving as a control to the training other participants were
receiving) and others offering ideas of what they would have rather done or learned about during the lesson.

5.4.2 Self-reports on translator training

The eleven participants in Group C were asked the following question on the Exit Questionnaire: “Do you feel that this training helped you in using a translator? Please explain why or why not.” Table 5.4a presents a tabular summary of participants’ comments. As above, participants sometimes had more than one comment. Seven respondents gave at least one comment that the session (Appendices K and L) was at least partly helpful for them, while four indicated that the training session did not help them.

Of those who said the training was helpful, two mentioned that they believed training showed them why, or in what circumstances, translators were not good to use. Three participants indicated that the session taught them ways to use a translator effectively, with one comment saying the training helped to know what types of phrases are better to translate, another saying that they learned the translator helped translate vocabulary words, and a final commenter simply stating the training helped with using the translator “successfully,” without further details. Two comments additionally did not specify in what way the training was helpful, only writing that the session was “helpful” or “somewhat” (helpful) and not indicating whether it showed them how to use a translator, showed them strengths and pitfalls, or some other precise information that they felt was useful.
All four participants who said the training session was not helpful stated that they either already knew how to use an online translator, and/or that translators were easy to use and thus they did not need training in OT. One additional participant commented that translators “were confusing with tenses.” It was unclear if the respondent meant that this observation was made thanks to the training session (and whether s/he considered it to be a positive or negative outcome of the lesson), or if it was a more general comment about translators gleaned through their subsequent use and not specifically related to the training itself.

<table>
<thead>
<tr>
<th>Positive comments</th>
<th># of comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Realize flaws / inefficiency of translators</td>
<td>2</td>
</tr>
<tr>
<td>(Somewhat) helpful</td>
<td>2</td>
</tr>
<tr>
<td>Understand using translator successfully</td>
<td>1</td>
</tr>
<tr>
<td>Know which types of phrases translate better</td>
<td>1</td>
</tr>
<tr>
<td>Translator helped with vocabulary words</td>
<td>1</td>
</tr>
<tr>
<td>Didn’t know about online translators before</td>
<td>1</td>
</tr>
<tr>
<td>Found out which website was more useful</td>
<td>1</td>
</tr>
<tr>
<td>Negative comments</td>
<td></td>
</tr>
<tr>
<td>Translator usage easy / self-explanatory</td>
<td>2</td>
</tr>
<tr>
<td>Already knew their limitations</td>
<td>2</td>
</tr>
<tr>
<td>Already knew how to use them</td>
<td>1</td>
</tr>
<tr>
<td>Other comments</td>
<td></td>
</tr>
<tr>
<td>Translators are confusing with tenses</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 5.4a. Comments by Group C concerning translator training session

Based on self-reports, most participants felt the training session was helpful. This outcome had been expected, since the lesson presented information that is not commonly known, even among some L2 instructors, about potential strengths and weaknesses of OT output. Respondents indicated that they either learned when to avoid using a translator, which goes along with the suggestions of McCarthy (2004) for desired outcomes for such a lesson, or cases when the translator might yield useful results, in part informed by
Williams (2006). Several participants however mentioned the session not being helpful due to previous knowledge or experience with translators.

Groups A and B, which did not undergo translator training, were asked: “Do you feel you would have benefited from such training? Please explain why or why not.” These comments are presented in Table 5.4b. For Group A, which did not use a translator during the study, most comments indicated that they would have found translator training helpful. Five respondents specifically mentioned issues related to vocabulary, indicating that they felt they had limited vocabulary, or that there had been times during the study where they wish they had had some help with either looking up or remembering words. One respondent said s/he felt the writing level in French would have been more “complex” if s/he had used a translator, which could be related to a variety of features (vocabulary, content, etc.). Another respondent reported having troubles with self-expression in French, feeling that translator training might have aided in this regard. One interesting comment was from a participant in Group A who answered that using an online translator “couldn’t possibly harm me,” indicating either a lack of knowledge on possible problems translators can introduce into writing or a lack of confidence in the participants’ writing ability.

Of those who did not think translator training would be useful, two participants said this was so since they did not use the translator during the study; the question might have been worded more clearly to indicate this hypothetical situation would include both training and subsequent access to a translator. A final comment stated that access to a dictionary would have been useful instead. Overall, participants in Group A appeared interested in a translation training session, possibly since they were not allowed access to
the translator during the study and were either hopeful that an online translator or similar tool could help them with vocabulary and other problems they had encountered on their own during the study, or simply curious about online translators since they knew that this was the focus of the study and that other participants had been allowed to use or even attend a training session concerning OT use.

<table>
<thead>
<tr>
<th>Positive comments</th>
<th># of respondents Group A</th>
<th># of respondents Group B</th>
<th>Total count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low vocabulary level</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Forget verb conjugations</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Frustration / help needed with finding words</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Writing at too low a level</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Troubles expressing self</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Would make writing easier</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>It couldn’t possibly harm</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Would help, but not much</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Negative comments</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did not need a translator because didn’t use one</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>A dictionary of sorts would have helped</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Already know how to use a translator</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Translators are easy to use</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Already know translators aren’t accurate</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Other comments</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not sure, think can already use them well</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Not sure, but assume it would be beneficial</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Cultural lesson had interesting background information but not related to online translator</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Cultural lesson had nothing useful for translating</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Maybe, sometimes had to reword input to get correct verb</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>13</td>
<td>12</td>
<td>25</td>
</tr>
</tbody>
</table>

Table 5.4b Comments by Groups A and B about desire for translator training

Among respondents for Group B, which did have translator access for the experimental tasks, most participants either did not express interest in a translator training session, or stated they weren’t sure if such training would have been beneficial. Only one participant said that s/he thought training would be helpful, adding “but not much” without further explanation. This outcome is unexpected; it had been thought that those who used a translator but hadn’t attended the translator training session would have
desired such training. Instead, four participants reported not needing training due to already having enough knowledge on online translators, stating that they already knew how to use one or that translators were easy to use. Another person reported knowing that online translators weren’t accurate, assuming perhaps that the online translation lesson was going to focus solely on negative aspects of translation use, or not knowing that there are some instances where online translators produce correct results.

Four comments expressed uncertainty or ambivalence about the prospect of online translation training. Two participants said they believed they could use online translators well but were unsure if the training would have been helpful. Another participant reported being “not sure” about training but “assumed” it would be beneficial. Another commenter wrote that training “maybe” would be useful, citing the fact that the participant sometimes had to reword output from the translator during the study “to get the correct verb.” Lastly, two respondents appeared to misunderstand the question, commenting about the cultural lesson they had attended and not the possibility of instead attending a training session for translators, with a final participant not answering this question for unknown reasons.

In summary, most participants in Group A said they felt they would have benefitted from translator training had they been offered it, while Group B’s comments indicated they either believed training would not have been helpful for them or were not certain about its possible usefulness. The fact that Group B did outperform the control group on several subscores and never significantly underperformed as compared to Group A might explain why those who had access to the translator but no training did not feel as though they had missed out. Participants’ responses from Group C, which was the
only group to have received translator training, showed most students reported the training to be beneficial to them. This group as a whole had higher means globally on Task Two and on a number of subscores for this and the first experimental task, pointing to a possible positive effect of the training on these participants that would confirm most group members’ impressions.

5.4.3 Self-reports on translation use

Since it was not logistically possible for this study to record keystrokes or videotape participants as they used the translator, those who were allowed to access online translators were asked to describe their usage. There are several related questions on the Exit Questionnaire (Appendix N) that will be considered to help gauge the participants’ perception of their translator use during the study. Based on participants’ reports, all students in Groups B and C did use an online translator as instructed, but for varying uses and to differing extents, ranging from searching isolated words to translating entire paragraphs. Table 5.5 shows for what purposes respondents said they used the translator.

<table>
<thead>
<tr>
<th>Ways of using translator</th>
<th># of participants, Group B</th>
<th># of participants, Group C</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Looking up words</td>
<td>5</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>Looking up phrases</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Help with spelling / accents</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Translate entire sentences / paragraphs</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Look up gender / articles</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Help with verb conjugation / tenses</td>
<td>2</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Double-check work after writing</td>
<td>2</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Help with syntax / sentence structure</td>
<td>2</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Help write more quickly / easily</td>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Help with adjectives / agreement</td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Help remembering a word</td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Help convey meaning more directly</td>
<td>1</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5.5 Tabulation of self-reports for how translator was used
Some participants named one area, while others mentioned several. Although it is difficult to categorize the comments neatly, most participants mentioned using the translator to look up words or phrases. Several participants specifically mentioned using the translator to check their work (such as the meaning of a word or a grammar point) after they had initially attempted to write on their own.

By far, students most often reported using online translation for lexical and orthographical purposes. The focus for many students, at least according to self-reports, was to find or check individual lexical items. In most cases, it wasn’t clear from student comments whether they were using the online translator to verify the meaning or spelling of words they already knew, or instead to discover new words to use that would convey their meaning. Four comments centered on using the translator for help with spelling or accents, which is an interesting approach since the online translator is designed to convert from one language to another and not to assist with usage within a language.

Five participants reported going beyond searching for individual lexical items by seeking assistance with translating phrases. It was not specified if this meant looking up idiomatic expressions or simply strings of text longer than a word. Four additional participants admitted to translating entire sentences or paragraphs. One participant commented, “I probably translated 1/3 of my composition”; another reported, “I wrote out the story in English and had it converted to French.” This sort of usage was expected among some participants, as Zamel (1982) had reported a student writing text out in English and then translating it herself long before the advent of online translators to assist with this approach to writing. It is interesting that with 22 students in the groups allowed
translation use, more students did not adopt this strategy (or at least, did not report having done so).

While two of the students reported using Free Translation to translate whole sentences, it is interesting to note that another three students specifically mentioned “never” using it to translate whole sentences. The question did not attempt to elicit for what uses participants had decided not to consult the online translator, but these students (who were both in Group B, the group without training) nonetheless felt it important to note not translating sentence-long text. One student reported the reasoning behind avoiding OT for complete sentences: “I never typed in whole sentences to translate because I thought the grammar would come out wrong.”

A number of students mentioned using the translator for grammar-related purposes, such as getting assistance with conjugations, help with sentence structure, finding out the gender of a noun or proper adjective agreement with nouns, or checking grammar after writing. Two students also indicated using online translation to help with syntax or sentence structure, while another two indicated they used the translator to check their work after trying themselves in French. It is interesting to note that most participants who reported using the translator to produce or check grammar or syntax were in Group B. This may possibly be due to the fact that translator training received by those in Group C showed some examples of translators producing correct grammar or syntax while also presenting other cases where OT provided ungrammatical output or inappropriate word order. Two participants in Group C however said they used the translator to make it easier or quicker to write, without specifying however how they accomplished this.
Overall, most participants indicated they had used a translator for help with finding or remembering individual words or phrases, while some used it to assist with spelling, grammar, syntax, or checking their work. Four students admitted to writing entire sentences or paragraphs in English first and having the website provide them with a translation into French, while other participants specifically mentioned not having done so. These self-reports suggest that participants adopted a variety of strategies for using the online translator.

5.4.4 Self-reports on effects of online translator use

In addition to reporting how they used OT during the study, participants commented on their perception of the effects of using this tool. For question 3b) of the Exit Questionnaire (Appendix N), participants were asked if the translator helped their writing; if so, with what in particular it had helped them; if not, for what reason they believed it had not been helpful.

The researcher categorized respondent comments on perceptions about OT as positive, negative, or other. Overall, participants reported positive aspects of online translator usage more often than negative ones. Most students mentioned at least one positive item about their experience with using the translator for writing (27 total comments), while a number of unfavorable impressions were also given (14 total comments) and one comment was provided that could not be classified as either. The self-reports for Group B, who had no prior training, made more comments (24, of which 16 positive comments versus 8 negative) than the group that had received training, with Group C making 18 comments overall, of which 11 were deemed positive, six negative,
and one unclassified. One student in Group C also did not answer this question for unknown reasons.

A tally was done of reports made by participants concerning specific aspects they found to be positive or negative. Table 5.6 presents the information gleaned from their comments. Among the positive comments include the ability to look up or check individual words, spelling or accents; use the translator like a dictionary; and check correct word usage. It is difficult to neatly divide these types of comments into categories: “checking” a word for one person may mean assistance in making the correct lexical choice, while for another it may mean verifying the orthographical form, for example. For the latter, results from subscores would tend to support the impression that the translator did have a positive effect, since both translator groups outperformed the control group for Task Two and approached doing so on the first task. No students mentioned the translator as having a negative impact for spelling or accents, and as mentioned early, this topic was also not investigated in the literature.

For students who may have meant that their vocabulary usage was helped by the online translator — including those mentioned above, as well as three additional participants who reported that the translator helped them recall words and one stating the translator helped because of his/her limited vocabulary — a positive effect of online translation effect is not borne out in the group results from the writing tasks. Neither experimental task saw statistically higher scores for the translator groups in Vocabulary scores. Several participants mentioned negative impressions concerning translators for vocabulary, including translations being too literal or giving an incorrect meaning. It is possible the online translator gave mixed results, sometimes providing the correct word
and helping the participant, and other times not, resulting in either the participant rejecting the translator’s suggestion or including an incorrect word or expression in their writing.

<table>
<thead>
<tr>
<th>Positive</th>
<th># of respondents Group B</th>
<th># of respondents Group C</th>
<th>Total count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Look up / check individual words</td>
<td>3</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Look up / check spelling or accents</td>
<td>5</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Remembering words</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Check correct word use</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Conjugation / tenses</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Use like a dictionary</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Check phrases</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Helpful for people with limited vocab</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Include things participant would have left out</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Write French at a higher level</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Express meaning more directly</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Help finish writing task</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Subtotal</strong></td>
<td>16</td>
<td>11</td>
<td>27</td>
</tr>
</tbody>
</table>

| Negative                                      |                           |                           |             |
| Didn’t learn / retain phrases                 | 1                         | 2                         | 3           |
| Translator not good with grammar / conjugations| 1                         | 2                         | 3           |
| Translator not good with phrases             | 2                         | 0                         | 2           |
| Too easy / don’t have to think               | 1                         | 1                         | 2           |
| Translator too literal                       | 1                         | 0                         | 1           |
| Translator not accurate                      | 1                         | 0                         | 1           |
| Translator gave incorrect meaning            | 0                         | 1                         | 1           |
| Didn’t improve writing ability               | 1                         | 0                         | 1           |
| **Subtotal**                                  | 8                         | 6                         | 14          |

| Other                                         |                           |                           |             |
| Not many unknown words in translation         | 0                         | 1                         | 1           |
| **Total**                                     | 24                        | 18                        | 42          |

Table 5.6 Student self-reports on positive and negative effects of translator use

Another possibility, especially given the relatively high number of comments related specifically to words, is that some participants are still under the assumption that good performance in a foreign language depends mostly on finding the right words, as under Bland et al.’s (1990) Naive Lexical Hypothesis, instead of considering linguistic ability and production more broadly. It may also simply be the case that some participants felt reassured in their vocabulary choice because the computer confirmed,
whether correctly or not, the word they had already intended to use, while others noticed that some translated vocabulary was incorrect.

Two participants in Group B mentioned receiving OT help with verb conjugation or tenses, while three participants (two in Group B and one in Group C) mentioned the translator not being good with grammar. In spite of this, and although not mentioned by many participants, grammar was one area where translator groups performed particularly well as compared to the control group. Group B had higher mean scores on Remaining Grammar on both experimental tasks, while Group C did on the second. One possibility is that the translator helped some students more than others on this component, even though the overall results indicate a positive effect. Another possibility is that participants did not much focus on whether the grammar output was correct or incorrect, either assuming that the translator had gotten the translation right or concentrating more on other issues such as comprehensibility and content.

It had been predicted, informed in part by the writing model suggested by Luoma and Tarnanen (2003) and analysis by Zamel (1982) of writers from her study, that the translator might free students to concentrate less on form and more on content. Three comments focused specifically on issues that may be related to content or overall comprehensibility. One participant reported that thanks to using a translator, s/he was able to include things that would otherwise have been left out due to an inability to express it in French. Another comment mentioned the participant being able to write French at a higher level, while a third report credited the translator with helping to expressing meaning more clearly. On the other hand, however, one participant reported the translator being “too easy” to use, while another complained “I did not have to use
any real brain power because the translator did it for me.” One participant mentioned not coming across many unknown words in the translation s/he did. It was not clear whether this was meant as a positive or negative, but it is possible that this person meant the translator was not particularly useful or challenging.

A possible interpretation of comments such as these would be that the translator does in fact allow students spend more time or attention on content, but that students will not necessarily take advantage of this benefit. Some participants may have attempted to use OT to explore the language and improve their expression in the L2, while others were using the translator simply to complete the task sooner. Although difficult to interpret, this may have been what one respondent meant when reporting “The only thing that it helped me with was finishing the task.”

None of the participant comments directly addressed word order. Since there were no significant differences found in the syntax subscores on any of the writing tasks, it is possible that the online translator had little or no effect on this and the issue thus did not come to students’ attention or stand out as something noteworthy on which to comment.

In response to Research Question 4, there does not appear to be a direct correspondence overall between the effects of online translation use and participant reports of the effects. Most participants who used the translator had at least some positive comments concerning its effects. On Task Two, one of the translator groups had a higher mean than the control group; on Task One, both translator groups outperformed the control group on some measures. The relatively high number of negative self-reports by participants in the translator groups, however, cannot be explained by the results on the compositions. The control group never scored significantly higher than the translator
groups, with the latter actually performing better in a number of categories – including grammar, a component for which translator group participants had several negative comments despite the groups’ comparatively good mean scores. It is possible that students were not able to judge well on their own whether the output of the translator was correct or beneficial to them. That translator use did not significantly help improve scores for all features or overall on the first task participants completed may also explain in part why there were a number of both positive and negative views of some towards the effects of online translation. A closer look below at some samples of participant writing, coupled with their self-reports and the impressions of raters, should provide more insight.

5.5 Samples of participant writing and self-reports, rater judgments and comments

To finish the discussion of the data obtained in this study, examples of participant writing will be presented in light of judgments and comments made by raters, as well as participants’ own reports on their translator usage, in an attempt to get a more qualitative glimpse at the effects of OT use. In addition to a binary decision concerning translator usage, raters were asked to provide a justification describing the rationale behind their decision. They were additionally told they could, but were not required to, make any additional marks and comments on the composition as they read and scored it either to help them determine what scores to award on the components or to point out words or passages, whether erroneous or correct, that they found noteworthy.

A detailed analysis of rater comments as compared to participant writing and self-reports is outside the purview of the current discussion. Instead, a selection of compositions, accompanied by a sampling of rater comments, participant self-reports,
and the researcher’s observations, will be offered, looking both at samples of what were judged to be well-written and error-ridden writing. Because these cases were judged to be more interesting, the discussion will focus principally, but not exclusively, on cases where one or both raters incorrectly judged whether or not a translator had been used. While not exhaustive, this discussion should serve as an initial exploration at the complex relationship between L2 written production, online translator use, and participant and rater perceptions.

The six raters throughout this section will be referred to by letter (Raters A through F). Additionally, in order to protect the confidentiality of participants, any occurrences of personal names were changed to names that were gender-equivalent (i.e. replacing a woman’s name with a woman’s name so as not to affect adjective agreement in the text) and phonetically similar in the word-initial position (e.g. names beginning with a vowel were replaced by ones with a vowel so as not to affect elision) before submitting the compositions to raters for scoring. Each composition was given a unique number to allow the researcher to match compositions, ratings, and Exit Questionnaire information to each participant.

Since the appearance of the composition factors several times into raters’ comments, the sample compositions will be presented below as they were given to raters, including converted names and composition numbers, but preserving the original typography (fonts and sizes) submitted by the participant. Because compositions were written in French, explanations and translations (done by the researcher and not an online translator) will be provided in English for specific tokens mentioned by raters or the researcher.
5.5.1 Compositions with Type II errors

As mentioned earlier, raters made a total of 42 Type II errors, incorrectly judging that a translator had not been used when in fact it had. Of these, there were fourteen compositions for which one person rating the composition was able to identify OT usage while the other did not (accounting for 14 errors). In the other fourteen cases, however, both raters gave incorrect judgments (representing 28 errors). The discussion for these errors will focus on what one or both raters may have missed, as well as other factors that could explain not detecting an effect of online translation on the participant’s writing.

5.5.1.1 Compositions with Type II errors by only one rater

One example of a composition where rater opinions were mixed concerning translator usage is below. This essay was written by a participant in Group C on Task One (Appendix E). The composition was written with the aid of an online translator, but only one rater correctly identified OT usage.

8577

Bonjour Marie-Claire

J’aime mon université ici a L’Université d’Illinois! Il y a beaucoup de gens gentil et organizations. Aussi, j’aime les professeurs. ils sont très sympa et intelligents. Ils sont seulement peu d’entre les beaucoup de raisons que vous devriez venir ici. J’aime mone ecole mais la cuisine dans la dorms est méchante! Aussi Les salles de classe sont très grandes!

Le semestre dernier, mes amis et je suis allé à beaucoup de partis. Mais nous avons fait notre les devoirs aussi. Aussi, Le semestre dernier, Je suis allé à beaucoup de jeux de football et basketball. Ils étaient amusants! Il y a beaucoup de faire sur ce campus!

C’est important de venir à mon université parce que vous obtiendrez une bonne éducation ! Il y a beacoup bon professeurs que sont très intelligents. Aussi vous auriez un temps très amusant à ma université! Aucune autre université comparerà.
The first rater (Rater A) did not detect OT usage, simply writing “Good standard French expression.” The only item marked in the composition itself was the word “a” in the first sentence, with the note “ACC,” referring to the missing grave accent that should be present in the word à (meaning “at”) instead of “a” (“has’”). For this rater, no report was made of anything appearing to stand out as signaling translator use.

The other rater correctly judged the composition to have been written with a translator, marking two sentences in particular as being “OT”: “Ils sont seulement peu d’entre les beaucoup de raisons que vous devriez venir ici” (literally, “They are only few from among the a lot of reasons that you should come here”) and “Aucune autre université comparera” (“No other university will compare,” but with an incorrect use of the verb “comparer,” for which the subject is typically the agent). While the grammar is wrong in both cases, Rater B commented that these sentences’ “complexity [was] greater than other sentences in compo.”

The participant reported making frequent use of the translator: “I actually used the translator for quite a good portion of the assignment. This was the first time I have been told to use it so I took advantage. Due to my lack of knowledge in French I found it to be very easy & helpful.” A further examination of the composition offers other possible examples of OT use. In the first paragraph, “dorms” is given in English; it is possible the online translator did not recognize this familiar term, but it is also possible the student just decided to leave the term in untranslated. In paragraph 2, “partis” is an incorrect lexical choice, referring for example to political parties as opposed to celebrations (fêtes). Since the form is similar to English, however, it may instead be an attempt of francization or simply the use of a faux ami (false cognate). Another example in the second paragraph,
“mes amis et je suis allé” (meaning “my friends and I went”), has the incorrect first person singular pronoun (which should be moi) and the verb form agreeing only with the most immediate pronoun instead of the entire subject (mes amis et moi). In the researcher’s experience, although this error is not a definitive sign of OT use, it is more common that students will recognize that the subject is plural and use either a correct or incorrect plural auxiliary (mes amis et moi sommes allés or *mes amis et moi sont allés).

There are also several mistakes that would not be expected from a translator: in the first paragraph, “J’aime mone ecole” (“I like/love my school”) has a misspelled possessive adjective and a missing accent over the word école, while in paragraph 3 beaucoup (“a lot”) is misspelled “beacoup.” The fact that this is correctly spelled elsewhere in the composition could point to translator use on some portions and not others, learner inattention, or incomplete acquisition of this form. The examples given by Rater B, however, are the clearest indication of potential OT usage.

The following composition, written by a participant in Group B for Task One, was correctly judged by one rater but not by the other to be written with the help of an online translator.

#8505
French Study

Marie-Claire,

Bonjour! Ça va? J’attende l ‘Universite d’ Illinois, et c’est le meuiller. L’universite est plus bien que autre universities parce que les professors sont trés bien! Aussi, les etudients sont trés interesent. Mais, moi universite est grande, et je ne sais pas si tu aimes l’universite est grande. Enfin, l’universite moins souvent que autre universities parce que il est expensive.
Durer le semestre, j’attendre le foot-ball matches et leur sont interesent. Aussi, moi et mes amies sont faire du shopping dans la citie et il y a beacoup de magasins. J’ai etudier à la bibliothèque. Enfin, j’ai faire beaucoup d’activités ont offert.

Tu devries considère l’Universite d’Illinois! Il est important parce que tu fais recieve un bon education! Aussi, tu recontre des amis! Quand tu arrives à l’universite, tu seres salué. Si tu ne aimerais pas l’univeriste, tu retourner à la maison. Si tu deciderais attendre l’universite, m’appiler et j’aider tu avec trouver toi classes.

Au Revoir!

-#8505

The rater (Rater F) who judged the composition as being written with a translator noticed one of the “red flags” of translator usage mentioned by Luton (2003), or “yellow flags” as described above: a misspelled English word: “recieve.” Even though participants in Group C had been specifically warned during their training (Appendix L) that online translators have difficulties with misspelled words, this writer from Group B who had received no such warning apparently did not notice in the output that the translator had not successfully processed this word, at least in the rater’s estimation. Given there are other words immediately after the word that incorrect (“un bon education” instead of une bonne éducation) it is possible the participant put the (misspelled) English word on his or her own. Students do sometimes insert English words instead of leaving blanks or attempting to find an equivalent in French: “expensive” in the first paragraph might be a case of this, or the student may have thought “expensive” was a French word (a fairly typical error, in the researcher’s experience). Nevertheless, the presence of a misspelled English word, as noted by the rater for “recieve” here, is usually a good indication of possible OT use.
The second case mentioned by Rater F, however, is less clear-cut. The rater wrote, “their → leur (instead of they’re...).” Looking at the sentence, “j’attendre le football matches et leur sont interessent,” (meaning “I to wait for the soccer matches and there are interesting” but with several errors) it is also possible that the participant confused the possessive adjective “leur” with its corresponding subject pronoun “ils” without the use of the online translator. The presence of other nearby errors (lack of verb conjugation for “attendre,” word order for “football matches,” misspelling and lack of agreement of “interesant,”) would tend to indicate that a translator may not have been used in this case.

The other rater did not judge the composition to be translator-aided. Rater E commented that “grammatical errors, such as subject/verb agreement, adjective agreement, and article usage seem more like student errors” and noted or circled several misspelled words in French (such as “meuiller” instead of meilleur [“best”] in the first paragraph, and “beacoup” instead of beaucoup in paragraph 2). The word “recieve” is circled, but the rater either did not consider this to be an indication of translation use, or noticed the use of English but not the misspelling.

The participant who wrote this composition gave the following self-report: “I used the translator when I was unsure on the spelling of a particular word.” Spelling (presumably in French) was an issue noted by the participant, but the fact that the writer did not use the translator to check other French words may explain in part why one of the raters was not able to detect OT usage, since such errors in French would not be expected from a student using online translator.
Another example of a Task One composition which had been written with the aid of online translation but which divided the raters was the following, this time from a participant of Group C.

#8501

Bonjour Marie-Claire! Ça va?
Je crois que mon université est le meilleur l’université parce que les gens ici sont très sympas, et très intelligents aussi. Les facultés ici sont très bonnes aussi. Il y a beaucoup d’équipes et groupes on peut joindre et plusieurs cours on peut prendre aussi. Malheureusement, les autobus ici sont plus dangereux que les autobus aux autres universités. Les cours ici sont plus difficiles que les autres universités que plusieurs de mes amis ont pris beaucoup des cours à une université différent pendant nos vacances.


Cette décision est ta décision. Dans mon opinion, je crois que tu aimerais cette université le mieux. À ta place, je viendrais ici. Les gens ici sont le plus agréable, et les opportunités sont interminables ici. J’espère que tu fais le bon choix. La bonne chance faisant cette grande décision!!

À bientôt!

Although it may be somewhat difficult to discern when reading on a screen, this is the first case that will be presented where there was a change in font, going from Calibri 11 to Arial 11. Although not definitive, formatting changes such as this are possible signs of copied and pasted content, such as the output of an online translator. Neither rater commented on this, either not noticing or not finding it noteworthy since it occurred at the beginning of the composition and between two short sentences containing commonly used expressions (“Bonjour,” “Hello” and “Ça va?,” “How’s it going?”).
The rater judging that a translator had not been used (Rater A) commented “Good colloquial French expression,” only marking two words on the composition: circling “le” in “le meilleur l’université” and underlining “joindre.” The rater correctly deciding that a translator was used (Rater B) did so with hesitation: it appears the rater first wrote “Yes,” then crossed it out to write “No,” before crossing this out to again write “Yes.” The final choice appears to have been made largely due to the last sentence in the final paragraph, with the rater underlining it and drawing an arrow between it, adding the comment “‘making’ (OT en faisant, I think)? But fewer GEN. [general] errors than most.” The sentence “La bonne chance faisant cette grande décision!!” (meaning, “The good luck making this decision”) is not correct and, as judged by Rater B, appears to be translator-created. Although not mentioned by the rater, the presence of the definite article “la” before “bonne chance” (“good luck”) seems unusual since it is ungrammatical both in French and English. The use of the gerund “faisant” (“making”) could be a calque from English, but since the gérondif is used less frequently in French than its counterpart in English, the student is less likely to have acquired this form as part of his or her active language production. For these reasons, translator use might be suspected.

The fact that it was difficult for judges to identify the use of the translator may be due to the fact that the participant who wrote this composition reported on the Exit Questionnaire using the translator sparingly: “I did actually use the translator for the tasks but only once or twice for each task and only for 1-4 words each time. Mostly for vocabulary or expressions I didn't know.” There are other passages that might be the result of consulting a translator, one example of which is “Les facultés ici sont très bonnes aussi.” If the student meant “The faculty here are very good also,” the online
translator mistranslated “faculty” (le corps enseignant or simply les professeurs) by the false cognate “facultés” (which can mean “schools,” “universities,” or “departments”). However, despite the change in font at the start of the composition, which may or may not have been due to pasting text in from the translator, there seem to be no definitive signs of translator use, which may in part explain the split decision concerning translator use.

The composition below was written for Task Two by a participant in Group C.

Cher Jean-Pierre,

J’entends que vous voulez être ingénieur. Je dois vous décrire la profession pour que tu peux choisir sagement. Si tu aimes les maths, tu aimerais l’ingénierie. Tu peux résoudre les problèmes. Mais, si tu n’aimes pas les maths, ne devenir pas ingénieur. Et aussi, on ne parle pas beaucoup avec les autres. On passe le temps dans un labo.


L’ingénierie est une profession importante dans le monde. Nous construisons des bâtiments, l’électronique, les voitures, et beaucoup d’autres. Le travail est dur, mais tu peux gagner beaucoup d’argent et apprendre beaucoup de choses. La choix est la vôtre.

Sincèrement,

Etudiant 7425

Raters were again divided on this composition. Rater C, who correctly noted translator usage, wrote, “spelling too good.” Comments for a number of compositions by various raters indicated that spelling, accents, grammar, or occasionally vocabulary were “too” good (or similar terminology: “perfect,” etc.) to have been written without translator assistance. There are no spelling mistakes in this participant’s essay, and only
one error with accents: “profésion,” used twice, should have no accent. Rater C also commented that “a couple of translations seem like they’d come from a translator,” without specifying which ones. Another area of concern for this rater was the varying use of formal and informal address; the person scoring the task was “not sure what to make of mixing tu/vous,” including within the same sentence (referring to “Je dois vous décrire la profession pour que tu peux choisir sagement,” meaning “I must describe for you [formal] the profession so that you [familiar] can choose wisely,” as well as the misuse of the indicative mood instead of the subjunctive for the third verb).

The second scorer however did not apparently share these concerns. Rater D commented that “the student would have known all of the vocabulary he uses in the essay.” Based on the participant’s self-report, however, there were grounds for suspecting translator usage: “Yes, I used it prodigiously. I usually used it to check the gender of the words I was using. Occasionally I used it to translate a full sentence, but only if I thought it was a simple, common sentence.” While the composition does show an error in gender (“la choix,” “the choice,” which should be masculine in French), if OT use was “prodigious” it might explain why there are no spelling mistakes and few other form-related errors. Even with this knowledge, however, it is difficult to discern clear cases of translator use. Another more subtle lexical error perhaps not noted or at least not mentioned by raters, “l’électronique” (the term for the field of electronics) instead of les appareils électroniques (for electronic devices, which is likely what was meant here), was probably due to translator use. The low occurrence of errors may help explain why one of the two raters did not believe OT had been used.
A final translator-assisted composition which divided raters will be presented below. A participant in Group C wrote this essay for the second experimental task.

Participant Number: 7433

Bonjour Jean-Pierre,

J’espère que tu es bien! Je suis très excité que tu veux être avocat aussi! J’ai voulu être avocat depuis onze ans. Pourquoi est-ce que tu veux être avocat? Si tu es la même est moi, tu veux aider les gens! Les travaux pour nous sont plus difficile et chronométrer consommer. Mais aussi, tu peux faire beaucoup d’argent et nous pouvons aider les gens qui a les problèmes.

C’est important pour tu parleras avec les gens dans notre profession, parce qu’ils disent que vous la logistique d’est avocat. Aussi, je suis allée à pièce de tribunal et j’ai parlé avec les juges. Ilsissent quelle est important pour toi est un avocat futur. Aussi, quand tu es à l’école tu devrais dans un club d’avocat où tu peux comprendre de notre profession future!

Je pense que aux Etats-Unis a besoin de beaucoup de h’net avocats. Je pense que tu devrais avocat toujours! Je sais que il y a un profession difficile, mais si tu mets le temps bien, c’est facile! Aussi, tu es faire quelque chose bien pour le monde! Tu peux changer le future si tu es avocat bien! Tu aimeras aux Etats-Unis beaucoup, et je pense que tu aussi aimerais ton profession. Alors, je ne peux pas attendre de tu recontrer. Au revoir!

Mieux,

7433!
Rater B, who correctly identified this composition as being written with the help of OT, wrote, “1) ‘time consuming’ 2) the logistics? Mieux (‘Best.’),” with numbers and circles indicating the corresponding parts of the text. The use of “*chronométrer consommer*” (literally, “to time to consume” instead of the intended “time-consuming”) does look to be translator produced or aided: perhaps the original English phrase was “Jobs for us are more difficult and time consuming” and the translator did not know how to parse the final two words, which can each serve multiple grammatical functions depending on context.

It is unclear what the participant meant exactly for the second example cited by the rater: “parce qu’ils disent que vous la logistique d’est avocat” literally means “because they say that you the logistic of is lawyer.” It is difficult to confirm translator use for this sentence, but the presence of the preposition “*d’*” elided and before a verb might indicate a translator-, and not human-, grammatical construction. The last example given by the rater, “Mieux,” appears to be a literal translation of the common English email closing, “Best” (short for “Best wishes” or “Best regards”), which would be consistent with translator use since it is giving a literal translation of a word without regard to context, which it either wasn’t provided or did not correctly detect.

The rater who did not judge the composition as being written with translator assistance did nonetheless have some suspicions. Rater A wrote: “I have some doubts because of words like ‘chronométer consommer,’ or things like ‘la logistique d’est…’ it confuses me, but very incomprehensible at times.” Despite citing two of the three cases the other rater mentioned, however, this rater came to a different and incorrect conclusion about OT usage. The participant who wrote this composition provided the following self-report in the Exit Questionnaire: “I used the translator when I needed to know a vocab
word or phrase I did not know. I did not use it very often — only when needed.” The fact that the student reported not using the online translator frequently might explain in part the indecision and eventual erroneous decision of the second rater.

As can be seen from the examples above, there were some cases where raters had divergent opinions concerning OT use, with one rater commenting on features that the other rater either did not notice or did not judge to be sufficiently suggestive of the influence of an online translator on the participant’s writing. For other compositions, both raters noticed and remarked on the same aspects but came to different conclusions about translator use. Spelling, vocabulary choice, and grammatical forms were among some of the issues noted that factored into rater decisions about translator use, with the number and type of errors playing into the determination.

5.5.1.2 Compositions with Type II errors by both raters

In addition to cases where rater judgments differed, online translation use for other compositions went undetected by both scorers. For the following example, written by a participant in Group B for Task Two, neither rater correctly identified the use of OT.

Bonjour Jean-Pierre,

Si vous voulez faire une demande d'emploi dans un musée d'art, Europe est le lieu meilleur parce qu'il a beaucoup d'artistes et d'œuvres d'art importantes. Mais les États-Unis offre beaucoup de musées d'art aussi. Ce travail est bon si vous aimez l'art parce que vous pouvez étudier que vous aimez. Aussi, vous pouvez voyager beaucoup pour préparer vos sujets et écrit d'eux. C'est difficile d'avoir une famille si vous voyagez souvent, et nous ne gagneons pas beaucoup d'argent. Si vous aimez l'art ceci est le travail pour vous.

Quand j'ai fait les études à l'université, je lisais beaucoup de livres et assistais beaucoup de conférences dans les arts. Aussi, pendant les vacances, j'ai voyagé avec mes amis à Europe
pour voir les grands musées et étudie l'art là. J'ai appris à étudier et écrire bien, et il m'a aide obtenir un travail dans un musée d'art, et apprendre à explorer le monde d'art aussi.

C'est important faire le chose que vous aimez, mais je suggère que vous trouviez un travail dans un musée au Europe. Les musées sont meilleurs, et il y a plus d'eux. Aussi, vous avez une grande sélection d'artistes et les types d'art pour choisir. Cependant, si vous concentrez sur l'art des États-Unis c'est le lieu pour vous. Je pense que vous préféreieriez l'Europe. Bonne chance trouvez un travail!

Sincèrement,

#7424

One scorer (Rater C) mentioned “Word order” as one reason OT use was not suspected. In particular, “le lieu meilleur” in the first paragraph was indicated, meaning “the best place” but with “meilleur” incorrectly positioned after the noun. Rater C additionally cited inconsistent articles and prepositions as reasons for not believing that the writer had used a translator, circling the space before “Europe” in the first paragraph where a definite article should be, and circling the incorrect prepositions before “Europe” in the second and third paragraphs (which should both be en). Rater D found that “nothing in this essay points to the fact that the student might have used a translator.”

With the knowledge that a translator was used, it is still difficult to find traces of its effect on the composition. In the first paragraph, the sentence “Si vous voulez faire une demande d'emploi dans un musée d'art” has an extra space in it between the words “emploi” and “dans” which could indicate that “faire une demande d'emploi” ("apply for a job") or “dans un musee d’art” (in an art museum) were directly translated or adapted from the pasted output of a translator. A similar space is found in paragraph 2 for the phrase “obtenir un travail” (“to get a job”). These may simply be errant spaces due to
the participant taking a pause between thoughts or pasting in his or her own text from another part of the composition, and are not definitive signs of translator use.

Based on the participant’s comments, the translator was not used often: “I used it for certain vocabulary I was almost sure of to see where the accent marks went — also, if I forgot the exact spelling of a tense I would check it on the translator.” Both raters gave the student a score of 5 on spelling and accents, so if the participant’s self-report matched actual performance, it would indicate that the strategy may have been beneficial to the final product without raising the suspicion of OT use among raters.

A member of Group B wrote the following composition for the second experimental task. Both raters incorrectly believed that a translator had not been used.

7487
Salut Jean-Pierre,

Les travaux de physicien dans les Etats-Unis, ils sont les meilleurs dans le monde. Il y a beaucoup de travaux parce que la physique grandit et il y a beaucoup de laboratoires. Malheureusement, beaucoup de laboratoires reçoivent d’argent du gouvernement, alors les travailleurs doivent être américains. Aussi, on doit étudier beaucoup à l’université.

Pour la physique, on peut recevoir au moins un degré du maître. Alors, si on veut être physicien, on doit étudier beaucoup pour les bonnes notes. Aussi, on doit aimer la physique si on veut être physicien. Finalement, on doit pouvoir faire les maths et écrire bien.
Si tu aimes les maths et la physique, tu dois être physicien aux Etats-Unis. Si tu ne les aimes pas, tu ne dois pas être physicien. Dans les Etats-Unis, on trouve beaucoup de travaux et beaucoup d’argent pour la physique. La chose la plus important, c’est que tu aimes ton travail.

Au revoir, 7487

Rater C noted that the “vocab in this essay is quite simple,” remarking that a better translation for “Master’s degree” might have been found with the aid of OT. The other rater, while not commenting on the phrase, drew boxes around “degré” and “maître.” It is possible however that this phrase was in fact obtained via online translation, since both are literal translations of polysemic words that might work in some contexts, but not in this set expression. The fact that the definite article is used with the preposition de might also be a hint of OT: “degré du maître,” literally “degree of the master.” The participant did not use the article in the first sentence, “Les travaux de physicien” for “physicist jobs” (literally “jobs of physicist”) even though it is likely this phrase was not taken directly from the translator given the incorrect plural of travail. “Degré du maître” is in fact the translation given when “Master’s degree” is entered into Free Translation’s translator (SDL, 2012), so the rater’s belief that an online translator might provide a better equivalent (such as maîtrise or the more recent master) is not supported in this case.

Rater D noted the incorrect form “travaux” (“jobs”) in the third paragraph, which is also found in the first paragraph as mentioned above. There are other scattered errors as well (“reçoivent d’argent” instead of reçoivent de l’argent for “receive money”; “la
chose la plus important,” “the most important thing,” with missing adjective agreement) that are typical of intermediate-level learners, and no clear example that would cause concern in terms of translator use. A possible explanation for the inability to discover traces of the online translator may be due to its limited reported use; according to the participant, “I avoided using it for complete phrases because I knew in a secondary way that it would be prone to making errors when translating phrases.” If the participant used the translator for isolated cases instead of longer texts, there simply may not have been enough translated text to be noticed by raters.

Another example of a composition for Task Two where both raters thought no translator had been used but for which the participant (Group B) actually had used OT can be found below:

#7483


J’aime parler avec beaucoup les gens différents. J’ai une passion pour apprendre des cultures différentes. Ma passion aide me reste très intéressé dans mon travail à la YMCA. Aussi, je suis très confortable parler aux étranger et se demander les donations. J’aime le YMCA et je combattrai pour les droits civils de tous gens.

Jean-Pierre, je pense que tu aimais le travail à la YMCA ou un autre NGO aux États-Unis. Le gens sont très cher est ils sont beaucoup l’argent pour faire don de. Aussi, il récompense pour travailler avec les étudiants et les engage dans l’action civile. Si tu voudras le mariage gai, tu travaillais pour un NGO!

One of the two raters for this composition, Rater A, commented simply that the composition showed “Good expression w/some common pitfalls.” Rater B underlined
“j’élève de l’argent” in the first paragraph, literally, “I raise some money,” but using a verb for “raise” in the sense of caring for children or breeding livestock. This rater also bracketed the phrase in paragraph 3 “ils sont beaucoup l’argent pour faire don de” (literally, “they are a lot the money in order to make donation of”), with the rater writing “prep. placement error.” Since the assumed meaning, “they have a lot of money to donate,” does not end in a preposition in English, the ungrammatical presence of the preposition “de” at the end of the French sentence might be a sign of translator usage (particularly considering that the student used the Anglicism “donation” in the preceding paragraph). These and other errors, even though noted by the rater, were not convincing proof that a translator had been used.

In the Exit Questionnaire, the participant who wrote this composition reported relying fairly heavily on the translator, giving the following comment about translation use on the two experimental tasks: “I did use it. I used the translator for phrases I didn't know. I translated things also just to double check my work. For these 2 reasons I probably translated 1/3 of my composition.” With this knowledge, there are other portions of the composition that might be suspected as output from a translator: “NGO” (an English abbreviation for “Non-Governmental Organization” that the translator left in English instead of giving the corresponding ONG in French), “le mariage gai” (spelled with an “i” as the French word for “happy/gay” is, instead of with a “y” used both in French and English to refer to homosexuals), “j’aime faire mes idées devenir de vrais programmes” (meaning “I like/love to make my ideas to become real programs,” with an incorrect use of the causative faire but correct use of the article “de” in front of plural
adjective, a more advanced grammar point that even native speakers do not always produce).

At the same time, there are a number of grammatical, orthographical, and diacritical errors throughout the composition; Rater B underlined “inteligents,” which had proper adjective agreement but whose base form has one “l” instead of two; for “je pense que te aimais” in paragraph 2, the rater circled “te” (the second person familiar object pronoun instead of the proper corresponding subject form tu) and underlined the second verb, which is in imperfect instead of the conditional mode (aimerais, “would like”). There are numerous other minor errors not commented on: accents (paragraph 1, “interessants”), missing plural markers (“le gens” in paragraph 3), etc. The presence of such errors, coupled with the fact that the participant reported using the online translator selectively albeit frequently, may explain in part why raters did not definitively judge the composition to be written with the aid of a translator.

Written by a member of Group B for Task One, the following composition also escaped detection despite the presence of several possible marks of online translation that were not caught by raters.

Participant Number 8503

Salut Marie-Claire!

J’adore l’Université d’Illinois! Les étudiants sont plus intelligents que les étudiants à des autres universités. Aussi le campus est le plus beau. Mais, l’université n’est pas perfectionner. Il y a plus l’argent de police que des autres universités. Aussi il neige tous les jours pendant l’hiver.

Le semestre passé, j’étais active! Il y a 1,000 organisations d’étudiants sur beaucoup de sujets different. Avec les Girl Scouts, j’ai enseigné des filles danses internationales et l’art. Aussi, j’ai travaillé à le YMCA d’université. Le YMCA est un NGO que a commencé dans 1872. Sur le weekend, mes amis et moi, nous sommes sorti à un bar ou un soireé.
Te adorais l’université d'Illinois! C’est un grand école avec beacoup d’étudiants grands! Si tu as étudié ici, tu serais heureuse parce que tu étudiais la chemise avec mon camarade de chambre! Nous nous amuserions!!

--Ta amie

Of the two scorers, Rater C noted errors with spelling, pronouns, and vocabulary as reasons why OT usage was not suspected, while the other (Rater D) commented that the vocabulary seemed level-appropriate. Both raters noted the incorrect word was used for “chemistry,” written as “chemise” (“shirt”) in the text instead of chimie, with Rater D commenting that a translator might have “given a proper translation for ‘chemistry’.” While this is likely the case, there are another mark of possible translator usage that did not appear to draw raters’ attention to OT use: in the first paragraph, the English adjective “perfect” is translated by the French infinitive “perfectionner,” an incorrect usage of a word in a different grammatical category than the original, resulting in a lexical item that a student at this level would be less likely to know. One rater circled this, and another marked a “V” (presumably for “vocabulary”), but it appears neither judged this to be a clear enough sign of translator use to affect their final judgment.

As might be suspected given the vocabulary and topic in this composition, this student is actually the same as the one who wrote #7483 above, and who self-reported translating “1/3 of my composition.” Four different raters scored this participant’s experimental tasks, two for each composition written; none of the four judged that a translator had been used. If it is in fact the case that the participant used OT “for phrases I didn't know” and “to double check my work,” this was done in a way that did not draw the attention of raters as presenting unusual language use for a student of this level on either experimental tasks.
A final case of a sample composition where the two raters believed there had been no OT, when in reality a translator was used, is presented below. This composition was written by a participant from Group C for Task Two.

Bonjour Jean Pierre,

Je voudrais vous parler de mon occupation. Je serai un comptable et je pense que vous devez considérer mon occupation pour vous aussi. Il ya beaucoup des avantages : mon occupation a beaucoup de sécurité puisque le monde toujours aura besoin des comptables. Dans cette économie beaucoup des personnes sont vidés et la sécurité dans l’occupation est très importante. Un autre avantage est qu’on a équilibré entre le travail et la vie. Mais il y a aussi des négatives de cette occupation. Quelques personnes pensent que cette occupation est un peu barbante parce qu’il y a beaucoup des nombres. Aussi beaucoup des personnes ne travaillent pas comme un comptable pour plus de cinq ans.

J’ai prépare beaucoup pour ce travail. Ma spécialisation à l’université était comptabilité. Aussi j’ai fait un stage pour un été au bureau de comptabilité. Une autre chose que j’ai fait pour la préparation est que j’ai parlé avec des comptables avant j’ai décidé mon occupation. Pour avoir l’expérience de direction j’ai participé avec les activités à l’université aussi.

Je pense qu’il est très important que tu remplie une application pour mon occupation aux Etats Unis. Si vous aimez les numéros et les vous aimez travailler avec une équipe vous l’aimerez beaucoup. Une autre option est que vous pouvez remplir une application pour travailler dans une banque ou peut-être avec la spécialisation de finance. J’espère que je vous aide avec mes conseils et bonne chance avec trouver un travail.

Participant #7434

One of the longer compositions submitted, this essay was believed to be entirely participant-produced by both raters. One of the raters, Rater E, cited “errors in articles” as one indication of student error: the three instances of “beaucoup des” (“a lot of the,” instead of beaucoup de, “a lot of”) were circled in the first paragraph, along with the “un” of “Je serai un comptable” in the second sentence (an incorrect calque of the indefinite article likely due to its presence in the corresponding English sentence, “I will be an
accountant.”). The other rater noted the presence of *faux amis*, adding however that “they’re not completely red flags.” Rater F circled “*remplir une application*” in the third paragraph likely to illustrate this point, since neither the verb nor noun are used in the French equivalent of the expression “to fill out an application.” An additional comment was made by Rater F: “The grammar is rather advanced, so if this student used a translator, s/he did it very well.” Unlike some raters, who have suggested in examples above and below that high grammatical accuracy can be a sign of translator use, this comment indicates that a certain level of grammatical sophistication could serve to lessen the likelihood that the rater would suspect OT.

In consideration of this last comment, a look at the writer’s reported use of the translator is interesting: “I used it for key words and a couple of sentences. I used it more with the second composition because I used more complex phrases.” One of the few to note explicitly an evolution in their OT use during the study, this participant mentioned increased use of the translator for the second experimental task. In spite of this, neither rater recognized the translator-aided text as being such. One rater did however note a high level of grammar usage, which would coincide with the participant’s comment that the sentences in this composition were more “complex.”

Knowing that a translator was used, it is still difficult to select phrases that may have been obtained through the translator, although it is possible to suggest some indicators: in the first paragraph, “*occupation*” (instead of *profession* or *métier*) could be the work of a translator or simply the use of a *faux ami*, while the word *vidés* does not have a clear meaning here. Among many things, the verb *vider* can mean “to empty (out)” or “to drain”; the word in context could be a deformation of *viré(e)s*, meaning
“fired.” Another sign of translator use could be that only one sentence contains the second person familiar form: “Je pense qu’il est très important que tu remplie une application,” (paragraph 3, presumably meaning “I think that it is very important that you fill out an application”), but since the subsequent subjunctive is incorrectly spelled, the switch to “tu” may be learner error; it is also not uncommon for L2 learners of French to confuse *tu* and *vous*. Assuming the participant did perform more OT with this task as reported, this use is rather difficult for the raters and researcher to detect, either due to OT accuracy or participant skill in selecting or adapting the resulting output.

In cases where neither rater correctly determined online translation use, participant self-reports varied from minimal to frequent OT use. In some cases, the presence of some errors judged to be typical for L2 learners appeared to lead raters away from some phrases or features that could have indicated the use of online translation; in other cases, the participant may have used the translator sparingly or judiciously enough that there were no telltale marks of OT in the final product.

### 5.5.2 Compositions with Type I errors

Although less common, there were also 16 Type I errors, wherein one or both raters incorrectly assumed translator usage where there had been none. For eight compositions, raters’ judgments differed, while in four cases (accounting for eight Type I errors) both raters incorrectly assumed a translator had been used. Since translator use was not allowed for these compositions, due to the task (Written Pretest or Posttest, in which no translator was allowed across groups) or the participants’ group (Group A

176
serving as a control on the experimental tasks), only the raters’ comments and researcher’s observations will be considered.

**5.5.2.1 Compositions with Type I errors for only one rater**

Below is an example of a composition written by a participant in Group A for Task Two that was misidentified as a translator-aided essay by one rater but correctly judged to be free of OT influence by the other.

7423

Cher Jean-Pierre,

Bonjour! Ça-va bien? Je travaille dans le théâtre. Le théâtre est bon endôit à traviller. Le viva de théâtre est très exantrique. Tu travaille avec bon artistes. Le théâtre est ne bon pas tout temps, les heurs est long et les actrices est mal.


Salute!

Rater A only made a few marks: in the first paragraph, there is the abbreviation ACC over the word “endôit” (for endroit, “place”), indicating the unneeded accent in this word; a line and the letters VOC over “le viva” (presumably meant to be la vie, “life”); and SP above “J’étduie” at the start of the second paragraph, marking the misspelling of
j’étudie (“I study”). These errors, which do not seem typical of OT, likely contributed to the rater’s correct judgment that the translator was not used.

The other rater, however, thought there were a pair of mistakes that did not seem to be the result of human error. Rater B commented: “j’amour le théâtre most/all beginning speakers are very familiar with ‘J’aime ...’ – and would not change to noun form.” While the presence of a word having more than one part of speech in English being used in the wrong grammatical category in French can sometimes be a sign of translator usage, this is not necessarily the case. Anecdotally, the researcher has seen students in test situations (without computer or dictionary access), both in writing and orally, incorrectly use the word “amour” as a verb. Although it would be hoped that novice-high-intermediate-low learners would not make such an error, it appears to have been the case here. A second comment reads, “also, the word ‘make’ in English seems odd, but an OT would have, I think, caught it?” The presence of the correctly-spelled “make,” as the rater wondered, mostly likely would have been translated had it been inputted into the online translator, with the misspelling of other words in the same sentence (“soutaite” for souhaite, “wish,” and “chiox” for choix, “choice”) not disproving but adding additional, potential disconfirmation of OT use on this task.

The following composition written for the Pretest by a participant of Group C was erroneously judged by one of the two raters to be written with the help of OT use.

Participant 6322

Bonjour Bernard,


Rater C, who incorrectly suspected translator use, commented, “Spelling is perfect, ‘long distance’ adj. agreement is not,” with a line drawn towards the first sentence of paragraph 3. While there are no spelling errors, the composition is rather short and, in the opinion of the researcher, not very ambitious in terms of vocabulary or structures attempted. It is possible the participant simply limited himself or herself to words whose spelling was known.

While it is true that translators do not always make agreement with remote adjectives (mentioned in Williams, 2006 among others), in this case the participant appears only to have had ambiguous use of the word vacances (“vacation” in English), sometimes treating it as a singular and other times as a plural. Additionally, there is a case of “short distance” agreement not being made: “la classe été annulé” (meaning “class was canceled”), which also has a missing auxiliary for the passé composé and would seemingly be unlikely to have been produced by a translator.

The other rater correctly determined that no translator was used. Rater D commented, “if this student had used a translator, she might have found a better word for ‘sleigh,’” as well as “bonhomme de neige for snowman.” The presence of English words that are not misspelled, in particular items the student would not be expected to know but which could have entered into an online translator, can be a sign that a translator was not used or at least a reason to reserve judgment until other factors are considered.
Below is a final example of a composition where only one rater made a Type I error, written by a student in Group A for the Posttest.

6320


Les vacances d’hiver est important pour les étudiants. Mais, une moins est toute longue. Les vacances devraient deux ou trois semaine. Après quatre semaine il est difficile retourner à l’universitaire. La vie à la maison avec la famille est important pour étudiants. Tous les temps faisaient difficile partir.

Rater A, who correctly determined that a translator had not been used, explained this decision by writing, “Just had grammar & WORD/vocab choice.” For this rater, none of the errors appeared to be out of the ordinary for a student of this level. The second rater, however, believed that a translator had been used. Rater B wrote the following: “Je n’ai pas travaillé depuis…contains no errors; most other sentences have at least 1; and depuis is a more difficult construction.” This assessment would be dependent, however, on what the participant meant: while the sentence “Je n’ai pas travaillé depuis mes vacances” is correct from a grammatical standpoint (meaning “I haven’t worked since my vacation”), a possible and perhaps more likely interpretation is that the student was instead trying to say, “I didn’t work during my vacation,” which would be expressed in French as Je n’ai pas travaillé pendant mes vacances. It is possible that the rater,
knowing that some tasks were written with the aid of the translator, was specifically looking for suspect structures and overlooked the possibility that the sentence, while on the surface appearing accurate, was actually an incorrect utterance in the given context, the result of a common confusion between the words *depuis* and *pendant*.

Lastly, although the rater did not explicitly say that this influenced the judgment, the word “*vacnaces*” is misspelled and circled by the rater in the first paragraph, while the word is correctly spelled in other occurrences: differences in spelling of the same word can sometimes raise suspicion of translator use. In this case, however, the mistakes represent normal L2 production without the aid of a translator.

### 5.5.2.2 Compositions with Type I errors by both raters

For some tasks, both raters thought they detected a trace of translator use when in fact there had been none. The following task, composed for the Written Pretest by a participant of Group A, is one such example.

Participant 9675

*Bonjour Corine,*

*Salut Corine, comment ca va?* Je pense que l’été pourrait amusant parce que tu n’a pas besoin aller tes cours. Aussi, vous vous dépêcheriez et rester avec tes amis. L’été pourrait être enneuyeux si vous avez besoin de travailler. Aussi, l’été va très rapide lorsque vous vous amusez.

Il faut que avoir un été long parce que les gens auraient plus temps se dépecher. Il est nécessaire s’amuser parce que c’est aider avec la santé. Si nous aurions un été long, nous plus nous amuserions. Enfin, vous n’avez pas besoin aller à l’école trop tôt.

Au Revoir

One rater pointed to the capitalization of “J” in “J’ai fait une croisière” (“I went on a cruise”) in paragraph two as a sign of translator use, although adding “maybe just transfer from English.” Rater A bracketed the entire sentence and wrote, “VOC is good, pronouns are good – (whereas they aren’t / are missing in rest.).” The writer did include the word “croisière” in this sentence, which is not a common word, but it is unclear what the rater meant by the difference in pronouns besides the capitalization of “J” and perhaps the improper construction “Il faut que avoir un été long” at the beginning of the third paragraph. Capitalizing the subject pronoun je is a common error among Anglophone L2 learners of French and while unusual typographical features can be a warning sign of translator use, it turned out not to be the case here. For the second example, the participant may have just confused two structures: il faut que followed by subject and a verb in the subjunctive mood, and il faut followed by an infinitive. Neither of these would generally be good indications that a translator might have been used.

The other rater, who also incorrectly judged the composition to be written with OT, cited “sentences like ‘libre temps’ ‘parc amusant’ or lack of preposition,” with the phrase “plus temps se dépecher” underlined. None of these examples mentioned by Rater B, however, necessarily points to translator usage: the first representing English syntax (“free time”) and likely a calque, whereas translators generally seem to place postpositive adjectives correctly. The second example cited by the rater may be an attempt at a
neologism for “amusement park” (as it does not say “*parc d’amusement” or “*parc de distractions,” which might be expected translator attempts if the term parc d’attractions is not in its word bank). The third example may just be a calque from an English structure which does not require a preposition (“more time to hurry”); this interpretation would be supported by the fact that dépêcher is accented incorrectly, which an online translator would not be expected to do, and the fact that in the first paragraph, “tu n’a pas besoin aller” is, in addition to a conjugation error, also missing a preposition in French where none is used in English: “you don’t need to go.”

It is possible that the false positives for this essay were due in part to the nature of the exercise in which the raters were engaged: scoring compositions for a study about online translation. Raters may have been looking out for instances in which a translator might have been used, sometimes suspecting OT use for reasons that do not appear to be justified in the text.

Below is a Posttest composition written by a participant of Group B which both raters incorrectly judged to be written with the aid of an online translator:

6325

Chèr Bernard,

J’écoute que tu écris un article sure les vacances en hiver aux Etas-Unis. Pour les étudiants, les vacances peut-être très amusants. Dans l’étas d’Illinios, on visite la grande ville de Chicago pour le nuit. Si on a de l’argent, on peut aller à l’état de Floride ou Colorado pour faire les vacances. Les vacances en hiver peuvent être un peu ennerveux aussi. On doit rester avec ses parents (pour garder de l’argent) et visite la famille (Ils m’ennuient). Et aussi, il fait froid! La froide est la mauvais pendant les vacances.

L’hiver dernier, j’ai rester chez mes parents à Mount Prospect. J’ai visité les amis et nous sommes sortis encore. Ma famille est un peu fousse (mais, pas mes parents) et je n’ai pas
voulu les voir beaucoup. Mon grand-père est le plus cool, et nous dinons deux ou trois fois. Je n’ai pas travaillé.

Les vacances sont un peu lents. Je crois qu’ils doivent changer les vacances pour que les étudiants puissent revenir à l’université plus tôt. Ils peuvent créer plus d’autre vacances pendant le semestre. Peut-être ils peuvent créer une semaine pour étudier les examens finaux.

Sincèrement
Student 6325

One rater (Rater A) appeared to vacillate several times on the decision, writing at first “No & Yes” but then crossing out “No &.” The rater gave comments as to the reasons for possible OT use: “sentences like “J’écoute que tu écris” (I hear you are writing…),” as well as arguments against: “No b/c misspellings like fousse,” “bad grammar agreement.” At the end, Rater A wrote, “I have to wonder if the student used a translator on a few sentences. Overall, I’d say no,” but then crossed out the last sentence, and added to the side, “I think the student used it on a line or 2,” but without indicating what a second case might have been. It should be noted that the line that rater singled out, “J’écoute que tu écris,” actually contains what is anecdotally a fairly common learner mistake: confusing the verbs écouter (“to listen”) and entendre (“to hear”), that a human writer would be more likely to commit than the machine translator.

The second rater, who also judged the composition’s writer to have been aided by an online translator, commented: “Some non-native speaker of French errors (e.g. faire les vacances), but also perfect use of the subjunctive ‘pour qu’ils puissent…’).” As seen in the discussion for Type II errors, the presence of grammar that seems to be too accurate was sometimes perceived to be a sign of OT use. It is not uncommon however
for students to have some amount of variation in accuracy for various structures they are acquiring. While the expression “faire les vacances” (literally, “to make/do vacation”) appears unusual, it is not clear what the rater thought would have been the original English expression that would have caused the online translator to produce this sentence. It is possible the writer was confusing faire un voyage (“to take a trip”) and aller en vacances (“to go on vacation”), although it cannot be known for certain. Other errors, such as the misspelling of États-Unis (“United States”), the lack of agreement between “vacances” and “amusant” (“fun”), the nonexistent word “ennerveux” (possibly an unintended mix between ennuyeux “boring” and nerveux “nervous”), among many others in this paragraph and others, appear to be more consistent with L2 learner errors than online translator influence.

Another composition was misjudged by both raters scoring it. This essay was also composed for the Posttest, but this time by a participant in Group A.

6330

Bonjour Bernard,

Je m’appelle Jacques et je suis étudiant à l’université d’Illinois. Les vacances en hiver est fantastique pour les étudiants américaines parce que nous n’allons pas à l’université pour trente jours et nous voulons notre familles. Mais vacances n’amusant pas quand tu ne peux pas vouloir votre amis. Plus tu ne peux pas aller aux boums à l’université.

En hiver, je suis allé avec ma famille à ma grand-mere maison. Aussi je sus allé au cinema avec mes amis puis nous sommes allés faire du ski un soir. Mais j’ai travaillé aussi pendant les vacances en hiver.

C’est très importante pour les étudiants aux universités avaient vacances longues parce que nous travaille beaucoup pour l’université. Aussi très étudiants vont l’universités plus que cinq heures des familles. Les vacances longues sont bonnes pour les étudiants parce que ils peuvent vouloir leur familles ou amis.
Both scorers (Raters C and D) cited the phrase “ma grand-mère maison” as indicative of translator usage. However, an error such as this was not encountered in the literature describing OT, nor by the researcher in sample sentences. The construction mirrors the corresponding English word order (“my grandmother’s house”) while dropping the possessive for “grandmother”. One possible explanation for this error is that since the phrase already contains one possessive (“ma”), or because the student did not know how to process the other type of possessive (normally with the possessor after the noun in French, preceded by the preposition de), the participant did not formulate this phrase correctly and reverted to English word order. Free Translation, on the other hand, does render the phrase correctly into French: *la maison de ma grand-mère* (SDL, 2012).

Rater C notes that the spelling and accents are “nearly perfect” and that the use of “vouloir” in paragraphs 1 and 3 “for ‘miss’ seems like a translator.” However, it is not clear that the student meant “miss;” they may instead have confused vouloir (“to want”) and voir (“to see”), an error that an online translator would not have been likely to make. The second rater cited the presence of “‘boum,’ which is slang” as an additional indicator of OT. The presence of a single term, particularly one that is likely to have come up in the students’ previous coursework or composition writing prior to the study, does not necessarily indicate translator usage, particularly if a term is more colloquial or slang. While there is an unexpected extra space in the final sentence between “pour” and “les étudiants” (“for [the] students”), the break is not at a point where online translation would be expected since both of these words are well-known and also used together in the first paragraph. Other lexical choices (confusing “très,” “very,” for beaucoup de, “a lot of,” in paragraph 3) and grammatical errors (mismatch in number between possessive
adjective and noun for “notre familles,” “our families,” in the first paragraph) seem typical of student error. Despite raters’ judgments, there was in fact no translator used permitted on the Posttest.

Although less frequent than Type II errors in this study, the presence of Type I errors might initially seem alarming since it may indicate that some instructors are incorrectly identifying student writing as being written with the aid of an online translator when it is actually not. While this may be true in some cases, the presence of a relatively small number of false positives in this study (33s out of a total of 212 compositions written without the aid of a translator) may partly be explained by the fact that the raters were specifically looking out for OT-generated errors in the context of a study investigation online translators.

5.5.3 Compositions correctly judged

As mentioned in the Chapter 4, raters accurately judged whether or not a translator was used 70.70% of the time. A number of ways in which raters correctly identified OT usage have already been discussed through an examination of Type I and Type II errors. A brief sampling of cases where both raters correctly identified translator use or non-use will be considered for additional discussion.

5.5.3.1 Compositions with translator use correctly judged

The following composition for Task One was written by a participant in Group B with the help of an online translator, and raters correctly classified it as such.
Bonjour Marie-Claire,

L'université d'Illinois est mieux que les autres écoles parce qu'il a des gens amicaux. Aussi, l'université a de grands universitaires et un beau campus. L'université coûte cher et il sent comme poop de vache quelquefois. Cependant, l'université est un grand lieu pour apprendre et rencontrer de nouveaux gens.


Marie-Claire, c'est important vous venez à l'université d'Illinois si vous pouvez recevoir une bonne éducation. Le campus a beaucoup d'argent est investi dans lui et tous les nouveaux bâtiments sont de pointe. Le campus est plein d'occasions passionnantes et de grands gens. C'est un hasard merveilleux pour venir à l'université. Marie-Claire, venir à l'université d'Illinois, vous ne le regretterez pas !

Based on rater comments, it was not difficult to determine that an online translator was used. One instructor scoring this composition, Rater E, noted the “advanced grammatical constructions,” citing the phrase “de grands universitaires” which features the correct use of the article “de” in front of a plural adjective (even though the word is presumably a mistranslation of “academics” referring to scholarly endeavors, as opposed to “scholars” as the French text would indicate). Two translator-influenced lexical choices mentioned by the rater are “j'ai joint une fraternité” for “I joined a fraternity” (where in French you would normally devenir membre, or “become (a) member,” of an organization) and “barres” for a drinking establishment (which in proper French is spelled as in English). Rater F also noted this error, as well as “degrés” (instead of notes to mean an academic “grade”) and “mare” (instead of piscine, for “(swimming) pool”).
While the raters noted the examples listed above, others can be offered as well: in the first paragraph, “un grand lieu” for “great” place (lexical error); in the next paragraph “Durer” instead of Durant (the translator seemingly mistaking the preposition “during” for a gerund and translating it with an infinitive in French) and “j’ai obtenu toujours de bons degrés” (an attempt to say “I always got good grades”: in addition to the lexical errors, the adverb is misplaced).

The fact that the two raters successfully recognized translator use is not surprising, given the writer’s self-report for OT during the experimental tasks. “I did use the translator. I wrote in the text box the paragraph I wanted translated. I wrote out the story in English and had it converted to French. I used it often, as I thought was supposed to be the case.” Based on the participant’s comments, it is possible that the entire composition was translated by Free Translation. No instruction was given to participants to use the translator “often”: the prompt stated, “you are to use the translator located at http://freetranslation.com to help you in writing your composition.” It was expected that, just as is supposed to be the case in real-world usage, some participants would decide to seek help through using the translator more often than others. This participant appears to have interpreted the instructions to mean frequent use, which may explain in part why both raters correctly identified the influence of the online translator on the composition.

The composition below, also for Task One, was submitted by a participant in Group C and correctly identified by both raters as written with the help of an online translator.
1. Bonjour Marie-Claire! Mon université je bon parce qu’il y a tant de types différents de gens ici. C’est aussi bon, parce qu’il fournit une bonne éducation. C’est les écoles moins désirables qu’autres, parce qu’il obtient très froid ici. C’est aussi moins désiré parce que c’est si grand il intimide quelquefois.

2. Il y a beaucoup de choses amusantes à faire sur le campus. Durer le semestre mes amis et je suis allé à tous les jeux de football. C’était vraiment amusant. J’ai joint aussi une association d’étudiantes et rencontré beaucoup de gens vraiment agréables. Une autre chose amusante que nous aimons faire est va du patiner de glace sur les weekends. Il y a toujours tant de choses différentes à faire ici !

3. Je pense que vous devriez venir à mon université parce que je pense que vous l’aimez tout comme je fais. Il y a toujours des gens hors et toujours choses à faire. Aucun t l’est seulement un lieu amusant, mais il fournit une très bonne éducation pour les étudiants. Les professeurs sont très amicaux et serviables. Je pense vraiment que vous l’aimez ici.

The first scorer (Rater A) noted the mistranslated “il obtient très froid” for “it gets very cold” (lexical error for the verb: obtenir instead of devenir) and the sentence “c’est si grand il intimide parfois”, for which the online translator mistook the English “it is intimidating” as a present progressive (and thus translated into French as the present) instead of a copula followed by an adjective ending in –ing (which, unlike many circumstances, can simply be translated in this case by the corresponding gerund-derived adjective in French: intimidant).

Rater B commented, “tant de = sophisticated for a speaker who makes an error like je bon.” After some reflection, the researcher discovered that “Je bon” (literally, “I good”) in the first paragraph is likely a translator-induced error due to a typo in English: “Mon université je bon” would be “My university i good,” with the translator interpreting the “i” to be the personal pronoun “I” when the writer presumably meant “My university is good.” The rater also commented on the misuse of the preposition in “sur les week-
ends” (paragraph 2), although this usage can also be found among students not using a translator. Another phrase, in the third paragraph, merited a question mark by the rater: “Aucun t l’est seulement un lieu amusant.” The researcher believes this to be yet another mistake due to typographical error: in this instance, a misplaced space in the text written in English and then presumably entered into the online translator: “No t only is it a fun place” (instead of “Not only…”) yields the phrase found in the participant’s writing with the stray “t” (SDL, 2012).

The writer reported on the Exit Questionnaire, “I used the translator for the writing tasks, but only for specific phrases I did not know how to say.” A phrase such as “my university is good,” however, would be expected to be in the linguistic reach of most students at this level. In addition to the ones listed above, another such phrase may have been “vous l’aimerez tout comme je fais,” (for “you’ll like it just like I do”). Assuming this is a case of OT, the translation did get the idiomatic expression “just like” correct, which would likely be advanced for a student at this level, but did not correctly translate the auxiliary “I do” as a disjunctive pronoun as would be correct in proper, colloquial French. Although the participant reported using the translator “only for specific phrases,” the resulting composition had enough marks of translator usage to catch the raters’ and researcher’s attention, including two mistyped phrases leading to what would likely be incomprehensible output to someone not familiar with English or OT.

For the following composition, written for Task Two by a member of Group C, both raters again correctly identified translator usage.

7484

Bonjour Jean-Pierre,
J’espère que tout est bon. Aussi, j’espère a devenir un ingénieur un jour. Dans ingénierie, on gagner plus d’argent. Dans ingénierie, on peut faire une différence dans le monde. Ingénierie est bonne, mais le travail est très difficile. Aussi, l’argent est les moins important, si ce n’est pas la meilleure carrière.


C’est important à devenir un ingénieur parce qu’il y a beaucoup de problèmes dans le monde. L’ingénierie aide les problèmes. L’Amérique a des bonnes écoles. Vous devez venir à l’école à Amérique.

Au revoir,

7484

Although it is not as easy to see on a screen, both raters noticed on the printed copy they scored that there is a change in font size starting in the middle of the last sentence of the second paragraph (after the comma) and continuing until the end of the composition. As noted above, such a change in format can be an indication that outside text, e.g. that from an online translator, was pasted into a document. One of the raters (Rater C) noted that after the change in typography, there are “no missing articles/accents.” The other rater notes the use of “the word ‘emploi’ instead of ‘travail’ which might be more appropriate for this level of French.” While it is possible that not all the text after this point is translator-generated — the phrase “L’Amérique a des bonnes écoles” (“America has good schools”) stands out, for example, since other samples of presumed OT output have contained the correct usage of the article de before a plural adjective — and while there are a number of errors prior to this that do not appear to be caused by online translator usage (such as missing accents on the preposition à in several
instances), the change in font size was the salient feature that raters reported that led them to believe a translator may have been used.

While it is difficult to say, the comments of the participant about translator use again may not necessarily line up with the observed output on the composition. The student wrote, “I used the translator like a dictionary. I often have trouble saying exactly what I want to say in French so I change to another sentence, the translators [sic] allowed me to find the word or two I didn't know and make the essays flow better.” It would appear that towards the end of the essay, the participant may have used the translator for more than just an isolated word or two. It appears that the student used OT to obtain the translation for “vous devez obtenir un emploi,” (most likely by inputting the five-word phrase “you have to get a job”), pasted the results into the composition, and then continued typing on his or her own without noticing the change in font size. Regardless of the exact amount of OT used, it was enough for raters to correctly identify its effects.

A final composition which was judged by both raters to be written with the aid of OT was the following, produced by a participant in Group C on the second experimental task.

#7481

Bonjour Jean-Pierre!

Je pense que c’est très important pour toi, tu poser ta candidature pour un emploi ici. Je crois ça parce que les opportunités ici sont plus grandes que les opportunités en France. La technologie ici est le plus avancée. Aussi tu peux apprendre le plus ici parce que les écoles médicales sont le mieux. Dans mon opinion, tu devrais travailler ici, mais la décision est jusqu’à tu.

While there are other features of this composition that merit discussion, this composition was another example where the font size changed, although only one of the two raters noted this visual clue that a translator may have been used. The scorer (Rater E) commented, “Although I’m not sure if the entire essay was written with the help of a translator, I’m pretty certain that the last paragraph was […] largely because the font changes.”

One item that both raters noted involved what seems to be an OT error in the final sentence: the expression “jusqu’à tu” (literally, “up to/until you”) standing out as an indicator of translation use, with Rater F adding the comment “ha ha!” for an error that would not typically be made by a student, particularly if not given access to a dictionary or translator. The translator training session that this participant from Group C underwent warned about translator errors with converting English prepositions and particles to French. Although this specific usage was not included in training, examples such as “write down” and “What’s up?” were (Appendices K and L). Either the participant did not retain this information from training or did not see this as a similar enough case to be vigilant about possible errors in the output.

Rater G also noted “j’ai recommandé trouvant” (an ungrammatical way of expressing J’ai recommandé de trouver, or “I recommended finding”) as another OT-produced sentence. Although occurring before the change in font size, the presence of a gerund in French, incorrect usage here for a verb form not normally emphasized in
beginning and intermediate textbooks and relatively limited in use compared to the English equivalent, may be an additional sign of translator use.

The participant’s self-report mentioned using the translator “mostly for vocabulary or expressions I didn’t know,” but it is possible actual use may have been more extensive. It is interesting to note that the font change in the third paragraph occurs in mid-sentence, and that most errors stop at this point. By contrast, the first paragraph sees among multiple errors: ones with verb forms (the second verb of “Je peux travailler,” “I can work,” being conjugated presumably in the present instead of left in the infinitive) and pronouns (“je peux leurs aider,” I can help them, with the incorrect object pronoun). The second paragraph also shows incorrect pronoun usage (“L’était” with the direct object pronoun l’ instead of the demonstrative pronoun e’) and placement (“ont m’aidé” instead of m’ont aidé for “helped me”), all of which unlikely to be caused by online translation. The contrast between these types of errors, which raters appeared to expect from a student of this level, and errors in the third paragraph, containing errors not believed to be typical, would suggest participant use of the translator for some but not all of the text, an observation in line with the first rater’s comments.

Raters were correctly able to identify a number of essays as being written with the aid of an online translator, noticing a number of orthographic, lexical, grammatical, and even visual features of compositions that suggested OT influence on L2 writing.

Participant self-reports for translator use sometimes seemed to align with observed characteristics of their writing, while it is possible that some participants, whether knowingly or not, may have understated the extent to which they availed themselves of the online translator during the writing process.
5.5.3.2 Compositions with translator non-use correctly judged

Lastly, a brief discussion will be presented about compositions for which no translator was used and raters correctly deduced the lack of OT influence. A participant in Group B wrote the following composition for the Pretest.

9605

Corine,

Bonjour! Las vacances dans Étas Unis est malheureuse et drôle! L’ été est malheureuse parce que tu amis travile. Aussi, tu travelies. Mais, l’été et drôle parce que il y a activités tu faire! Las vacances dans Étas Unis est très important!

Quelle été, j’ai travelier au “golf course”, et j’ai aimer la travelier! Parce que, moi (boss) était sympa et intérsente. Avec moi famille, nous avons traviler à la Wisconsin, parce que nous veuonts vouloir mes grand-parents. Avec mes amis, j’ai faire du shopping et nous avons manger à las restaurants!

Quand l’été, une vanance est important! Quand tu es (on) une vanance, tu ne pense pas l’école! Aussi, tu fais bronzer et tu livres une roman! Enfin, tu vouloes toi amis tu ne vouloir pas quand nous sont l’école! Et, tu travelies, mais tu économie la monie.

As both raters corrected determined, no translator was used for this composition.

Raters C and D each commented on the numerous misspellings present in the composition, with each of them circling or marking with an “S” over a dozen misspelled words throughout the entire composition. Additionally, one of the scorers (Rater C) noted the use of “English in parentheses” and “franglais” (a mix of French and English) as indications the writer did not use an online translator. Examples of this can be seen in paragraphs 2 and 3: “(boss)” and “(on),” as well as the use of “golf course” in quotation marks in the second paragraph. Although not commented on by raters, but as seen in a
few compositions, there is also some possible influence of Spanish (with the recurrent presence in this essay of the definite article “*las*” instead of the French *les*, for example).

In addition, Rater D commented that the student “probably did not use a translator because there are so many misused/misspelled words.” Similar to previous cases, the student’s production was judged not to be aided by OT due to spelling errors committed by the participant. It is interesting to note the participant’s consistent use throughout the composition of various morphological forms of the incorrect “*travelier*” instead of *travailler* (to work), while reserving the form “*travile*” as an apparent idiolectical *faux ami* for “to travel” (*voyager* in proper French). While surprisingly systematic, these errors would suggest an online translator was not used, as the translator would likely have either provided the correct form of the words, or a correctly-spelled but misused French word.

The composition that follows was written for Task Two by a participant in Group A and was correctly judged by both raters to have been composed by the student without use of an online translator.

7420

Bonjour Jean Pierre! Mon emploi est très bien! Je dépense mon argent tout le jour!
Les personnes sont très sociable! Mais, le travaille est difficile. Les heures sont longues.


For this essay, Rater D noted the presence of an untranslated English word in the second paragraph as a sign that the participant was not allowed translator access: “If this
student had used a translator, she would have found a better translation for
‘accountancy.’” As seen previously, participants using the online translator reported
making use of OT for unknown vocabulary words.

This rater also commented that had a translator been used, “spelling/verb
conjugation might’ve improved.” Some items underlined by the scorer include “j’ai
interresè” in the second paragraph (the participant’s version of “I am interested,” with
incorrect verb choice that puts the sentence in the active instead of passive voice, and a
form of intéressé(e) with incorrect accents, spelling, and possibly agreement) and “si
vous veux” in the last paragraph (with the formal second-person pronoun mismatched
with the informal second-person singular verb conjugation). Rater C also noted these
“spelling errors, verb forms that don’t exist” and “SVA” (subject-verb agreement). The
fact that these and similar mistakes are present without any signs that normally call
attention to possible translator use appeared to have led both raters to determine correctly
that no translator was used.

This composition, not written with OT, was submitted by a participant from
Group A on Task One.

Salut Marie-Claire! Ca va? Ca va bien. Mon universitaire est le plus de monde. A la
universite de l’Illinois il y a jolis hommes. Illinois est le plus parce que il y a tres bien clubs. Illinois
est moins parce que tres grand. Aussi il y a les cours que sont difficil.


For this composition, Rater B judged that a translator was not used in part due to “spelling errors.” In the first and third paragraph, two incorrect attempts at the word difficile (“difficult”) are circled, a word that poses many learners issues in spelling. The word bien is misspelled (“bein”) and misused in paragraph 3, typical of inconsistency found among learners who have incompletely acquired forms. The word “beer” is underlined in the second paragraph as an example of an “English word spelled correctly” inserted into the composition by the learner. While there is the presence of a misspelled word from English in the first paragraph, “l’Illinois,” the fact that it is part of a phrase that has errors in French (“la universite de l’Illinois,” an attempt to say “the University of Illinois” but with missing elision on the first definite article and the absence of an acute accent on université) weigh against the possible influence of OT here.

Rater A mentioned that “lots of franglais” was found in the composition. While no examples are given, this is perhaps in part due to the word “beer” and a number of American or local references in the composition (“clubs,” “Phi Sigma Pi”). The penultimate sentence contains a preposition at the end of a sentence, which also may be a sign of influence from the participants’ native language. Earlier examples of misplaced prepositions in other compositions were found before verbs and were considered likely to be an effect of OT use since they did not correspond to prepositions present in the original corresponding expressions in English. This instance of an improperly placed
preposition however is in sentence-final position and in a context where modern
colloquial English would use a preposition (“clubs you dance in/at”). In the researcher’s
estimation, this mistake would be characteristic of learner error not caused by the effects
of online translator use.

Lastly, the essay below, correctly identified as not written with an online
translator, was composed for the Posttest by a participant in Group C.

#6383

Bonjour Bernard! Comment vas-tu? J’adore l’hiver! Pour les étudiants américains, nous n’avons
pas les classes alors beaucoup des étudiants rendent visiter ses parents. Aussi, les Américains
adorent Noël ! Mais, quelques étudiants n’aiment pas l’hiver parce ils doivent travailler. Pour
exemple mon ami doit travailler à le McDonalds. Finalement, quelques étudiants n’aiment pas
l’hiver parce qu’il est long et ses parents ne sont pas drôle.

L’hiver dernier, je suis allée mes parents à Nashville. Je m’ai détendu parce que à l’université je
étudiais toujours. J’ai lu quelque romans et j’ai regardé le télé. Aussi, j’ai dormi beaucoup de
temps ! Avec mes amis, je suis allée au cinéma et aux restaurants. Nashville est très amusant à
le soir. À ma maison, j’ai parle avec mes parent and sur Noël, nous sommes allés le cathédral.

Je pense que les étudiants doivent avoir une longue vacance d’hiver. L’hiver est un temps pour
se détendre. La vie à l’université est très hectique et beaucoup des étudiants ne dorment pas
souvent! Bernard, l’hiver est un temps pour l’amusement. Est-ce que tu veux travailler sur
Noël ? Non, tu ne veux pas! Il est important que les étudiants aient une longue vacance d’hiver.

Rater D offered the following thoughts on this composition: “This essay uses
vocabulary commonly taught in basic university French classes. It doesn’t use English
but it doesn’t use overly complicated French expressions, either. I would expect an essay
like this from one of my Intermediate II students on an exam.” Some typical errors noted
or underlined in the text of the composition include the following example in paragraph
one: “beaucoup des étudiants rendent visiter ses parents” (an attempt to say “a lot of students visit their parents” / “[...]pay a visit to their parents”), with the infinitival form of the verb visiter instead of the noun visite, along with a missing preposition à needed in this expression and an incorrect possessive adjective with a singular instead of plural possessor. For the second paragraph, “je étudiais” (“I studied”) has a missing elision of je before a vowel sound. In the third paragraph, “une longue vacance” (meaning “a long vacation”) represents an improper and L2 learner-authentic use of the singular (which means “vacancy”) in lieu of the plural. All of these errors, in the opinion of the rater, are ones that students do not need an online translator to make, with the rater suggesting this essay falls somewhere between other comments seen previously suggesting participant output was too good or too bad to be written with OT.

The second rater, while also judging that a translator had not been used, focused on “verb forms” and “accent that don’t exist in French.” For the former, Rater C circled “Je m’ai détendu” (“I relaxed”) with an incorrect auxiliary for a pronominal verb, an error that had been underlined by Rater D; several examples of “á” (instead of à, at/in) were circled throughout the composition by Rater C in support of the assertion that there were diacritical marks included in the essay which are not found in French. For this rater as well, these mistakes did not result in suspicion of OT usage. Other errors not pointed out by either rater (“beaucoup des étudiants” in lieu of beaucoup d’étudiants; a missing que in one instance of the expression parce que, “because”; “Pour exemple” instead of “par exemple” to mean “for example”; missing plural agreement on drôle, “funny”) are all relatively minor mistakes that an online translator would not probably make and which would not likely cause comprehension issues if encountered by a native speaker.
5.5.4 Summary of writing samples and comments

This initial attempt at a qualitative description of possible effects of online translator use on student writing revealed that a number of factors, including spelling, vocabulary, and grammatical accuracy and sophistication, as well as visual cues such as font size, led raters to their judgments concerning online translation use. As seen above, while the presence or absence of OT use was correctly identified in the majority of cases by both raters, judgments in some cases differed among raters, and in several cases both raters gave an incorrect decision on this. Despite raters undergoing a one-hour training session, this result was not entirely unexpected given the facts that this was the first known study of its kind to ask raters to judge student compositions for translator usage, and that no foolproof way of detecting the various ways and extents OT can be used has been discovered, if even possible. In a statistically-significant number of cases, however, raters agreed on whether or not a translator had been used. A discussion of a sampling of rater comments and participant self-reports showed that online translation use was generally consistent with participants’ description of their use, although in several cases participants reported using a translator to a lesser or greater extent than raters or the researcher could detect.

5.6 Conclusion

After a discussion of results obtained during the study, a summary can now be presented of the findings as they relate to the research questions (Table 5.7). Two of the five questions outlined prior to the study can be answer affirmatively. First, raters were able to tell with statistical significance the difference between compositions written
without the aid of a translator and those written with; even though there were a large number of Type II errors, overall raters correctly judged translation use to a statistically meaningful level. Second, the use of online translation positively affected raters’ perception of participant performance on some specific features (Overall Comprehensibility, Content, Spelling and Accents, and Remaining Grammar) for one or both experimental tasks.

<table>
<thead>
<tr>
<th>Research question 1a: Does the use of online translators by L2 writers result in quantifiably different global scores on compositions written in the L2?</th>
<th>Finding</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>No (Null hypothesis confirmed)</td>
<td>Group C (+translator +training) had statistically higher global scores than Group A (-translator -training) on Task Two. No other significant differences globally.</td>
<td></td>
</tr>
<tr>
<td>Research question 1b: Of students who use an online translator, do those who have been trained in the use of online translators achieve higher scores on L2 compositions than those who have not received such training?</td>
<td>Finding</td>
<td>Notes</td>
</tr>
<tr>
<td>No (Null hypothesis confirmed)</td>
<td>No significant differences found between Group B (+translation –training) and Group C.</td>
<td></td>
</tr>
<tr>
<td>Research question 2: Does online translator usage affect perceived performance on any specific features of L2 learner writing, including comprehensibility, vocabulary, syntax, grammar, spelling, and content?</td>
<td>Finding</td>
<td>Notes</td>
</tr>
<tr>
<td>Yes (Null hypothesis rejected)</td>
<td>Group B and/or C significantly outperformed Group A on comprehensibility, grammar, spelling, content on one or both experimental tasks. No statistical difference among groups for vocabulary or syntax.</td>
<td></td>
</tr>
<tr>
<td>Research question 3: Can instructors detect with statistical significance the difference between a text written with the aid of an online translator and one written without such aid?</td>
<td>Finding</td>
<td>Notes</td>
</tr>
<tr>
<td>Yes (Null hypothesis rejected)</td>
<td>Very highly significant correlation between judgment and actual usage overall. Statistically significant correlation on Task Two, and very highly significant correlation on other three compositions.</td>
<td></td>
</tr>
<tr>
<td>Research question 4: Do student self-reports over the positive or negative effects of online translator usage correspond to observed effects of online translator usage on L2 writing?</td>
<td>Finding</td>
<td>Notes</td>
</tr>
<tr>
<td>No (Null hypothesis confirmed)</td>
<td>Qualitative analysis showed some self-reports fit with quantitative results obtained, while others did not, or appeared to go counter to results.</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.7 Research questions and findings

Three other research questions must be answered negatively. Overall, online translator use did not produce quantifiably different results on global scores, although the group that had translator training and access (Group C) outperformed the control group (Group A) on the second of two experimental tasks. Of the two experimental groups, there was no statistically significant difference between the one that received training and
the other that did not (Group B); although Group C was the only one to significantly outperform the control group on a global score, and more often than Group B had a higher mean on component scores as compared to the control group, there was never a significant difference globally or by component between Groups B and C. Lastly, participant perceptions of the effects of online translation use, although sometimes aligning with actual results, did not consistently do so.

Of remarks made by participants, there is one final type of comment that was found to be particularly important and which may also have influenced some views on the effectiveness of online translation. Four respondents addressed whether or not they believed the online translator would be helpful to them beyond the tasks they had completed. Of those who mentioned the issue, all of them had a negative outlook on any long-term benefits they may have received. Of these, three respondents mentioned the problem of retention: they did not believe that they had actually learned or could remember anything from their use of the translator. Another participant felt that his or her writing ability had not improved through OT use, without specifying however any specific area. Although participants using a translator did outperform those in the control group in some components for Tasks One and Two, and the group with training in online translation had a higher overall mean than the control group, in the subsequent Posttest all groups returned to being statistically similar in performance overall and on all component scores. While further research would be needed to confirm this, any positive effects of online translation on student writing may be short-lived and possibly confined to the task at hand, with several participants themselves specifically reporting that they
felt they had not learned or retained anything new, or that their writing had not improved, after using an online translator.
Chapter 6

Conclusions

6.0 Presentation of chapter

This final chapter will outline limitations of the current study, propose a number of avenues for future research, and offer some implications and conclusions concerning the future role of online translation as it relates to second language acquisition and pedagogy.

6.1 Limitations

This section discusses several limitations to the current research which must be taken into consideration.

The first of these concerns the population pool chosen for the study. Only students who were in their final semester of French, with no intention of pursing future coursework in French, were recruited for participation. This restriction was added out of an excess of caution, as there was no previous study found detailing any long-term effects of translation use that could have given participants in the translator groups an unfair advantage or disadvantage as compared to other students who had not been allowed to use a translator, both in the study and in the general student population. The study was done near the end of the semester, after all assigned compositions for the courses had been completed, so as to affect future performance as little as possible. Since participants volunteered for the study, there may also have been a self-selection bias since students with less interest in the topic of online translation, or with preconceived ideas about its effectiveness, may have been less likely to participate and chosen instead to do the
alternative extra credit. In addition, all participants were third- and fourth-semester French students from the same institution, meaning caution must be exercised before attempting to generalize the results to other levels, languages, and instructional settings.

Potential limitations also involve the make-up and backgrounds of the participants themselves. While groups were not found to be statistically-significantly different from each other in reading comprehension or writing at the start of the study, participants had varying experiences with French, others language, and translators. One participant reported starting to learn French as early as sixth grade, while others had just started within the previous year, and yet others not reporting their previous experience with French. Some students listed no other languages besides English and French being spoken or learned, while others had experience in ones as varied as German, Gujarati, Hindi, Japanese, Spanish, Telugu, and Wolof at various levels of fluency. Participants also reported a number of different majors and minors, including undecided, although interestingly there were no reported foreign language majors (possibly due in part to the restriction against further coursework in French).

The study’s design also had some limitations imposed by the number of participants. Due to the small number of participants both able and willing to participate, it was not possible to have a fully-balanced design. Under ideal circumstances, and as originally envisaged, it would have been desirable to have four groups for the study: -translator, -training; -translator +training; +translator -training; and +translator +training. Had participants been divided into four groups for the current study, however, any results obtained would have been less likely to be meaningful since each group’s n-size would have been below ten. Since it was not anticipated that translator training
would have a significant effect on a group which would not have subsequently been
allowed to use the translator, this condition was eliminated. Additionally, there were
originally 34 participants in the study, but one student did not attend the final session.
This would have left 33 participants and equal groups, and attrition of one participant in a
five-week study of this size is not high. However, another student was excluded since his
or her native language was not English. As a result, Group A (the control group) had one
fewer participant \( n = 10 \) than the other groups (each \( n = 11 \)), resulting in slightly
uneven groups.

Additionally, time and logistical constraints affected the Written Pretest and
Posttest. Most parts of the study were completed during 50-minute class periods, so the
study needed to be limited to five weeks in order to ensure students could complete
requirements for their course. On Tasks One and Two, participants performed a 20-
minute prewriting activity similar to ones used in the Department for other composition
assignments. However, a prewriting activity was not possible for the Posttest since
students needed to complete the Exit Questionnaire afterwards that day. One option
would have been to eliminate the prewriting entirely, including for the experimental
tasks; it was deemed desirable however to keep the prewriting for the experimental tasks
to make them as similar as possible to compositions students had previously written for
their French at the university. The decision was therefore made not to give a prewriting
activity for either the Written Pretest or Posttest, but to keep them for the experimental
tasks. In this way, the conditions for the non-experimental tasks (no prewriting) were as
similar to each other as possible, as well as those for the two experimental tasks
(prewriting). Having all or no compositions, whether experimental or not, preceded by a
prewriting activity might be advisable for future studies of this kind. A delayed posttest would have also been desirable to track any long-term effects of online translation use on writing.

An unanticipated problem occurred with the translator training session. Shortly before the session was given, the location of the Babel Fish translator changed. Instead of being hosted at AltaVista as it had been for over ten years, the service was moved to Yahoo! Due to this, a last-minute decision was made to switch the translator used by participants in the study to http://freetranslation.com. While the change was noticed in time to update URLs in the instruments for the Written Pretest and Posttest, the translator training session still contained examples mostly from Babel Fish which were not updated. Although several sample sentences were translated the same as previously, apostrophes were no longer properly processed in the translation engine (and still aren’t as of writing), resulting sometimes in translations such as “I’ ; m affamé” for “I’m hungry” (Yahoo 2012). As a result, the translator training session involved several translations that were erroneous in ways not foreseen initially. It is possible that this affected student perceptions of, and subsequent use of, online translators. Although not a desirable outcome, since all students encountered these mistranslations, were exposed to other correctly translated sentences, used a different translator on the experimental tasks (which did process apostrophes correctly) to which they had also been exposed to in training, and had a list summarizing potential strengths and pitfalls of online translation (Appendix L), the training session was still considered appropriate.

Another limitation involves potential student access to unauthorized tools. While no participant mentioned doing so, there is a possibility that students, whose keystrokes
were not tracked, could have accessed an online translator or other sites for their compositions even when they were not allowed to do so. No cases were found that showed translator usage on tasks or among participants where no OT was allowed. Additionally, some students could have activated Microsoft Word’s autocorrect feature in French to assist in their writing. In future research, measures will be taken to limit the possibility of either of these eventualities occurring.

There were two mistakes in one of the instruments. On questions 3b and 4b of the Background Questionnaire, the instructions for questions soliciting answers for prior computer (including online translator) usage stated “You are not currently enrolled in French, but give your best estimate for your typical usage the last semester you had French.” There had been a trial run of the study on a limited number of volunteers who were no longer enrolled in French. When the questionnaire was adapted to the current study and given to terminal (non-continuing) students instead of students who had already finished their coursework, this wording was not changed. It is possible that the wording may have affected participant responses to these questions.

Lastly, all raters were native English speakers. It had been hoped that the results for native and non-native instructors could be compared, particularly for Overall Comprehensibility and judgments on if a translator had been used, since native-speaker perceptions on language use could be particularly insightful in addition to the ratings of non-native language instructors. Since one rater used half points on some of the ratings and the other did not complete ratings, their scores could not be used.
6.2 Future research

The findings of this study offer what is believed to be the first in-depth look at the effects of online translation on L2 student writers. While providing results that indicate the use of online translators did not negatively affect student performance – and in some cases allowed groups using a translator even to achieve higher mean scores than those not using a translator – much remains open for explanation in this nascent field of investigation.

Further research could confirm the positive effects of online translation on certain aspects of L2 written production, both among students trained in translator usage, and those that are not. Since several measures in the current study approached but did not reach statistical significance, a larger study building off of the design of the current one but with more participants might yield more conclusive and robust results. Furthermore, a longer study involving more experimental writing tasks could confirm whether students using online translation, and particularly those trained in its use, do in fact improve their scores with more time, as the results of the current study suggest.

The presence of delayed posttests with no access allowed to online translators could also gauge any long-term effects of OT training or use a month or longer after experimental tasks. In addition to self-reports, the use of other methods of determining the ways participants used the online translator (such as capture of keystrokes, recordings of participants’ activity either through a video camera or screen capture, think-aloud protocols) could be informative towards determining in what ways, at what parts of the writing process, and to what extent individual students use the online translator.
The inclusion of native speaker or faculty raters, who in many cases may have less day-to-day teaching experience with typical L2 student production at the novice-high to intermediate-low level, but who often have more experience in the language or grading/rating than graduate TAs, might give additional insight into how student writing is perceived in terms of overall comprehensibility and other measures.

While the current analysis looked at rater scores as the main indicator of participant performance on writing tasks, a detailed qualitative look at the differences between the compositions of students who did and did not use online translators would also be justified. Comparing the mistakes and output of students in the translator groups on compositions where OT was allowed, to those written by the control group on the same tasks, could shed some more light on why the former’s compositions were rated higher on some features, or why the control group never significantly outscored the translator groups on the experimental tasks. If participants are tracked in their translator usage, a rater’s impressions of specific words or phrases judged to be translated could be lined up with the student’s actual usage to see if rater suspicions proved to be correct not only on whether a translator was used on a given composition, but about specific portions of the text. Such an investigation might help explain cases where instructors think a translator is used when it is not, or vice versa, and train instructors in identifying translation use either to discourage OT use or help students use it more effectively.

Another analysis might contrast the features of pre- and post-tests for an individual student in online translation groups to the same features for that student on experimental tasks, tracking their use of vocabulary, spelling, etc., across tasks to see what changed when a translator was used. Additional, targeted questions on an exit
questionnaire or interview could expand on those from the current study to explore further student attitudes and impressions about the online translator’s effects overall, or specific characteristics of their writing.

The current study looked at terminal students at the novice-high / intermediate-low level. Future research on other populations of lower or higher proficiency is indicated. Students just beginning in French may not perform as well due to an inability to make informed judgments about OT output, while students majoring in French or at a higher level of fluency may be able to use online translators more effectively than those in the current study. Students more familiar with, or planning to pursue further coursework in, a language may be less motivated to enter entire chunks of language into the translator, as some in the current study reported doing, and instead use the translator as a tool to target specific items. If participants have a better knowledge of how the language works, they may be better at hypothesis checking and less likely to include in their writing any translations that are not proper from a lexical, grammatical, or semantic perspective, with OT freeing up more time to create and focus on communicating a message in the language.

Research could also be conducted on students who are taking a course about translation. Beginning translation students learning about making effective translations between languages, or more advanced students more well-versed in translation techniques and methodologies, may be able to make better use of an online translator as compared to students who are not knowledgable about translation.

Another promising area of research would be a comparison between students using an online translator and those using a free electronic dictionary (such as that found
A number of participants in the current study reported using the online translator like a dictionary to look up individual words, while others said they would have liked access to a dictionary in addition to or instead of online translator access. A comparison could be done among students trained in online dictionary usage and those trained in online translation to see the relative performance of each overall or on certain component scores as compared to a control group with access to neither tool.

While Free Translation was used in the current study, Google Translate appears to have become more well-known currently and, as noted before, is reported to have a remarkably large corpus from which to draw. With the increasing importance of mobile devices outside and inside the classroom, a number of other free translation solutions exist on a variety of mobile platforms. A study investigating L2 student performance on different online translators, or comparing student translation strategies on different devices, could produce enlightening results.

6.3 Implications and conclusions

Students have been using online translation for L2 writing since its inception in the late 1990s. Unlike most of the literature encountered, which has generally started with the stated or implied supposition that online translation is unuseful, inadequate, unethical, or even harmful to L2 students, this research has endeavored to adopt a fairly agnostic approach towards investigating this real and presumably prevalent phenomenon both inside and outside the foreign language classroom.

This study was an attempt to provide an ambitious first look at online translation, not necessarily from the point of view of measuring translator accuracy or trying to
determine how best to catch or discourage translator usage, but from the perspective of understanding how students use online translators and what effects, be they positive or negative, they might have on L2 written production.

The results of this study do not support the assumption that online translation use leads to poorly-written compositions. The control group, statistically similar to both translator groups at the onset of the study, never significantly outsored either translator group globally or on individual features that were rated. On the other hand, one or both translator groups, while not consistently outperforming participants who didn’t use the translator, did however achieve statistically higher scores on several measures. The results do however suggest that instructors can usually, but not always, tell if a translator has been used.

The simple fact that students who use an online translator do not necessarily perform worse, and may in some cases perform better, than those who write on their own without the help of a machine does not however suggest that online translation should be encouraged, allowed, or even tolerated in the L2 classroom. Several students who used a translator for the study reported themselves that they had doubts as to whether they had actually gotten anything of lasting value out of their translator-assisted writing.

It is important for those of us in the fields of SLA and pedagogy to make informed decisions about what policies we set to maximize our students’ learning experience. Technology can be the source of tools that foster students’ achievement, or may prevent or hinder their betterment. Students may get a higher grade if they copy-and-paste history papers from the Internet, but they will not have learned as much as if they had researched and written the paper themselves. On the other hand, in the not-too-distant
past, students were not allowed to use electronic calculators in the math classroom because it was believed they would hinder student progress. These devices are now commonplace and often required for students of a certain level, and on certain tasks, as a useful way to expand their knowledge and understanding of a field that, like language, seems foreign to many.

As technology and access to it continue to evolve, our approach to the issue of computer translation in student writing should encourage fair, ethical, and pedagogically-sound policies that foster student learning, while adapting to the reality that online translation is widely available and may actually help improve student production – particularly if they are using it to explore language (as a calculator might help a student explore complex mathematical concepts) instead of simply to accomplish the task at hand as quickly as possible (as a prewritten essay obtained online might allow a history student to do).

While this study is offered as one step in this direction, research and discussion is called for to investigate further how students are already using online translators; how providing information and instruction affects or discourages this usage; how informed or uninformed translation use impacts student writing in the short term; and how using an online translator, whether according to or in spite of instructional policies, may impact long-term language acquisition. These considerations should help determine what role, if any, online translation should have going forward in the L2 classroom.


Appendix A

Consent Form – Online Translation Study

Consent Form – Online Translation Study

You are invited to take part in a study concerning second-language writing. The research I am conducting is aimed at looking at the possible effects online translators may have on composition writing in French. Thanks to your participation, it is hoped that more can be understood about any potential benefits or drawbacks of advanced beginner and intermediate level language learners using online translators while writing. My name is Errol M. O’Neill, and I am a graduate student in the Department of French at the University of Illinois at Urbana-Champaign. I will be conducting this study under the supervision of Dr. Peter Golato, a professor in the Department of French at UIUC.

You have been chosen as a potential participant because you will have completed FR 103, FR 104, or FR 106 by the end of the semester; you are at least 18 years of age; and you do not intend to continue taking courses in French in the future. The study will take place on the campus of UIUC.

If you decide to participate in the study, you may be asked to complete a background questionnaire, take a series of pre-tests and post-tests, undergo a training session on the computer, and complete two tasks in which you will write a composition for each based on the topic provided. The total time for completing all parts of the study is expected to be approximately five and a half hours. You will complete about 4-5 hours of the study during class time, and 30 to 80 minutes outside of class time, depending on what group you are placed into. For each day you participate, you will receive 5 extra credit points. If you complete all 5 days, you will receive an additional 5 points, for 30 extra credit points total. Those who choose not to participate will be offered alternate extra credit opportunities in class, which will take approximately the same amount of time to complete as the tasks in the study. Each alternate activity will be worth 5 extra credit points, and non-participants who complete all 5 activities will receive an additional 5 points, for 30 extra credit points total. The total extra credit possible to both participants and non-participants is equivalent to 3% of total points for the course grade.

You are not obligated to respond to all the items on the background questionnaire or tasks you are given. Please note that in the event of partial completion or discontinued participation, you will only receive remuneration (in the form of extra credit) as indicated above for the parts of the study you have completed fully.

Your decision whether or not to participate in this study is completely voluntary and will not affect your grades (since there are alternative ways offered to earn the same amount of extra credit) or your relationship with UIUC or the Department of French. If you wish to discontinue your participation in the study, you may do so at any time by simply informing the researchers: myself (Errol M. O’Neill) or my supervisor (Dr. Peter Golato).

Only the researchers will have access to your identity and personal information for the sole purpose of data collection for this study. In order to protect your complete confidentiality, participants will be assigned a code number and code name which will be known to no one other than the researchers. In this way, any data used from the study will not be personally identifiable by anyone, including your instructor.

There is a small risk of mild, temporary fatigue from concentrating on the various tasks or from looking at the computer screen. Participants are free to take a pause periodically while working to alleviate any such discomfort. There is also a possibility that the use of online translators may have a positive or negative effect on the processes and strategies you use while writing in French. Since all participants are non-continuing students who are done with any composition assignments in their final semester of French, any positive or negative outcomes should not affect academic performance. It is hoped that your participation will contribute to knowledge about how the use of online translators might help or hinder students writing in a foreign language.

Questions about this research study should be directed to the researchers, Errol M. O’Neill, at eoneill@uiuc.edu or to my supervisor, Dr. Peter Golato, at pgolato@uiuc.edu. They may also be contacted by phone in the Department of French at (217) 333-2020. Questions about your rights as a research participant should be directed to the UIUC Institutional Review Board Office, which is located at 528 E. Green Street, Suite 203, Champaign, IL 61820 (MC-419) and which may be reached by calling (217) 333-2670, or writing to irb@uiuc.edu. You are welcome to call collect if you identify yourself as a research participant.

You will receive a copy of this consent form. Thank you very much for your participation.

YOUR FULL NAME (printed) __________________________________________

I have read and understood this form and volunteer to participate in this study.

YOUR SIGNATURE __________________________ DATE __________________

SIGNATURE OF RESEARCHER ______________________________________
Appendix B

Background Questionnaire

Self-report questionnaire

Please answer the following questions. Your responses are completely confidential.

• Background information

1) What is your native (first) language? ___________________________________________

2) In what country were you born? ___________________________________________

3) Have you lived in any country besides the one in question 2? If so, please specify where and length of residence.

4) What other language(s), including French, do you speak or have you learned or studied? (If none, please write NA)

<table>
<thead>
<tr>
<th>Language</th>
<th>When you started learning it</th>
<th>Place(s) you learned it</th>
<th>Fluency (be specific)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If you speak or have studied or learned additional languages, please continue your list on the back of this sheet.

5) What is/are your major(s)?

________________________________________________________________

6) What is/are your minor(s)?

________________________________________________________________

7) What is your (expected) graduation date?

________________________________________________________________

8) What motivated you to take French initially? (Check all that apply)

_____ To fulfill language requirement  _____ Interest in French language  _____ Interest in French culture

_____ Spanish (or other language) was full  _____ Interest in languages in general  _____ Interest in culture in general

_____ Other (please specify) __________________________________________

9) What future plans do you have concerning French? (Check all that apply)

_____ Take the next course in sequence (FR 104, FR 205, etc.)

_____ In my (future) profession

_____ Travel abroad

_____ I do not plan to use French in the future

_____ Other (please specify) __________________________________________
**Computer usage**

1) Please rate how frequently you used computers in **English** for the following purposes during the last semester you were enrolled.

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>To “surf” the Web for fun in English</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To play games in English</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To read (news, information) in English</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To check/write emails in English</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To look up information for class in English</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To look up individual words in a dictionary for English</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To type essays/papers for class in English</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To look up information for class in English</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To look up sample phrases/expressions in a search engine (Google, etc.) to incorporate in your writing in English</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To translate text to/from English (Babelfish, etc.)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Other (please specify)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

2) Please rate how frequently you use computers in **French** for the following purposes during the last semester you were enrolled.

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>To “surf” the Web for fun in French</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To play games in French</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To read (news, information) in French</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To check/write emails in French</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To look up information for class in French</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To look up individual words in a dictionary for French</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To type essays/papers for class for French</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To look up information for class in French</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To look up sample phrases/expressions in a search engine (Google, etc.) to incorporate in your writing in French</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>To translate text to/from French (Babelfish, etc.)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Other (please specify)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
3) Did you ever use a computer during classtime for French?

_____ Yes _____ No

• If Yes, please answer a through c below. If No, please go to the next page.

a) Rate how frequently:

_____ Once a month or less _____ Two to three times a month _____ Once a week or more

b) Indicate which tool(s) you used for French during classtime and how often you did so. You are not currently enrolled in French, but give your best estimate for your typical usage the last semester you had French.


\begin{tabular}{|c|c|c|c|c|c|}
\hline
\multicolumn{6}{|c|}{Never | Rarely | Sometimes | Often | Always} \\
\hline
\textbf{Word processor (Word, etc.)} & 1 & 2 & 3 & 4 & 5 \\
\hline
\textbf{Système D (or other, similar} & 1 & 2 & 3 & 4 & 5 \\
language writing software) & \\
\hline
\textbf{Cultural websites} & 1 & 2 & 3 & 4 & 5 \\
\hline
\textbf{Your course website} & 1 & 2 & 3 & 4 & 5 \\
\hline
\textbf{Online dictionary to look up} & 1 & 2 & 3 & 4 & 5 \\
individual words & \\
(WordReference, etc.) & \\
\hline
\textbf{Online translator (Babelfish, etc.)} & 1 & 2 & 3 & 4 & 5 \\
to translate text & \\
\hline
\textbf{Online grammar exercises} & 1 & 2 & 3 & 4 & 5 \\
\hline
\textbf{Other (please specify)} & 1 & 2 & 3 & 4 & 5 \\
\hline
\end{tabular}

c) Please list any other computer experience you have had, if any, involving French in the classroom.
4) Have you ever used a computer outside of class time for French?

_____ Yes _____ No

• If Yes, complete a through c below. If No, skip to the following page.

a) Rate how frequently:

_____ Once a month or less  _____ Two to three times a month  _____ Once a week or more

b) Please check which of the following you used for French outside of class time and circle how often it was used. You are not currently enrolled in French, but give your best estimate for your typical usage the last semester you had French.

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>___ Word processor (Word, etc.)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>___ Système D (or other, similar language writing software)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>___ Cultural websites</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>___ Your course website</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>___ Online dictionary to look up individual words (WordReference, etc.)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>___ Online translator (Babelfish, etc.) to translate text</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>___ Online grammar exercises</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>___ Other (please specify)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

C) Please list any other computer experience you have had, if any, involving French.
5) Have you ever used a computer for other foreign language classes (not French)?
   ____ Yes  ____ No
   • If Yes, please indicate your computer use below. If No, you have finished the questionnaire.

Please rate how frequently you use computers for languages other than English or French for the following purposes during the last semester you were enrolled.

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Never</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
</tr>
</thead>
<tbody>
<tr>
<td>____ To “surf” the Web for fun</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>____ To play games</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>____ To read (news, information)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>____ To check/write emails</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>____ To look up information for class</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>____ To look up individual words in a dictionary</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>____ To type essays/papers for class</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>____ To type/fill out other homework assignments</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>____ To look up sample phrases/expressions in a search engine (Google, etc.) to incorporate in your writing</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>____ To translate text to/from the language (Babelfish, etc.)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>____ Other (please specify)</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Thank you for your participation!
Appendix C

Reading Pretest (Excerpt)

Pretest: Reading

Directions for Part A: Each incomplete statement is followed by four suggested completions. Select the one that is best in each case.

1. En Normandie, les pommes mûres sont déjà toutes tombées par ------.
   (A) écrit
   (B) terre
   (C) avion
   (D) coeur

2. Dans les restaurants chics, on n’utilise pas de ------ en papier.
   (A) ceintures
   (B) serrures
   (C) serveuses
   (D) serviettes

3. Il vaudrait mieux que vous ------ à l’heure.
   (A) rentrez
   (B) finissez
   (C) soyez
   (D) partez

4. Regarde donc ce chêne là-bas. Celui-là, c’est vraiment un ------ arbre.
   (A) belle
   (B) beau
   (C) bel
   (D) beaux

   (A) au-dessus
   (B) par dessus
   (C) au-delà
   (D) par delà

6. . . . avoir entendu la nouvelle, Martine est rentrée chez elle en toute urgence.
   (A) En
   (B) Pour
   (C) Après
   (D) Car

7. Voilà toutes les robes qu’on nous a . . .
   (A) donné
   (B) donnée
   (C) donnés
   (D) données

Directions for Part B: In each of the following paragraphs, there are blanks indicating that words or phrases have been omitted. First, read through the entire paragraph. Then, for each blank, choose the completion that is most appropriate, given the context of the entire paragraph.

Questions 8–14

Quand j’étais enfant, j’adorais (8) animaux. Ma passion allait surtout aux chiens: je ne pouvais pas (9) voir un sans me (10) pour aller le caresser; je (11) un tas de livres (12) parlaient des chiens ou les mettaient en (13). Malheureusement mes parents ne (14) pas mon enthousiasme.

8. (A) l’
   (B) d’
   (C) les
   (D) des

9. (A) en
   (B) y
   (C) lui
   (D) le

10. (A) précipitant
    (B) précipitais
    (C) précipité
    (D) précipiter

11. (A) dévorais
    (B) mangeais
    (C) dénonçais
    (D) mordais

12. (A) que
    (B) qui
    (C) dont
    (D) lesquels

13. (A) scène
    (B) chemin
    (C) plateau
    (D) étage

14. (A) joignaient
    (B) manquaient
    (C) ravissaient
    (D) partageaient
Appendix D

Written Pretest

Written Pretest

Write a 3-paragraph composition on the following topic. Please follow the directions as stated below. You will have a total of 30 minutes to complete your composition.

Topic:

You are writing a letter to Corine, a student from Paris, France who is writing an article about summer vacation at American universities. (3 paragraphs, minimum of 4-5 sentences each)

1) Greet Corine, then give 2 reasons why summer vacation can be interesting/fun for American students and 2 reasons why it can be boring/annoying.

2) Talk about some interesting activities you did this past summer, either for work or with your family and friends.

3) Convince Corine why it is important for students to have a long summer vacation or why it would be better for there to be a shorter (or no) summer vacation.

Important Note: Remember, you are not allowed to use any outside help to assist you in your writing, either online or from anyone else (including the proctor). You should type your response in Microsoft Word and save the file to your desktop. Please raise your hand to let the proctor know once you have finished or if you experience technical problems.
Appendix E

Task One

You are writing a letter to Marie-Claire, a student from Avignon, France who is considering studying at your university in the United States. (3 paragraphs, minimum of 4-5 sentences each)

1) Greet Marie-Claire, then give 2 reasons why your university is better/more desirable than other schools and 2 reasons why it is worse/less desirable than other schools.

2) Talk about some interesting activities you did last semester while attending your university, either at school for class or in town with friends.

3) Convince Marie-Claire either why it is important for her to come to your university or why she might consider another school.

Important Note: Remember, you are to use the translator located at http://freetranslation.com to help you in writing your composition. You may not use any other programs, websites or other help to assist you in your writing, either online or from anyone else (including the proctor). You should type your response in Microsoft Word and save the file to your desktop. Please raise your hand to let the proctor know once you have finished or if you experience technical problems.
Task One

Write a 3-paragraph composition on the following topic. Please follow the directions as stated below. You will have a total of 30 minutes to complete your composition.

**Topic:**

You are writing a letter to Marie-Claire, a student from Avignon, France who is considering studying at your university in the United States. (3 paragraphs, minimum of 4-5 sentences each)

1) Greet Marie-Claire, then give 2 reasons why your university is better/more desirable than other schools and 2 reasons why it is worse/less desirable than other schools.

2) Talk about some interesting activities you did last semester while attending your university, either at school for class or in town with friends.

3) Convince Marie-Claire either why it is important for her to come to your university or why she might consider another school.

**Important Note:** Remember, you are not allowed to use any outside help to assist you in your writing, either online or from anyone else (including the proctor). You should type your response in Microsoft Word and save the file to your desktop. Please raise your hand to let the proctor know once you have finished or if you experience technical problems.
Appendix F

Task Two

Task Two

Write a 3-paragraph composition on the following topic. Please follow the directions as stated below. You will have a total of 30 minutes to complete your composition.

**Topic:**

*You are writing a letter to Jean-Pierre, a student from Lyon, France who is considering applying for a job in your field in the United States. (3 paragraphs, minimum of 4-5 sentences each)*

1) **Greet Jean-Pierre, then describe 2 things that are potentially good about jobs in your field and 2 things that are potentially negative about jobs in your field.**

2) **Talk about some important things you did, either at school or elsewhere, which helped you to prepare for having a job in this field.**

3) **Convince Jean-Pierre either why it is important for him to apply for your field in the United States or why he might consider another field (either related or completely different) instead.**

**Important Note:** Remember, you are to use the translator located at [http://freetranslation.com](http://freetranslation.com) to help you in writing your composition. You may not use any other programs, websites or other help to assist you in your writing, either online or from anyone else (including the proctor). You should type your response in Microsoft Word and save the file to your desktop. Please raise your hand to let the proctor know once you have finished or if you experience technical problems.
Task Two

Write a 3-paragraph composition on the following topic. Please follow the directions as stated below. You will have a total of **30 minutes** to complete your composition.

**Topic:**

*You are writing a letter to Jean-Pierre, a student from Lyon, France who is considering applying for a job in your field in the United States. (3 paragraphs, minimum of 4-5 sentences each)*

1) Greet Jean-Pierre, then describe 2 things that are potentially good about jobs in your field and 2 things that are potentially negative about jobs in your field.

2) Talk about some important things you did, either at school or elsewhere, which helped you to prepare for having a job in this field.

3) Convince Jean-Pierre either why it is important for him to apply for your field in the United States or why he might consider another field (either related or completely different) instead.

**Important Note:** Remember, you are **not** allowed to use any outside help to assist you in your writing, either online or from anyone else (including the proctor). You should type your response in Microsoft Word and save the file to your desktop. Please raise your hand to let the proctor know once you have finished or if you experience technical problems.
Appendix G

Posttest

Posttest

Write a 3-paragraph composition on the following topic. Please follow the directions as stated below. You will have a total of 30 minutes to complete your composition.

Topic:

You are writing a letter to Bernard, a student from Cannes, France who is writing an article about winter vacation at American universities. (3 paragraphs, minimum of 4-5 sentences each)

1) Greet Bernard, then give 2 reasons why winter vacation can be interesting/fun for American students and 2 reasons why it can be boring/annoying.

2) Talk about some interesting activities you did last winter, either for work or with your family and friends.

3) Convince Bernard why it is important for students to have a long winter vacation or why it would be better for there to be a shorter (or no) winter vacation.

Important Note: Remember, you are not allowed to use any outside help to assist you in your writing, either online or from anyone else (including the proctor). You should type your response in Microsoft Word and save the file to your desktop. Please raise your hand to let the proctor know once you have finished or if you experience technical problems.
Appendix H

Pre-writing Activity I

**Pre-writing Activity I**

Today you will write a composition based on a theme. This activity is designed to help you prepare for this task.

Please write your answers in French. Your answers will not be scored, but doing your best to answer them will help you prepare for the task you will do today.

I. *Votre passé et votre présent:* Répondez aux questions suivantes.

| a) Où est-ce que vous êtes allé(e) le week-end passé? Avec qui? Citez plusieurs endroits (*places*). [3 phrases] |
| b) Qu’est-ce que vous avez fait à ces endroit(s)? (étudier, dormir, etc….) Citez plusieurs activités. [3 phrases] |
| c) À votre avis, quelle classe est plus difficile, les maths ou le français? De toutes les classes, quelle classe est la moins difficile? Pourquoi? [3 phrases] |
| d) Quel genre de film est moins intéressant pour vous, la science-fiction ou les films d’amour? De tous les films, quel film aimez-vous le plus? Expliquez. [3 phrases] |
II. Vos amis
Pensez à un ami (homme) et une amie (femme) que vous connaissez. Ecrivez dans la boîte des qualités (bonnes ou mauvaises) pour décrire la personnalité, l’apparence physique, et les intérêts de chaque personne. Si vous préférez, vous pouvez inventer une personne et faire une description de cette personne inventée. Il n’est pas nécessaire d’écrire des phrases complètes.

<table>
<thead>
<tr>
<th>Personne</th>
<th>Ami:</th>
<th>Amie:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personalité</td>
<td>[2 qualités]</td>
<td></td>
</tr>
<tr>
<td>Apparence physique</td>
<td>[2 qualités]</td>
<td></td>
</tr>
<tr>
<td>Intérêts</td>
<td>[2 intérêts]</td>
<td></td>
</tr>
</tbody>
</table>

III. Vos vœux (wishes)
Maintenant, parlez des choses que vous désirez faire dans la vie. Ça peut être des choses concrètes (des objets, etc.) ou ça peut être des choses plus abstraites (l’amour, etc.). Voici des expressions que vous pouvez utiliser pour indiquer vos vœux pour vos amis: être important, vouloir, préférer, souhaiter, aimer, espérer, désirer, être nécessaire, être préférable, etc. [5-6 phrases. Si vous finissez quelques minutes en avance, continuez à écrire s’il vous plaît]
Appendix I

Pre-writing Activity II

Pre-writing Activity II

Today you will write a composition based on a theme. This activity is designed to help you prepare for this task.

Please write your answers in French. Your answers will not be scored, but doing your best to answer them will help you prepare for the task you will do today.

I. Votre passé et votre présent: Répondez aux questions suivantes.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>b) Avec qui avez-vous habité (comme enfant ou comme adulte)? Qu’est-ce que vous avez fait avec ces personnes? Citez plusieurs activités. [3 phrases]</td>
<td></td>
</tr>
<tr>
<td>c) À votre avis, quel est plus important, l’argent ou l’amour? Quelle est la chose la moins importante dans la vie? Expliquez. [3 phrases]</td>
<td></td>
</tr>
<tr>
<td>d) Quel moyen de transport est moins cher en général, la voiture ou le bus? De tous les moyens de transports, quel transport utilisez-vous le plus? Pourquoi? [3 phrases]</td>
<td></td>
</tr>
</tbody>
</table>
II. Vos amis
Pensez à deux amis (hommes) et deux amies (femmes) que vous connaissez. Ecrivez dans la boîte leur spécialisation, leur profession future (ou actuelle), et un passe-temps/activité favori. Si vous préférez, vous pouvez inventer une personne et donnez les détails de cette personne inventée. Il n’est pas nécessaire d’écrire des phrases complètes.

<table>
<thead>
<tr>
<th>Personne</th>
<th>Ami 1:</th>
<th>Ami 2:</th>
<th>Amie 1:</th>
<th>Amie 2:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Études</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Profession</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passe-temps</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

III. Vos vœux (wishes)
Maintenant, parlez de comment vous voyez le monde dans 10 ans. Qu’est-ce que vous désirez pour l’avenir (the future)? Quelles choses sont importantes pour vous personnellement? Pour la société en général? Voici des expressions que vous pouvez utiliser pour indiquer votre vision: être important, vouloir, préférer, souhaiter, aimer, espérer, désirer, être nécessaire, être préférable, etc. [5-6 phrases. Si vous finissez quelques minutes en avance, continuez à écrire s’il vous plaît]
Appendix J

Typing Accents for French Class

Typing Accents for French Class

Average PC*

é (e accent aigu) control + ', then e by itself
à (a accent grave) control + ', then a
c (e accent grave) control + ', then e
ô (o accent circonflexe) control + shift + 6, then o
â (a accent circonflexe) control + shift + 6, then a
î (i accent circonflexe) control + shift + 6, then i
ï (tréma) control + colon (:), then i
c (cédille) control + comma (,), then c

*These codes are guidelines for use with Microsoft Word that should work with versions of Office 97 and later. Codes should be similar for other programs in Windows.

Alternate accent codes on PC for French
Press ALT + the corresponding number at the same time.

<table>
<thead>
<tr>
<th>ALT+</th>
<th>Character</th>
<th>ALT+</th>
<th>Character</th>
</tr>
</thead>
<tbody>
<tr>
<td>0156</td>
<td>œ</td>
<td>0140</td>
<td>Ø</td>
</tr>
<tr>
<td>0224</td>
<td>â</td>
<td>0192</td>
<td>À</td>
</tr>
<tr>
<td>0226</td>
<td>à</td>
<td>0194</td>
<td>À</td>
</tr>
<tr>
<td>0228</td>
<td>ë</td>
<td>0196</td>
<td>À</td>
</tr>
<tr>
<td>0231</td>
<td>ç</td>
<td>0199</td>
<td>Ç</td>
</tr>
<tr>
<td>0232</td>
<td>è</td>
<td>0200</td>
<td>È</td>
</tr>
<tr>
<td>0233</td>
<td>é</td>
<td>0201</td>
<td>É</td>
</tr>
<tr>
<td>0234</td>
<td>ê</td>
<td>0202</td>
<td>È</td>
</tr>
<tr>
<td>0235</td>
<td>ñ</td>
<td>0203</td>
<td>Ê</td>
</tr>
<tr>
<td>0238</td>
<td>ï</td>
<td>0206</td>
<td>Ï</td>
</tr>
<tr>
<td>0239</td>
<td>î</td>
<td>0207</td>
<td>Î</td>
</tr>
<tr>
<td>0244</td>
<td>ò</td>
<td>0212</td>
<td>Ô</td>
</tr>
<tr>
<td>0246</td>
<td>ô</td>
<td>0214</td>
<td>Ô</td>
</tr>
<tr>
<td>0249</td>
<td>ù</td>
<td>0217</td>
<td>Ú</td>
</tr>
</tbody>
</table>

Adapted from:
http://www.class.uh.edu/lac/accentcode.html

Macintosh Computers

é (e accent aigu) option + e, then e by itself
à (a accent grave) option + ', then a
c (e accent grave) option + ', then e
ô (o accent circonflexe) option + i, then o
â (a accent circonflexe) option + i, then a
î (i accent circonflexe) option + i, then i by itself
ï (tréma) option + u, then i
c (cédille) option + c
œ (o dans l'e) option + q
Appendix K

Training Session – Lesson Outline

I) Online look-up tools for foreign languages
A) Online dictionaries
1) Advantages
   • Free of charge
   • Easy to use: enter a word and voilà!
   • Often several definitions, examples for each word or phrase to help fit the context
2) Disadvantages
   • Have to pick the right word/phrase
   • Not as complete as most full-sized dictionaries
   • Doesn’t do conjugation, agreement, etc. for you
B) Online translators
1) Advantages
   • Free of charge
   • Easy to use: enter word, phrase, sentence, paragraph...and voilà!
   • Nothing to choose from, only one word/phrase comes out
   • Attempts to do agreement, conjugation, etc.
2) Disadvantages
   • No options you can choose from: it decides for you
   • Errors in word choice, agreement
   • Unable to detect context, subject matter
   • Against the rules for many other departments/instructors

II) The good, the bad, and the unknown
A) Warnings about translator usage
1) Go to http://freetranslation.com/. Enter any text, hit “free translation”, then read the cautionary note.
2) Babel Fish: hidden in the Help section
   “As an automatic translator, Babel Fish works best when the text you wish to translate uses proper grammar. Slang, misspelled words, poorly placed punctuation and complex or lengthy sentences can all cause a page to be translated incorrectly. Expect Babel Fish to allow you to grasp the general intent of the original, not to produce a polished translation.” (Source: http://www.altavista.com/help/babelfish/babel_help)
3) Free online translators were created to help sell translation software (like a shareware version of software, to give an idea of how the paid version is like), but have grown immensely in popularity.
4) Target audience: prospective clients to buy professional translation software, or possibly people reading a text in a foreign language trying to get the general idea — not students trying to write in a foreign language.

B) Native-language judgments
1) Easier to understand and spot errors in your native language than in a foreign language.
2) Go to http://freetranslation.com
3) Set translator to “French to English”
4) Enter *Ma voisine est sympa, mais elle est trop curieuse.*
5) What is the result you get in English? Do you see any mistakes?

6) Sympa: very common word in French, but informal (sympa = sympathique).
   Apparently not in Freetranslation’s dictionary.
7) Would you be able to understand this sentence in English if you didn’t know a
   foreign language?
8) Go to http://babelfish.yahoo.com/
9) Set translator to “French to English”.
10) Enter *Ma voisine est sympa, mais elle est trop curieuse.*
11) What is the result you get in English? Do you see any mistakes?

12) Elle: can refer to people or things. Babel Fish didn’t figure out from context that
    “elle” refers to “voisine.”
13) Would you be able to understand this sentence in English? How is it different
    from Freetranslation’s? Are either of these translations satisfactory to you?
    Why or why not?

C) Foreign-language judgments
1) As a non-native speaker of French, you may be able to notice some mistakes or be
   able to understand even if there is a mistake, but harder than in your native language.
2) Go to http://freetranslation.com
3) Set translator to “English to French”
4) Enter *What a beautiful car!* 
5) What is the result you get in French? Do you see any mistakes?

6) “What” at the beginning of a sentence often indicates a question (“What are you
   doing?” could use “que” / “qu’est-ce que”), the translator didn’t get that this is an
   exclamation (context).
7) If your teacher got this sentence, do you think he/she would understand it?
8) Go to http://babelfish.yahoo.com/
9) Set translator to “English to French”
10) Enter *What a beautiful car!*
11) What is the result you get in French? Do you see any mistakes?

12) Babel Fish translated the right idea in a grammatically-correct, colloquial manner.
    Online translators are NOT always wrong.

D) Much remains unknown concerning online translators. They are generally owned by
private companies who do not reveal the exact workings of their software. The general public,
students, and even most teachers know little about how they actually work, but we can look at
examples to try to understand better.

III) Example translations from English to French.
    Now we’re going to look at some translations from English to French to see which ones are
    correct, could be correct depending on context, or are incorrect (at least partially).
A) Enter the following sentences into Babel Fish (http://babelfish.yahoo.com), selecting “English to French”.

1) *I don’t really like your house.*
   • Correct / Could be correct / Incorrect?
   • Word order in English and French
   • Corrected sentence/Other options?

2) *I really don’t like your house.*
   • Correct / Could be correct / Incorrect?
   • Word order in English and French
   • Corrected sentence/Other options?

3) *She wrote it down.*
   • Correct / Could be correct / Incorrect?
   • Phrasal verbs in English (e.g. to wake up, to sit down, etc.) don’t exist in French
   • Other options?

4) *She woke up.*
   • Correct / Could be correct / Incorrect?
   • Phrasal verbs in English (e.g. to wake up, to sit down, etc.) don’t exist in French
   • Corrected sentence/Other options?

5) *What’s up?*
   • Correct / Could be correct / Incorrect?
   • Familiar/colloquial expressions in English and French
   • Corrected sentence/Other options?

6) *What a drag!*  
   • Correct / Could be correct / Incorrect?
   • Familiar/colloquial expressions in English and French
   • Corrected sentence/Other options?

B) Now, let’s look at FreeTranslation. Enter the following sentences into FreeTranslation (http://freetranslation.com), selecting “English to French”.

1) *I gotta go.*
   • Correct / Could be correct / Incorrect?
   • “Incorrect” grammar in English
   • Corrected sentence/Other options?

2) *She don’t understand.*
   • Correct / Could be correct / Incorrect?
   • “Incorrect” grammar in English
   • Corrected sentence/Other options?

3) *She’s written a new book.*
   • Correct / Could be correct / Incorrect?
   • Contractions in English
   • Corrected sentence/Other options?
4) It’s written in English.
   • Correct / Could be correct / Incorrect?
   • Contractions in English
   • Corrected sentence/Other options?

5) I have lived here for 4 years.
   • Correct / Could be correct / Incorrect?
   • Different tenses in French and English
   • Corrected sentence/Other options?

6) When I was young, I would always go there.
   • Correct / Could be correct / Incorrect?
   • Different tenses in French and English
   • Corrected sentence/Other options?

7) I’m 21.
   • Correct / Could be correct / Incorrect?
   • Different verbs/expressions in French and English
   • Corrected sentence/Other options?

8) I’m never wrong.
   • Correct / Could be correct / Incorrect?
   • Different verbs/expressions in French and English
   • Corrected sentence/Other options?

9) a television
   • Correct / Could be correct / Incorrect?
   • Isolated words in French and English
   • Corrected sentence/Other options?

10) set
    • Correct / Could be correct / Incorrect?
    • Isolated words in French and English
    • Corrected sentence/Other options?

IV) Sheet: Online Translation—Some possible strengths and pitfalls.

V) Conclusion
   A) Online translation and this study
   B) Questions, comments

   Thank you again for your continued participation!!! 😊
Appendix L

Online Translation — Some possible strengths and weaknesses

A. What online translators often can do.
   1. Translate words that have one meaning in isolation
      (for example, the television=la télévision)
      
   2. Translate some words in context that have a small number of common, clear/distinct meanings
      (for example: I saw the man=J’ai vu l’homme)
      
   3. Translate some common, straightforward expressions correctly or with the right general idea
      (for example, What a bummer!=Quelle déception!)
      
   4. Make basic agreement between nouns and nearby adjectives and verbs
      (for example, She went=elle est allée We’re tired=Nous sommes fatigués)
      
   5. Translate tenses correctly when they’re the same in French and English
      (for example, I like the Internet=j’aime l’Internet)
      
   6. Give a general idea or “gist” of the meaning of a sentence or longer passage

B. What online translators often cannot do or what they have problems doing.
   1. Translate words that have a lot of meanings or uses
      (for example, set=ensemble in only some cases)
      
   2. Translate many specialized or technical words in context
      (for example: the escape key on a keyboard is not la clef d’évasion)
      
   3. Translate most expressions and phrasal verbs
      (for example: She wrote it down. is not Elle l’a écrit avalent. What a drag! is not Quelle drague!)
      
   4. Make correct agreement between nouns/adjectives and subjects/verbs, especially if they aren’t nearby
      (for example: She was, as you know, very intelligent. should have fem. adjective agreement in French)
      
   5. Translate tenses when the tenses are different between French and English
      (for example: I’ve been doing...since)
      
   6. Recognize misspelled or misused words
      (for example: their vs. they’re vs. there vs. thier)
      
   7. Deal with proper nouns (people’s names, place names, etc.) effectively
      (for example: Bill’s house ≠ la maison de la facture)
      
   8. Give an accurate, polished translation

C. Based on what we just saw, online translators have mixed results. Sometimes they give a completely correct translation, other times there are little mistakes here and there, and other times someone reading the translation wouldn’t be able to make heads or tails of what comes out. If you know what translators generally are good at doing or what they’re bad at doing, it can help you to use them more effectively in a given situation.
Appendix M

Cultural Lesson

Introduction to Translation and Interpretation

Adapted from Laura K. Lawless, http://french.about.com

What are they? What's the difference?

Introduction to Translation and Interpretation

Translation and interpretation are the ultimate jobs for people who love language. However, there are a lot of misunderstandings about these two fields, including the difference between them and what kind of skills and education they require. Both translation and interpretation (sometimes abbreviated as T + I) require superior language ability in at least two languages. That may seem like a given, but in fact there are many working translators whose language skills are not up to the task. You can usually recognize these unqualified translators by extremely low rates, and also by wild claims about being able to translate any language and subject.

Translation and interpretation also require the ability to accurately express information in the target language. Word for word translation is neither accurate nor desirable, and a good translator/interpreter knows how to express the source text or speech so that it sounds natural in the target language. The best translation is one that you don't realize is a translation, because it sounds just like it would if it had been written in that language to begin with.

Translators and interpreters nearly always work into their native language, because it's too easy for a non-native speaker to write or speak in a way that just doesn't sound quite right to native speakers. Using unqualified translators will leave you with poor-quality translations with mistakes ranging from poor grammar and awkward phrasing to nonsensical or inaccurate information.

And finally, translators and interpreters need to understand the cultures of both the source and target languages, in order to be able to adapt the language to the appropriate culture. In short, the simple fact of speaking two or more languages does not necessarily make a good translator or interpreter - there's a lot more to it.

Discussion questions. Please answer in your own words the following:

1) What are some ways the article mentions that you can tell a translator may be not qualified to do professional translation?

2) Is word-for-word translation normally good? Explain.

3) Is it better to translate from your native language into a foreign language, or from a foreign language into your native language? Why?

4) Besides the language itself, what else should a good translator have knowledge of?
Translation vs Interpretation

Translation and interpretation share the common goal of taking information that is available in one language and converting it to another. So what is the difference between translation and interpretation? It's very simple. Translation is written - it involves taking a written text (such as a book or an article) and translating it in writing into the target language. Interpretation is oral - it refers to listening to something spoken (a speech or phone conversation) and interpreting it orally into the target language. So you can see that the main difference is in how the information is presented - orally in interpretation and written in translation.

This might seem like a subtle distinction, but if you consider your own language skills, the odds are that your ability to read/write and listen/speak are not identical - you are probably more skilled at one pair or the other. So translators are excellent writers, while interpreters have superior oral communication skills. In addition, spoken language is quite different from written, which adds a further dimension to the distinction. Then there's the fact that translators work alone to produce a translation, while interpreters work with two or more people/groups to provide an interpretation on the spot during negotiations, seminars, phone conversations, etc.

Translation and Interpretation Terms

Source language
The language of the original message.

Target language
The language of the resulting translation or interpretation.

A language - Native language
Most people have one A language, although someone who was raised bilingual may have two A languages or an A and a B, depending on whether they are truly bilingual or just very fluent in the second language.

B language - Fluent language
Fluent here means near-native ability - understanding virtually all vocabulary, structure, dialects, cultural influence, etc. A certified translator or interpreter has at least one B language, unless he or she is bilingual with two A languages.

C language - Working language
Translators and interpreters may have one or more C languages - those which they understand well enough to translate or interpret from but not to.

Translators and interpreters should only work into the languages that they write/speak like a native or very close to it. It is very rare for anyone to have more than two target languages, although having several source languages is fairly common.

Discussion questions. Please answer in your own words the following:
1) What is the main difference between translation and interpretation?
2) What is the difference between source and target language? Between A, B, and C languages?
3) Is it more common to have a several source languages or several target languages? Why?
Types of Translation and Interpretation

General translation/interpretation is just what you think - the translation or interpretation of non-specific language that does not require any specialized vocabulary or knowledge. However, the best translators and interpreters read extensively in order to be up-to-date with current events and trends so that they are able to do their work to the best of their ability, having knowledge of what they might be asked to convert. In addition, good translators and interpreters make an effort to read about whatever topic they are currently working on. If a translator is asked to translate an article on organic farming, for example, he or she would be well served to read about organic farming in both languages.

Specialized translation or interpretation refers to domains which require at the very least that the person be extremely well read in the domain. Even better is training in the field (such as a college degree in the subject, or a specialized course in that type of translation or interpretation). Some common types of specialized translation and interpretation are

- financial translation and interpretation
- legal translation and interpretation
- literary translation
- medical translation and interpretation
- scientific translation and interpretation
- technical translation and interpretation

Types of Translation:

Machine translation
Also known as automatic translation, this is any translation that is done without human intervention, using software, hand-held translators, online translators such as Babelfish, etc. Machine translation is extremely limited in quality and usefulness.

Machine-assisted translation
Translation that is done with a machine translator and a human working together. For example, to translate "honey," the machine translator might give the options *le miel* and *chéri* so that the person could decide which one makes sense in the context. This is better than machine translation, and some argue that it is more effective than human-only translation.

Screen translation
Translation of movies and television programs, including subtitling (where the translation is typed along the bottom of the screen) and dubbing (where the voices of native speakers of the target language are heard in place of the original actors).

Sight translation
Document in the source language is explained orally in the target language. This task is performed by interpreters when an article in the source language is not provided with a translation (such as a memo handed out at a meeting).

Discussion questions. Please answer in your own words the following:
1) How do good general translators and interpreters stay current for their jobs?
2) What experience would one expect a specialized translator to have?
3) What is the difference between machine translation and machine-assisted translation? Screen translation and sight translation?
Should I Become A Translator?
Adapted from an article by Arthur Borges, July 10, 2005 (Source: http://www.translatorscafe.com/cafe/article39.htm)

Good translators may come in any size, shape and color: neurotic but idealistic language teachers, laid-off factory workers, ex-army commandos in from the cold, sharp disbarred lawyers, retired physicians and poetic alcoholics. We are all human. We all have both failings and the strengths that flourish from them: you are one of us to at least some extent. The Trinity of Translation is Language Proficiency, Specialization and Writing Skills:

Language Proficiency
Language proficiency is about feeling comfortable in two wordworlds: have you got it or are you prepared to get it by spending at least two years in a country that speaks your source language (i.e. the language you want to translate out of)? This is far more than speaking one language at home and another in your environment or pocketing a degree in a given second language. Not only do different folks use different words and grammars, they also think differently, their emotions respond to different stimuli and they have different moral value systems.

Languages are like flesh: they are subject to the law of birth, growth, old age and death. Immigrant parents lose contact with the evolution of their native tongue by living outside their native culture: you have to go back there for at least two years to get the hang of how folks think, act, feel, talk, gesticulate and generally operate. Or you have this degree in Tibetan from the University of Hintertupfingen that didn’t teach you any of the Chinese loanwords in the terminology of ATM maintenance training. The baseline is about becoming bi-cultural: learning the mindset behind the words.

Specialization
Specialization is about hands-on knowledge of anything from basketweaving and meatpacking to phased array radar technology and offshore oil services. In short, get really focused, ideally, by listing one foreign language, one target language and one area of expertise. Choose the area of expertise from your job history: a smart bilingual bookkeeper should be able to translate accounting, a smart factory worker should be able to do industrial user manuals. If you’ve got a degree, check with friends with family—you have readymade mentors all around you. Relatives and friends with expertise can also explain the fuzzy parts of any sentences you are translating, which is critical to minimizing mistranslations.

You WILL make mistakes: we only murder virtual doc files but by analogy, the more patients a surgeon has killed, the higher her/his skill levels.

Writing Skills
Writing skills are about how many different ways you can express the same thought in properly spelt and punctuated sentences. You have to enjoy writing. You have to enjoy playing with words and figuring out the meaning and intentions behind the words. Translation is a lonely job and it helps if you can get playful about the words you handle: can you stop and wonder why aircraft have no wing nuts? Or why they have cosmonauts, official state atheism and censorship while we have astronauts, separation of church and state, and news management?

Discussion questions. Please answer in your own words the following:
1) According to the article, do translators have to be language majors? Explain.
2) What is the most important way to build up your proficiency in another language? Why?
3) How can you choose a specialization for translation?
4) Is it bad to make mistakes when translating? Explain.
5) What should you like doing in order to be a good translator?
PC Skills
But there’s more. Documents arrive in different formats: how much of Microsoft Office can you exploit? Some terms are special to an industry or even a company: how deep are you willing to dig to find the right match in the target language? Being a bit of a neatness freak helps too: are you a perfectionist about layout, spelling, carriage returns, numbering, spelling, grammar and, just before you do the final SAVE, can you go through a document to remove all the extra spaces after the periods and all the extra line feeds on the last page? Translators are keyboard warriors: you need all 10 trigger fingers. Documents come in different formats based on different software: pirate or buy it and learn it.

Living With Your CAT
Also relevant are CAT tools, or computer-assisted translation software. You can still survive without it for the time being but if you have it, your chances of securing commissions improve nicely. That said, increasing numbers of agencies and clients expect you to buy that software and then use it against you to pay you a lower word rate.

Deadlines
Deadlines are sacred: miss one and your market value falls to that of a monolingual kindergartener holding a freshly opened box of brand new crayons. One marketing study reports that every unsatisfied customer talks to at least 11 potential customers; another study says the 11 is 15. Negotiate the longest deadline you can, but once you’ve committed to a date, honor it. Before you commit, make sure you have all the assurances you need on your end to get the job done.

Free Translations
‘Pro bono’ or ‘voluntary’ translations are a good way to start off. Contact your town hall and surf the Internet to identify local non-profit organizations that may need your services. Because they are local, you can get detailed feedback on the quality and presentation of your translations, secure job references and build up a network of professional contacts that way: all of them can connect you to paid work and they will out of gratitude sooner or later.

Perks
The perk is working for yourself with freedom to manage your time as you see fit, within the limits of your deadlines. The perk is being able to live anywhere in range of the Internet. You can live in Siberia and translate for Miami or work for Hintertupfingen and live in the Himalayas. The perk is inside insight to leading-edge research, business deals, technology and whatever—and a peek at how these things are interconnecting to shape daily reality and the world around you. The perk is developing a binocular vision of the world through a deeper understanding of contrasting mindsets and value systems—the more you understand them, the more aware you become of your own. The perk is sharpening and expanding your natural curiosity. The perk is doing something you love.

Discussion questions. Please answer in your own words the following:
1) Explain at least two aspects of being a “keyboard warrior”.
2) What is CAT? What advantage and disadvantage is given for using it?
3) Why are deadlines important?
5) Do any of the perks listed sound particularly good or bad to you? Why?
Appendix N

Self-report Exit Questionnaire

Please answer the following questions as instructed. Give as much detail as possible in your answers. Your responses are completely confidential, so please feel free to be completely honest. You can use the back or a separate sheet of paper if you need more room.

1. Do you feel the pre-writing activities before Task One and Task Two helped you in writing your composition?
   Circle: Yes / No

   If Yes, please explain in what way(s) you think they were helpful:

   If No, please explain what would have been more helpful for preparing for the composition:

2. Did you receive translator training during this study? Circle: Yes / No

   If Yes, do you feel that this training helped you in using a translator? Please explain why or why not.

   If No, do you feel you would have benefited from such training? Please explain why or why not.

Please proceed to the next page
3. Were you instructed to use an online translator during this study? Circle: **Yes / No**

   If **Yes**, answer the following two question sets. If **No**, skip to question 4.

   a) Did you actually use the translator for the writing tasks? If you **did** use the translator, please explain how you used the translator, how often you used it, and how much text you tended to translate (please mention if this varied). If you **didn't** use the translator, why you did not choose to do so? Please explain your reason(s).

   b) If you **did** use the translator, do you feel the use of the translator helped in your writing? If you feel using a translator was helpful please explain what specific areas this training helped with. If not, please explain what reason(s) you felt it was not helpful.

4. Did you attend a session with a cultural reading in English? Circle: **Yes / No**

   If **Yes**, did you feel this reading helped prepare you for writing in French? Explain why or why not.

   If **No**, please explain what type of reading or other activity would have been more helpful.
5. What other comments do you have about the study as a whole, or specific aspects of the study that you haven’t mentioned above?

Thank you once again for your participation in my study! 😊
Appendix O

Debriefing

Debriefing

By taking part in this study, you have participated in research into the effects of using online translators while writing in French. It is believed this is the first study of its kind, and your results should help lead to a better understanding of the possible advantages and disadvantages of these tools. Since the research is ongoing, it is extremely important for you not to reveal the details of this study to other possible participants (in particular, people who are currently taking 100-level French and might finish their French coursework after FR 103, 104, 106, 133, or 134).

Here is a more specific explanation of the various steps of the study you took part in.

Part One: Pretests. The first pretest, a 3-paragraph composition written without the aid of an online translator, was given to measure your writing level in French before any treatments (training, translator usage, etc.) were given. It was scored by both native and non-native speakers based on overall comprehensibility, content, spelling, syntax, grammar, and vocabulary. The second pretest, the sample CLEP exam, was administered to help gauge your overall competence in French before any treatments were given. Listening comprehension, reading comprehension, vocabulary, and structure were all assessed.

Part Two: Translator training. Some subjects were given training in translator usage. More specifically, possible strengths and weaknesses of online translators were explicitly presented, and several exercises were given for participants to try to determine through guided exercises and discussion with the researcher and other participants other potential strengths and weaknesses. The other participants who were not given this treatment acted as a “control”. By having some participants receive this training and other participants not receive the training, the results should help determine what effect, if any, the training sessions had on composition writing.

Part Three: Tasks One and Two. Both of these tasks involved writing a 3-paragraph composition. Each was scored by both native and non-native speakers based on overall comprehensibility, content, spelling, syntax, grammar, and vocabulary. Approximately half of the participants were permitted to use an online translator, while the rest were not allowed access to a translator. Comparing the results of those who wrote the composition with and without a translator, and with or without training, should help determine the possible effect of each on composition writing.

Part Four: Posttest. The posttest, consisting of a 3-paragraph composition written without the aid of an online translator, was given to measure your writing level in French at the end of the experiment. The posttest was scored by both native and non-native speakers based on overall comprehensibility, content, spelling, syntax, grammar, and vocabulary. By looking at the results of the posttest, it is hoped the effects of translator usage and training can be measured by comparing the writing of those who had access to online translators and training to the writing of those that had access to one or neither.

A number of variables were examined through this study to try to account for any significant differences in results among participants. The main variables considered included the ones mentioned above, namely: access to an online translator; access to training; writing overall comprehensibility, content, spelling, syntax, grammar, and vocabulary at the start of the study; listening comprehension, reading comprehension, vocabulary, and structure at the start of the study. Other variables have also been considered for possible effects, including various personal background information such as previous experience with online translators and/or other computer technology, experience and motivation in studying French, and experience in using other languages.

If you would like to be informed by the researcher once the results to this study have been published, please inform him at coneill@uiuc.edu or autrefois@mac.com. To protect your privacy, no information on individual performance will be published or given out to anyone, but by reading the published results you can find out the overall results to this study. Thank you once again for your participation! 😊
Appendix P

Scoring Criteria for Compositions

Scoring Criteria for Compositions

1) After reading each composition, please evaluate it based on the following categories by noting on the composition the abbreviation for each category and the corresponding score which you believe is appropriate for that category.

• Overall comprehensibility (OC) — Could this writing be understood by a native French speaker with little or no knowledge of English?
  5 Yes, with no difficulty
  4 Yes, with a little difficulty
  3 Yes, with moderate difficulty
  2 Yes, with great difficulty
  1 No, it could not be understood

• Content (CO) — Are the ideas used sufficiently creative and appropriate to the subtopics outlined in the task?
  5 Yes, the writer creatively and appropriately addressed the theme by covering all 3 subtopics
  4 Yes, the writer creatively and appropriately addressed the theme by covering 2 of the 3 subtopics
  3 The writer creatively and appropriately addressed the theme by covering 1 of the 3 subtopics
  2 No, the writer did not creatively and effectively address any of the subtopics, but addressed the theme of the composition
  1 No, the writer did not creatively and effectively address any of the subtopics, nor did s/he address the theme of the composition

• Spelling / Accents (SA) — Are the letters and diactrical marks in words written as would be expected in standard or colloquial written French? (Note: This does not include conjugation or other grammatical usage.)
  5 Yes, all or very nearly all of the words used are written perfectly
  4 Yes, the vast majority of the words used are written correctly
  3 About half of the words used are written correctly
  2 No, the vast majority of the words used are written incorrectly
  1 No, all or very nearly all of the words contain mistakes

• Syntax (SY) — Is the word order used appropriate to standard or colloquial written French?
  5 Yes, word order throughout is appropriate
  4 Yes, word order is appropriate in the vast majority of cases
  3 Word order is appropriate about half of the time
  2 No, word order is inappropriate in the vast majority of cases
  1 No, word order throughout is inappropriate

• Remaining grammar (RG) — Is remaining grammar appropriate in relation to standard or colloquial written French? (including: tense/mood, subject-verb agreement, number/gender agreements, articles, negation, etc.)
  5 Yes, the remaining grammar throughout is appropriate
  4 Yes, the remaining grammar is appropriate in the vast majority of cases
  3 The remaining grammar is appropriate about half of the time
  2 No, the remaining grammar is inappropriate in the vast majority of cases
  1 No, the remaining grammar throughout is inappropriate

• Vocabulary (VO) — Are the words that have been chosen accurate and effective in standard or colloquial written French? (This does not include spelling, accents, or grammatical usage.)
  5 Yes, all or very nearly all of the words used are accurate and effective
  4 Yes, the vast majority of the words used are accurate and effective
  3 About half of the words used are accurate and effective
  2 No, the vast majority of the words used are not accurate and effective
  1 No, all or very nearly all of the words are not accurate and effective

2) After noting the score for each category, add up the scores and note the Total score (out of 30).

3) Do you believe this composition was written with the aid of a translator? Write "Yes" or "No" on the composition.

4) Please briefly explain below the composition (or on the back if you prefer / if more space is needed) your response to number 3 by citing a specific reason(s) or example(s) of errors present or lacking from the paper. Please be sure you're clear about what you mean.