

HARNESSING COLLECTIVE INTELLIGENCE FOR TRANSLATION: AN ASSESSMENT OF CROWDSOURCING AS A MEANS OF BRIDGING THE CANADIAN LINGUISTIC DIGITAL DIVIDE

Steven O'Brien

Supervisor: **Dr. Ajax Persaud**

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN ELECTRONIC BUSINESS TECHNOLOGIES



uOttawa

Telfer School of Management
Faculty of Graduate and Postdoctoral Studies
University of Ottawa

© Steven O'Brien, Ottawa, Canada, 2011

Table of Contents

List of Tables	vii
List of Figures	x
List of Equations.....	xii
Acknowledgments.....	xiii
Abstract.....	xiv
Chapter 1: Introduction	1
1.1 Rationale and Significance	3
1.2 Anticipated Contribution	6
Chapter 2: Theoretical Framework.....	7
2.1 Collaborative Online Language Transfer.....	8
2.2 Performance Criteria.....	9
2.3 Lightweight Peer Production	14
2.4 The Landscape of Document Engineering	18
2.5 Problem Statement.....	19
Chapter 3: Phase I (Translation Creation) - Methodology.....	21
3.1 Overview	21
3.2 Source Text.....	23
3.3 Qualitative Translation Goals (<i>Translation Brief</i>)	25
3.4 Amazon Mechanical Turk.....	25
3.5 Confluence Software.....	27
3.6 Paragraph Translation Process	27
3.7 Paragraph Evaluation Process.....	27
3.8 Revision Creation Process.....	28
3.9 Revision Evaluation Process.....	29
Chapter 4: Phase I (Translation Creation) - Analysis.....	30
4.1 Key Performance Indicators.....	30
4.2 Participation	30
4.3 Non-Participation.....	31
4.4 Time	32

4.5 Cost	32
4.6 Evaluation Selection.....	32
4.7 Diversity	33
4.8 Approximate Translation Quality.....	36
Chapter 5: Phase II (Translation Evaluation) - Methodology.....	37
5.1 Overall Translation Quality	37
5.2 Task Repetition	38
5.3 Knowledge of Translation Methods.....	38
5.4 General Translation Errors.....	38
5.5 Overt Errors.....	39
5.6 Usage Errors	39
5.7 Author Errors	40
5.8 Contextual Translation Errors.....	40
5.9 Translation Diversity	40
5.10 Machine Translation	41
5.11 Professional Translation.....	41
5.12 Crowdsourced Translation	41
5.13 Crowdsourced Translation Acceptance	41
5.14 Research Hypotheses.....	42
5.15 Hypothesis Testing.....	45
5.16 Partial Least Squares – Path Modeling with Smart PLS 2.0	45
5.17 Structural Equation Modeling using PLS-PM	46
5.18 Study Participants	48
5.19 Instrument and Measurement	51
Chapter 6: Phase II (Translation Evaluation) - Analysis	53
6.1 Overview	53
6.2 Hypothesis (H ₁)	55
6.3 Hypothesis (H ₂)	58
6.4 Hypothesis (H ₃)	61

6.5 Hypothesis (H ₄)	64
6.6 Hypothesis (H ₅)	66
6.7 Hypothesis (H ₆)	69
6.8 Hypothesis (H ₇)	73
6.9 Hypothesis (H ₈)	76
6.9 Crowdsourced Translation Acceptance	79
6.10 Rank Order Positively Influences Overall Translation Quality	84
6.11 Multivariate Model of Crowdsourced Translation	86
Chapter 7: Discussion.....	94
7.1 Limitations and Future Research	99
7.2 Conclusion.....	100
References	102
Appendix A – Confluence Software	109
A.1 Major Components.....	109
A.2 Confluence Host	109
A.3 Confluence Client.....	110
A.4 Types of Raw Data Generated by Confluence	111
A.5 Twitter Integration	112
A.6 Amazon Mechanical Turk Screens.....	113
A.7 Template Files.....	119
Appendix B – Phase I (Translation Creation) Data.....	125
B.1 Professional Translation (PT)	126
B.2 Crowdsourced Translation – Language Professionals – Task Repetition Value Two (CSTLP2)	127
B.3 Crowdsourced Translation – Language Professionals – Task Repetition Value Four (CSTLP4)	128
B.4 Crowdsourced Translation – Non-Language Professionals – Task Repetition Value Two (CSTNLP2).....	129
B.5 Crowdsourced Translation – Non-Language Professionals – Task Repetition Value Four (CSTNLP4).....	130

B.6 Machine Translation (MT)	131
B.7 CSTLP2 Rough Draft	132
B.8 CSTLP2 Discarded Revision #1	133
B.9 CSTLP4 Rough Draft	134
B.10 CSTLP4 Discarded Revision #1	135
B.11 CSTLP4 Discarded Revision #2	136
B.12 CSTLP4 Discarded Revision #3	137
B.13 CSTNLP2 Rough Draft.....	138
B.14 CSTNLP2 Discarded Revision #1.....	139
B.15 CSTNLP4 Rough Draft.....	140
B.16 CSTNLP4 Discarded Revision #1.....	141
B.17 CSTNLP4 Discarded Revision #2.....	142
B.18 CSTNLP4 Discarded Revision #3.....	143
B.19 Discarded Professional Translation #1	144
B.20 Discarded Professional Translation #2	145
B.21 Discarded Machine Translation #1	146
B.22 Discarded Machine Translation #2	147
Appendix C – Phase II (Translation Evaluation) Data.....	148
C.1 Selfish Selection and Overall Translation Quality.....	149
C.2 Unselfish Selection and Overall Translation Quality	150
C.3 Disinterested Selection and Overall Translation Quality.....	152
C.4 Data Tables for H ₂	153
C.5 Data Tables for H ₄	156
C.6 Data Tables for H ₅	159
C.7 Crowdsourced Translation Acceptance Data.....	162
C.8 Survey Instructions – Part I.....	163
C.9 Source Text Evaluation Survey.....	164
C.10 Target Text Evaluation Survey	165
C.11 Survey Instructions – Part II.....	166

C.12 Supplemental Evaluation Survey	167
C.13 Screening Letter	168
C.14 Statement of Informed Consent	169

List of Tables

Table 1 - Estimated Source Text Difficulty	25
Table 2 - Non-Participation	31
Table 3 - Reasons for Declining to Participate (by experiment).....	31
Table 4 - Approximate Time Measures	32
Table 5 - Cost Measurements.....	32
Table 6 - Approximate Evaluation Selection	33
Table 7 - Simpson's Index Measures.....	35
Table 8 - Simpson's Reciprocal Index Measures.....	35
Table 9 - Rankings of Approximate Translation Quality.....	36
Table 10 - Acceptable Thresholds for Interpreting PLS-PM	47
Table 11 - Phase II Participant Education	49
Table 12 - Phase II Professional Translation Work Experience	49
Table 13 - H _{1A} (Structural and Measurement Model).....	55
Table 14 - H _{1A} (Bootstrapping Results).....	55
Table 15 - H _{1A} (PLS Algorithm Results).....	56
Table 16 - H _{2A} (Measurement Indicators)	58
Table 17 - Changes in OTQ after Knowledge of Translation Methods (KTM).....	59
Table 18 - H _{2A} (Aggregated Changes in Overall Translation Quality after KTM)	60
Table 19 - H _{3A} (Structural and Measurement Model)	61
Table 20 - H _{3A} (Bootstrapping Results)	61
Table 21 - H _{3A} (PLS Algorithm Results)	62
Table 22 - H _{4A} (Measurement Indicators)	64
Table 23 - H _{4A} (Aggregated Overall Translation Quality)	64
Table 24 - H ₄ (MT vs. CST Aggregated Rank)	65
Table 25 - H _{5A} (Measurement Indicators)	66
Table 26 - H _{5A} (Overall Translation Quality by Experiment Type)	66
Table 27 - H _{5A} (Aggregated Overall Translation Quality).....	67
Table 28 - H ₅ (PT vs. CST Aggregated Rank)	68
Table 29 - H _{6A} (Structural and Measurement Model)	69
Table 30 - H _{6A} (Bootstrapping Results)	70

Table 31 – H _{6A} (PLS Algorithm Results)	70
Table 32 – H _{6D} (PLS Algorithm Results)	71
Table 33 – H _{6E} (PLS Algorithm Results)	72
Table 34 – H _{7A} (Structural and Measurement Model)	73
Table 35 – H _{7A} (Bootstrapping Results)	74
Table 36 – H _{7A} (PLS Algorithm Results)	74
Table 37 – H _{8A} (Structural and Measurement Model)	76
Table 38 – H _{8A} (Bootstrapping Results)	76
Table 39 – H _{8A} (PLS Algorithm Results)	77
Table 40 – Crowdsourced Translation Acceptance (Measurement Indicators)	79
Table 41 - Crowdsourced Translation Acceptance by Participant Type	83
Table 42 – H _{9A} (Structural and Measurement Model)	84
Table 43 – H _{9A} (Bootstrapping Results)	84
Table 44 – H _{9A} (PLS Algorithm Results)	85
Table 45 – Multivariate Model, Version #1 (Bootstrapping Results)	87
Table 46 – Multivariate Model, Version #2 (Bootstrapping Results)	88
Table 47 – Multivariate Model, Version #2 (PLS Algorithm Results)	90
Table 48 – Multivariate Model, Version #3 (Bootstrapping Results)	91
Table 49 – Multivariate Model, Version #3 (PLS Algorithm Results)	92
Table 50 – Selfish Selection Hypothesis (Structural and Measurement Model)	149
Table 51 – Selfish Selection Hypothesis (Bootstrapping Results)	149
Table 52 – Unselfish Selection Hypothesis (Structural and Measurement Model)	150
Table 53 – Unselfish Selection Hypothesis (Bootstrapping Results)	150
Table 54 – Unselfish Selection Hypothesis (PLS Algorithm Results)	151
Table 55 – Disinterested Selection Hypothesis (Structural and Measurement Model)	152
Table 56 – Disinterested Selection Hypothesis (Bootstrapping Results)	152
Table 57 - H ₂ (ACQ4)	153
Table 58 - H ₂ (OTQ14)	153
Table 59 - H ₂ (OTQ24)	154
Table 60 - H ₂ (OTQ34)	154
Table 61 - H ₂ (OTQ44)	155
Table 62 - H ₂ (Shift for Overall Translation Quality)	155

Table 63 - H _{4A} (ACQ1).....	156
Table 64 - H _{4A} (OTQ11)	156
Table 65 - H _{4A} (OTQ21)	156
Table 66 - H _{4A} (OTQ31)	157
Table 67 - H _{4A} (OTQ41)	157
Table 68 - H _{4A} (Overall Translation Quality – Five Tiers)	157
Table 69 - H ₄ (MT vs. CST Rankings in Six Tiers).....	158
Table 70 - H ₄ (MT vs. CST Rankings in Two Tiers)	158
Table 71 - H ₄ (MT vs. CST Aggregated Rank in Six Tiers)	158
Table 72 – H _{5A} (ACQ1)	159
Table 73 – H _{5A} (OTQ11)	159
Table 74 – H _{5A} (OTQ21).....	159
Table 75 – H _{5A} (OTQ31).....	160
Table 76 – H _{5A} (OTQ41).....	160
Table 77 – H _{5A} (Aggregated Overall Translation Quality - Five Tiers)	160
Table 78 - H _{5A} (PT vs. CST Rankings in Six Tiers)	161
Table 79 – H _{5A} (PT vs. CST Rankings in Two Tiers)	161
Table 80 - H _{5A} (PT vs. CST Aggregated Rank in Six Tiers).....	161
Table 81 - Crowdsourced Translation Acceptance - Five Tiers.....	162
Table 82 - Crowdsourced Translation Acceptance - Three Tiers	162

List of Figures

Figure 1 - Black Box Model	9
Figure 2 - High-Level Process Model	10
Figure 3 - Lightweight Peer Production Model.....	17
Figure 4 - Collaborative Online Language Transfer System Model	19
Figure 5 - Source Text	23
Figure 6 - Participant Qualification Test.....	26
Figure 7 - Structural Equation Modeling using PLS-PM.....	46
Figure 8 - Phase II Participant Workflow.....	50
Figure 9 - H _{1A} (PLS Algorithm Results)	56
Figure 10 - H _{3A} (PLS Algorithm Results)	62
Figure 11 - H _{6A} (PLS Algorithm Results)	70
Figure 12 - H _{6D} (PLS Algorithm Results)	71
Figure 13 - H _{6E} (PLS Algorithm Results)	72
Figure 14 - H _{7A} (PLS Algorithm Results)	74
Figure 15 - H _{8A} (PLS Algorithm Results)	77
Figure 16 - CTA1 Chart	79
Figure 17 - CTA2 Chart.....	80
Figure 18 - CTA3 Chart.....	81
Figure 19 - CTA4 Chart.....	81
Figure 20 - CTA5 Chart.....	82
Figure 21 - H _{9A} (PLS Algorithm Results)	85
Figure 22 - Multivariate Model of Crowdsourced Translation, Version #1	86
Figure 23 - Multivariate Model of Crowdsourced Translation, Version #2	88
Figure 24 - Multivariate Model, Version #2 (PLS Algorithm Results)	89
Figure 25 - Multivariate Model of Crowdsourced Translation, Version #3	91
Figure 26 - Multivariate Model, Version #3 (PLS Algorithm Results)	92
Figure 27 - Scalable Lightweight Peer Production with Post-Editing	96
Figure 28 - Confluence Data Files	111
Figure 29 - Sample Twitter Feed	112
Figure 30 - Pre-Qualification Screen	113
Figure 31 - Participant Qualification Acceptance	114

Figure 32 - Confluence Search Example	114
Figure 33 - Paragraph Translation Task Example	115
Figure 34 - Paragraph Evaluation Task Example (abbreviated)	116
Figure 35 - Revision Creation Task Example (abbreviated)	117
Figure 36 - Revision Evaluation Task Example (abbreviated).....	118
Figure 37 - Hit Type Template	119
Figure 38 - Answer Key Template.....	119
Figure 39 - Invitation Template.....	119
Figure 40 - Paragraph Translation Task Template	120
Figure 41 - Paragraph Evaluation Task Template (parts 1 and 2).....	121
Figure 42 - Paragraph Evaluation Task Template (part 3)	122
Figure 43 - Revision Creation Task Template	123
Figure 44 - Revision Evaluation Task Template.....	124
Figure 45 - Unselfish Selection Hypothesis (PLS Algorithm Results).....	151

List of Equations

Equation 1 - Simpson's Index	34
Equation 2 - Simpson's Reciprocal Index	34

Acknowledgments

"There are those who feel the urge to solve a problem, those for whom a problem becomes real, like a disorder which they have to get out of their system." –Karl Popper, The Nature of Philosophical Problems and Their Roots in Science, 1952.

I would like to thank my thesis supervisor, Dr. Ajax Persaud, for his invaluable contribution, support, and incredible patience; this research could not have been conducted without his wisdom and guidance. I am grateful for the opportunity to have been able to work with him on this project. I would like to thank Dr. Ferrand for establishing the Electronic Business Technologies program at the University of Ottawa. I would like to thank the faculty at the School of Translation and Interpretation for sharing their enthusiasm about the project with me.

To my parents, I would like to thank them for instilling in me the importance of setting goals in life, of believing in myself, and the utility of imagination. I would like to thank my children for letting me study and work on this research during so many of the days while they were little. Lastly, I would like to thank my loving wife, Christine, for all of her support, patience, and endless hours of listening to me ramble on about a topic that excites me.

I would like to acknowledge the research support provided by my thesis supervisor, Dr. Ajax Persaud. I would also like to thank the Research Ethics Board of the University of Ottawa and thesis examination board for their insightful comments and contributions.

Abstract

Web 2.0 technologies have been catalysts for the emergence of online collaboration and peer production. These disruptive technologies are among the drivers for business model changes in many sectors, including the language service industry. This thesis describes a framework in which the techniques of Lightweight Peer Production can be joined with the principles of Document Engineering for e-Business to enable a Collaborative Online Language Transfer System for crowdsourced translation. It is within the context of the Canadian Linguistic Digital Divide that this research is anchored.

This study attempts to shed light on the efficacy of crowdsourcing as a means of translating web content in Canada. Within, we seek to explore and understand if a model can be created that can estimate the effectiveness of crowdsourced translation as a means of bridging the Canadian Linguistic Digital Divide. To test our hypotheses and models, we use structural equation modeling techniques coupled with confidence intervals for comparing experimental crowdsourced translation to both professional and machine translation baselines. Furthermore, we explore a variety of factors which influence the quality of the experimental translations, how those translations performed in the context of their source text, and the ways in which the views of the quality of the experimental translations were measured before and after participants were made aware of how the experimental translations were created.

This study considers the work of both Language Professionals (translators) and Non-Language Professionals (bilinguals). We find a wide acceptance of the idea of crowdsourced translation to be performed by Language Professionals as a means of bridging the Canadian Linguistic Digital Divide. Furthermore, our study shows that this method of crowdsourced translation performed relatively well with respect to translation quality when compared against traditional translation methods and online machine translation methods. We find that Non-Language Professional translation evaluators are more accepting of crowdsourced translation than are their Language Professional counterparts.

In summary, this thesis will describe the theoretical framework and literature used to create a Collaborative Online Language Transfer System. Largely experimental and exploratory in nature, the thesis is divided into two distinct phases: Phase I (Translation Creation) and Phase II (Translation Evaluation). Within each phase, the methodology and analyses are described in detail. This thesis concludes with a discussion of the study's implications, limitations, and future research directions.

Chapter 1: Introduction

On the World Wide Web there exists a digital divide between Francophone and Anglophone Canadians (Clavet, 2002; OCOL, 2005; Lord, 2008). It is a division of human communication and expression that inhibits the people of Canada from exchanging their ideas and cultures with each other. In 2005, Canada's Commissioner of Official Languages published a report outlining the Federal Government's strategies for bridging this digital divide (OCOL, 2005). The first strategic goal described in this report is to "*Promote the provision of French Internet tools and content*" (OCOL, 2005, p. 1). Although much progress has been made through various endeavors, the Commissioner of Official Languages indicates that there are many challenges still remaining. Among the remaining challenges are the following:

"It is important to continue investing in the digitization of French content, as the digital gap between Francophones and Anglophones is still wide; it has also become urgent to start a worldwide project to digitize cultural heritage in the French language." (OCOL, 2005, p. 3)

"Research and development activities aimed at designing technolinguistic tools need to be accelerated, as these tools will contribute to increasing the productivity of those working in the language industry and allow all Canadians to fully profit from the resources found on the Web." (OCOL, 2005, p. 3)

The nature of the Internet and its technologies are rapidly changing. The emergence of Web 2.0 and its associated technologies is now ubiquitous as electronic business technologies, and therefore, it is necessary to consider Web 2.0 when considering solutions for problems related to the World Wide Web. This thesis proposes that the technique of crowdsourcing could be one approach to bridging the Canadian linguistic digital divide. However, for crowdsourcing to be considered or adopted, its effectiveness must first be demonstrated. In order to understand if crowdsourcing should be considered as one possible solution to the problem of bridging the digital divide between Francophone and Anglophone Canadians on the World Wide Web, this thesis proposes to establish a model for estimating the effectiveness of crowdsourced translations.

What is the operational definition of translation? "... [T]ranslation may be defined as the replacement of a text in the source language by a semantically and pragmatically

equivalent text in the target language" (House, 1977, p. 103). Translation therefore may be viewed as one of the primary means by which English language content on the World Wide Web may be transferred into French-language content on the World Wide Web. Some common approaches to translation involve the use of a professional translator or the use of technolinguistic tools, such as machine translation systems. It is already well established that machine translation methods usually provide inferior results to those provided by professional translators (Bowker, 2008). The magnitude of this digital divide contributes to the shortage of professional translators in Canada who are capable of bridging this digital divide (OCOL, 2005).

Crowdsourcing is defined as "the act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call, generally over the Internet." (Howe, 2008, para. 3). A specific and recent example of crowdsourcing and translation is IBM's n.Fluent project, which was started as an internal research project in 2008 (IBM, 2010). "Through a series of Crowdsourcing strategies and events, n.Fluent has successfully engaged and nurtured a vibrant and active world-wide pool of volunteer translators who are dedicated to innovation that truly matters!" (IBM, 2010).

This thesis hypothesizes that the techniques of crowdsourced translation may be capable of creating translation products of acceptable quality when compared against the lower quality results produced by machine translation methods and in some cases when compared against the expected higher quality work of professional translators (Snow, O'Connor, Jurafsky, & Ng, 2008). It is the purpose of this thesis to investigate the validity of using crowdsourced translations as a possible means by which the Canadian linguistic digital divide may be partially bridged. If it can be shown that this form of translation can be capable of producing translations of acceptable quality in a reasonable time at a reasonable cost, it may be useful for companies, individuals and government agencies to consider this approach as one means of bridging the digital divide that exists for Canadians online.

It is worth noting, however, that crowdsourced translation is not likely to be appropriate for all types of texts or translation services. For example, the translation of legal documents poses challenges of difficult terminology, which in some cases span incongruous legal systems and cultures (Prieto Ramos, 2002). Translators of legal terminology, in some cases, are obligated to also be practitioners of comparative law

(Kocbek, 2006). Generally, this type of knowledge and expertise is beyond the capabilities of most bilingual Canadians and therefore may be impractical for crowdsourced translation.

1.1 Rationale and Significance

The primary rationale for this study resides in the fact that there is a shortage of professional translators in Canada (OCOL, 2005). This shortage of professional translators in Canada is “a major obstacle in the translation and digitization of French content” (OCOL, 2005, p. 3). In some cases, the lack of professional translators in Canada has contributed to this digital divide (OCOL, 2005, p. 58):

“Certain federal institutions have given up on posting some scientific, technical or historical material on their Web sites because that information is available in one language only (usually English) and because it would be too expensive to have it translated professionally. For some, the problem goes beyond mere economics – it is difficult to obtain adequate translation for a technical document that uses terminology that few translators know.”

The second rationale for this study is that there is simply too much content to be professionally translated in an acceptable timeframe. The problem of too much web content is exacerbated by its rapidly changing nature. The task of translating this constantly changing content can be a significant problem for organizations and could be prohibitively expensive for some private companies, particularly smaller ones (OCOL, 2005, p. 29):

“The documents posted on-line on a daily basis in the large departments are too numerous, too diversified, and too specialized, and the approval processes were too decentralized and scattered to hope that the Webmasters and their teams could determine whether it was relevant to post a particular document, or assess the quality of the language used in the documents proposed for posting, where their primary concern is the upgrading and technical maintenance of the Web sites.”

A third reason for this study is that some substitutes for professional translation, like machine translation, although faster and relatively less expensive than professional translation, do not provide a sufficient level of quality (OCOL, 2005, p.60-61):

“For example, senior management at the Translation Bureau has rightly stated that it would be unreasonable, at this time, to use technolinguistic tools to translate the content of several government Web sites automatically,

because any drop in translation costs would inevitably come at the cost of a dramatic drop in quality."

This linguistic digital divide on the World Wide Web also extends beyond the boundaries of the Canadian federal government and has ramifications for businesses in general (Hutchins, 2001, p. 7):

"In addition to electronic messages, the amount of information available in text form on the Internet can now be counted in the hundreds of millions of pages, and it is growing exponentially at a very high rate. The non-English content is now estimated as approaching 60% of the total, and it is growing faster than English. There is no doubt that readers everywhere prefer to have text in their own language, no matter how flawed and error-ridden it may be, rather than to struggle to understand a foreign language text."

The point made by Hutes (2001, p. 7) can be understood for businesses in this way: being able to read content on the web in one's native language is necessary, but it is not entirely sufficient in order to obtain and maintain customers online. As businesses engage customers through the Internet, the issue of language quality becomes a significant determining factor as to whether or not customers will proceed with online purchases (DePalma, Sargent, & Beninatto, 2006). Translation quality is not only an important factor when viewing World Wide Web content, but also for consumers who receive products from companies that operate in different languages or cultures (Jorgensen, 2009). When a customer "is unable to properly use the equipment because the operating manuals are not properly translated or are not translated at all, you can expect to lose their repeat business. The next time the customer wishes to purchase your product, they will recall their dissatisfaction with the guides that accompanied your product and choose your competitor's offering instead" (Jorgensen, 2009, para. 12). Poor or non-existent translations associated to a company's products or services can have consequences that may impact "customer loyalty, compliance with regulatory requirements, worker safety, legal liability, increased calls to customer service, and overall customer satisfaction" (Jorgensen, 2009, para. 23).

How then can considering harnessing collective intelligence for translation potentially solve these problems of scarce resources, an overabundance of work, and the production of effective translation products with acceptable quality? While still in the realm of Web 2.0, crowdsourcing fundamentally differs from other aspects of Web 2.0-style distributed innovation. "Crowdsourcing has the 'Web 2.0'-style attribute of increased interactive

participation by large numbers of online users. But unlike user-generated content, social networks, and other popular trends, participants in a crowdsourcing ecosystem have little or no contact with each other. In particular, one worker cannot see the results of another's work" (Alonso, Rose, & Stewart, 2008, p. 10). In essence, in crowdsourced systems, participants act anonymously. There are other conditions which also differentiate crowdsourcing from other forms of Web 2.0 distributed innovation. Surowiecki posits that under the right conditions, crowds may be capable of making better decisions than any single individual, even an expert, but that in order for this to occur, the crowd must have the following characteristics: a *diversity of opinion, independence, decentralization*, and a means by which the group's private judgments can be *aggregated into a collective judgment* (Surowiecki, 2005).

"No matter who you are, most of the smartest people work for someone else", states Bill Joy, Cofounder of Sun Microsystems. Known as Joy's Law, his statement illustrates the idea for many organizations, "that in any given sphere of activity most of the pertinent knowledge will reside outside the boundaries of any one organization, and the central challenge for those charged with the innovation mission is to find ways to access that knowledge" (Lakhani & Panetta, 2007, p. 2). Therefore, the issue of scarcity **might** be addressed through crowdsourcing since the nature of crowdsourcing is to create an open call to people over the Internet to solve a specific problem: it is reasonable to assume that the total number of potentially qualified people (e.g. bilingual Internet users) is greater than the total number of professional Canadian translators. The issue of an overabundance of work **might** be also addressed through crowdsourcing for the same reason; the more people that there are to do work, the less work there is that remains.

In this context, crowdsourcing could also be seen as a way of creating a substitute product or service, e.g., a substitute translation. In Porter's Five Forces model, crowdsourcing acts as a threat of a substitute product or service. "A substitute performs the same or a similar function as an industry's product by a different means... Substitutes are always present, but they are easy to overlook because they may appear to be very different from the industry's product" (Porter, 2008, p. 17). Crowdsourcing, as a system of distributed innovation (Lakhani & Panetta, 2007), creates a different means for producing a substitute product. "The threat of a substitute is high if it offers an attractive price-performance trade-off to the industry's product... Of course, the substitution threat can also shift in favor of an industry, which bodes well for its future profitability and growth potential" (Porter, 2008, p. 17-18).

1.2 Anticipated Contribution

This research aims to examine if crowdsourcing may be a possible, if not effective, technique of helping to bridge the linguistic digital divide within Canada. The research conducted and described in this thesis may be useful for others who wish to further examine solutions to this or similar problems. We expect that the data collected and findings of this research may help to attain these goals.

This research may contribute to gaining a better understanding of the effectiveness of using non-professional bilingual individuals to create translated web content in Canada.

This research might lead to a better understanding of the effectiveness of using many language professionals to create collaboratively translated web content in Canada, as opposed to the common practice of a single translator working alone.

An anticipated contribution of this research could be a better understanding of how crowdsourced translations are accepted among both a non-professional bilingual community and also a language professional community. In terms of bridging the linguistic digital divide in Canada, this research will contribute to a better understanding of whether this technique would be accepted by a bilingual audience in general.

Lastly, this research will likely indicate whether this technique can be applied on a much larger scale in subsequent research activities. The nature of the technology and the software tool that was developed to coordinate the crowdsourced translation activities could, with minimal changes, be applied to additional language pairs, or to different types of source language texts.

Chapter 2: Theoretical Framework

To the best of our knowledge, no studies have been conducted that examine the use of crowdsourced translation as a possible means of bridging the Canadian linguistic digital divide. However, there are some known cases of businesses that have experimented with crowdsourcing different types or aspects of translation work (dotSUB LLC, 2009); (Sun Microsystems, 2008); (TED Conferences, 2009); (txteagle, 2009); (IBM, 2010). In most of these cases, the types of translation that have been crowdsourced are limited either to single word or terminology translation, software user-interface localization, or audio transcription and translation of online video subtitles. Within academia, there have been some studies that address crowdsourcing in general (Alonso et al., 2008; Brabham, 2008; Huberman, Romero, & Wu, 2008; Kaisser & Lowe, 2008; Kittur, Chi, & Suh, 2008; Su, Pavlov, Chow, & Baker, 2007; Van Den Ende, 2008) but fewer cases of crowdsourcing related research as it pertains specifically to languages (Snow et al., 2008). With respect to the specific issue of translation, our literature review did not uncover any peer-reviewed material that discusses the idea of using crowdsourcing as a means of bridging the linguistic digital divide that exists on the World Wide Web in Canada. Only two such literatures were found that address the issue of translation and crowdsourcing as its main focus (Eagle, 2009; Ledlie, Odero, Minkov, Kiss, & Polifroni, 2009). There is some literature published that discusses the use of wikis to assist with collaborative translation performed primarily by virtual communities of language professionals (Bey et al., 2008; Désilets, 2007; Huberdeau et al., 2008). In these cases, wikis were used as the main collaborative tools for professional translators in creating content. In some cases, the wikis supported the simultaneous translation of content into multiple target languages.

Considering the diversity of languages on the web today, the general state of Internet connectivity, and the lack of academic research focused on evaluating the efficacy of crowdsourced translations of web content, all of these aspects combine to create a unique research opportunity that may be useful in understanding and helping to address the specific Canadian version of this wider problem – bridging the linguistic digital divide.

We hypothesize that crowdsourced translation may be a possible means of producing acceptable quality translations of website content when compared with machine and professional translations. It is the purpose of this study to investigate the validity of using crowdsourced translation as a means by which the Canadian linguistic digital divide may be bridged. If this form of translation can be shown to be capable of producing acceptable

quality translations in a reasonable time at a reasonable cost, it may be useful for companies, individuals and government agencies to consider this approach as one means of bridging the linguistic digital divide that exists for Canadians online.

The next sections of this chapter describe the theoretical underpinnings of the system that was used in this study to generate crowdsourced translations. This type of system is herein referred to generally as Collaborative Online Language Transfer.

2.1 Collaborative Online Language Transfer

Web 2.0 technologies have been catalysts for the emergence of online collaboration and peer production. These disruptive technologies are among the drivers for business model changes in many sectors, including the language service industry. This section describes a conceptual framework in which the techniques of Lightweight Peer Production can be joined with the principles of Document Engineering for e-Business to enable a Collaborative Online Language Transfer System. Some of the key measurements needed in order to understand the performance of this system will be emphasized.

Historically, the process of transferring written content between languages has been in the hands of non-professional bilingual individuals (Garcia, 2009). Only since the middle of the 20th century has the process of translation moved from the domain of the amateur into the domain of the professional (Garcia, 2009). During this era of professional translation there have been many disruptive innovations, such as machine translation and translation memory. It now appears that the Internet and more specifically Web 2.0 technologies in particular have the most potential to drive new innovation (Garcia, 2009).

Professional translators, in light of these changes in the Web 2.0 era, struggle to cope with the seemingly endless supply of new content on the web (Garcia, 2009). It is not without surprise that language service providers have begun to look toward collaborative Internet technologies (Guyon, 2010, para. 27-28):

"If we were to use all the benefits that today's web-based tools, with their collaborative platforms, have to offer, we would have a system that would enable us to share a text of unlimited length among an unlimited number of translators and revisers... [w]e would have revisers starting work minutes after translators instead of hours later. Peer review would be almost instantaneous, and validation of questionable parts by peers would be done by experts in the field."

Purchasers of language services are also confronted by the challenges of the Web 2.0 era. Professional translation and localization methods, while of high quality, are “relatively slow and only affordable to big players on big projects” (Garcia, 2009, p. 204-205). The presence of free online machine translation websites such as Google Translate, Bing Translator and Yahoo Babel Fish represent quick and cheap substitutes for professional translation. However, machine translation is not without its drawbacks, most notably in the area of diminished quality. “... it is still the general consensus that it is not yet up to the standard required to be used for dissemination” (Garcia, 2009, p. 206).

Little yet is known about the nature of the system as imagined by Guyon (2010). This thesis refers to this class of system as Collaborative Online Language Transfer and defines it as an asynchronous web-based system enabling multiple bilingual individuals to transfer in parallel the linguistic content of a source text into a target text. In Figure 1, we see the simplest model of the translation process. Here, we see a source text being transformed by the translation process into a target text translation.

Figure 1 - Black Box Model



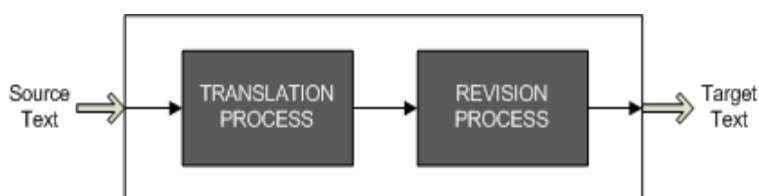
This chapter proposes a conceptual framework that uses the techniques and ideas of Lightweight Peer Production within the context and language of the field of Document Engineering for e-Business. The literature referred to within this conceptual framework is drawn on to explore the notion of a Collaborative Online Language Transfer System.

2.2 Performance Criteria

Which measures should be employed to determine how a Collaborative Online Language Transfer System performs? In design research, the performance of a designed artifact is made by observing it and assessing whether the designed artifact works or not (Lee & Nickerson, 2010). Conceptually, a Collaborative Online Language Transfer System should be judged in its performance primarily in three ways: first by the performance of the translation process itself; second, by the performance of the target text in the context of the source text; and third, by the performance of the target text in response to the qualitative translation goals (*translation brief*). The terms *source text* and *target text* are widely understood in the field of translation to refer to the input and output documents of the translation process.

The performance of the translation process can be measured quantitatively in several ways. We can measure quantitatively the time that the total translation process took and also any monetary benefits bestowed upon its participants. If intrinsic motivators were used to encourage participation in a Collaborative Online Language Transfer System, it may be possible to quantify and measure these as well. As noted by Guyon (2010), the translation process can be further subdivided into a distinct process of translation and a distinct process of revision; these sub-processes can be similarly measured quantitatively. In Figure 2, we see the original black box model in light of these distinct processes.

Figure 2 - High-Level Process Model



For this type of system, it is easy to conceive of additional quantitative measures for individual participant performance, for example, percentages of work completed and abandonment rates.

With respect to the performance of the target text, its quality cannot simply be measured by the absence of lexical or grammatical errors or by poorly translated semantics alone. For some types of translation, the main question of quality is whether or not the target text has the same impact as the source text (House, 1997, p. 24):

"The fundamental characteristic of a translation is that it is a text that is doubly bound: on the one hand to its source text and on the other hand to the recipient's communicative conditions... In other words, the equivalence relation equals the relation between a source text and its translation."

For example, if to translate the phrase, "I are to a good student when me not skippin' school" into a target language, if one were to assess the quality of target text without considering the quality of the source text, it may be easy to misinterpret the quality of the target text. The quality of the source text must be considered when assessing the quality of the target text; the relative differential quality of the two texts should inform how the target text's performance is interpreted (House, 1997).

The target text's performance should also be judged by how well or how poorly the qualitative goals of the translation were responded to by the translators; this is referred to

in the work of Nord (2006) as a *translation brief*. In essence, the qualitative goals of the translation (*translation brief*) are a set of instructions that are provided to the translators giving some context as to how to interpret and translate the source text. Finally, the performance of the target text can be measured as a *recipient evaluation* (Bowker, 2009); which means that it is the intended audience of the translation that should perform the assessment of the translation's quality. For example, the bilingual Canadian participants in this study could evaluate the target text outputs of this system in terms of cost, speed, and quality, determining whether the target texts were satisfactory.

Under which conditions should such a system's performance be measured? "The strategy is for a researcher to eschew the random collection of large quantities of data and instead assess the performance of the design or theory under a single, well-specified set of circumstances deliberately sought out to be challenging to it" (Lee & Nickerson, 2010). Clear qualitative indications of the target text's performance could be made by assessing how native speakers in the language of the target text respond after reading the target text. If the quality is acceptable, we posit that a native speaker would likely have positive impressions of the target text. For example, if the source text deals with marketing material encouraging purchasing behavior and the target text were of acceptable quality, the following types of questions may be clear indicators of the performance of the target text:

After reading the text, I feel comfortable contacting this company in my native language.

After reading the text, I feel confident in purchasing products or services from this company.

After reading the text, other native speakers would not be able to tell that this was a translation.

Another condition which may affect the interpretation of the performance of the target text is the extent by which the knowledge that a Collaborative Online Language Transfer System was used may bias the reader of the target text. Bowker (2008) observed that professional translation efforts are sometimes viewed as a way of preserving the linguistic culture of the target language (Bowker, 2008). This study identified a bias by language professionals who would "not accept texts that had been produced with the assistance of MT [Machine Translation] software" (Bowker, 2008). This study offers the explanation that for language professionals, the "fear of loss of livelihood or fear that they

will be forced to become post-editors could potentially explain these negative feelings” (Bowker, 2008). A related sentiment is expressed by Garcia (Garcia, 2009):

“The ‘hive’ model does away with professional translators altogether in preference to a mass of volunteers/amateurs. This model brings back the pre-professional era when translators were simply bilinguals with good subject knowledge, and the ability to transfer meaning between languages. This model would be supported by a few professionally trained translators occupying key terminological or QC [Quality Control] roles in the background.”

Little yet is known about the optimal compositional makeup of individuals performing translation in a Collaborative Online Language Transfer System, nor the extent to which the injection of non-professional bilingual individuals within the process of translation may bias the acceptance of the outputs from this type of system. As there exists a shortage of professional Canadian translators (OCOL, 2005), is it reasonable to assume it could be possible to create a crowdsourced translation with acceptable quality from non-professional bilingual individuals whose abilities in translation may likely be less than those of professional Canadian translators?

In this vein, Hong and Page (Hong & Page, 2004) ask, *“Can a functionally diverse group whose members have less ability outperform a group of people with high ability who may themselves be diverse?”* Known as the diversity trumps ability theorem, Hong and Page posit that this possibility exists under the right circumstances. The four characteristics are necessary in order for this theorem to be successful: “(i) Agents are intelligent: given any starting point, an agent finds a weakly better solution, and the set of local optima can be enumerated, (ii) The problem is difficult: no agent can always find the optimal solution. (iii) Agents are diverse: for any potential solution that is not the optimum, there exists at least one agent who can find an improvement. (iv) the best agent is unique” (Hong & Page, 2004). Viewing this theory in the context of this study raises the following questions: are there conditions in which crowdsourced translation can have similar or better quality, lower cost and shorter turnaround times than professional translation? If so, what are those conditions? If not, which conditions adversely affect the performance of crowdsourced translation?

For the problem of translation, Hong and Page’s criteria (Hong & Page, 2004) may be appropriate since non-professional bilingual individuals can be said to be intelligent

regarding the languages to be translated. Translation, as a problem, is difficult and this is evidenced by both the lack of quality translations produced from machine translation software (Bowker, 2008) and by the difficulty of translating documents that use "terminology that few translators know" (OCOL, 2005). No translator can be said to always find the optimal translation since "no definition of error gravity has been proposed on a scientific, theoretical, textological basis, and evaluators have to rely on ill-defined concepts such as 'complete failure to render the meaning' and 'essential part of the message'" (Williams, 2004).

Hong and Page's criteria may also be appropriate in the context of this study since the problem of translation is such that it is possible to improve upon a solution through the better use of language, grammar, terminology or fidelity to source text argumentation (Williams, 2004). Since the evaluation of translation quality is often "left to the discretion and experience of the co-ordinator/evaluator" (Williams, 2004), it is reasonable to assume that it is unlikely that multiple translations of the same source text by different people will result in identical "best" translations perceived a single evaluator.

2.3 Lightweight Peer Production

"What are the optimal structures for contributory practices?"
(Haythornthwaite, 2009, p. 1)

Collaborative Online efforts have begun to fall into two main categories for participation and production. Haythornthwaite (2009) identifies these categories as Heavy Weight Peer Production (HWPP) and Light Weight Peer Production (LWPP). Whereas HWPP refers to systems which employ a virtual community model, LWPP instead refers to systems which employ a crowdsourcing model (Haythornthwaite, 2009). To further this distinction, HWPPs are characterized by a small set of negotiated contributions requiring high investments of time and energy on the part of the participants, who are governed by gatekeepers, and coordinated through reciprocal interdependence. In contrast, LWPPs are characterized by a large number of atomistic, discrete and independent contributions, requiring low investments of time and energy on the part of the participants, who are governed by rule-based mechanisms, and coordinated through pooled interdependence (Haythornthwaite, 2009).

At first glance, LWPPs appear to be a better fit than HWPPs for a Collaborative Online Language Transfer System simply because LWPPs are designed in a way to maximize participation; this is a key feature of Guyon's who states that *"...we would have a system that would enable us to share a text of unlimited length among an unlimited number of translators and revisers"* (Guyon, 2010, para. 27). LWPPs encourage maximum participation by their very nature: they allow for anonymity, offer open membership, minimal effort to enter, provide tasks that are simple to perform and that are governed by clearly defined rules which are implemented by the system itself (Haythornthwaite, 2009). The characteristics of LWPPs have the added benefit of helping to *"alleviate coordination among contributors or contributions"* (Haythornthwaite, 2009, p. 2).

Conversely, HWPPs have characteristics that discourage maximum participation. HWPPs require of the participants a *"commitment to the enterprise as a whole, including internal processes as well as products, the social and emotional experience of the community, and its continued existence"* (Haythornthwaite, 2009, p. 2). In HWPPs, it is the participants themselves who determine the direction of the collaborative enterprise, including its *"goals, purposes and processes"* (Haythornthwaite, 2009, p. 3). In HWPPs, individual contributors examine the work of other participants and decide in an ad-hoc fashion how to contribute (Haythornthwaite, 2009). Characteristic of HWPPs are high-effort

membership and gate-keeping which act as mechanisms for the purpose of limiting participation, especially from those who are outside of the community (Haythornthwaite, 2009). Recent efforts in collaborative translation using wikis (Bey et al., 2008; Désilets, 2007; Huberdeau et al., 2008) can be classified and understood in this context as HWPP approaches to translation.

For the crowdsourcing model to be effective, the LWPPs must guide participation with rules embedded in the system itself. Surowiecki offers his view that the most effective crowds have rules governing the *diversity of opinion, independence, decentralization* and *aggregation* (Surowiecki, 2005). Diversity of opinion is defined as the cognitive diversity of the group. When a group contains too many people of a similar background, education or training, that group's collective judgement suffers because everyone tends to think in the same way (Surowiecki, 2005). Independence is defined as the capacity for people within the group to be able to express their individual judgements without the knowledge or influence of other member's judgements. Conversely, when members of a group are allowed the opportunity to influence other members, 'information cascades' may begin to develop, whereby the individuals in the group are paying more attention to what the group 'knows', rather than what they themselves know (Surowiecki, 2005). Decentralization is defined as when the decision-making of the group is not solely in the hands of a single individual, but rather that it is expressed by the whole of the group (Surowiecki, 2005). Aggregation is defined as when the information that resides in the hands of only a few individuals has the capacity to be expressed in the whole of the group's judgment and is not buried or hidden (Surowiecki, 2005).

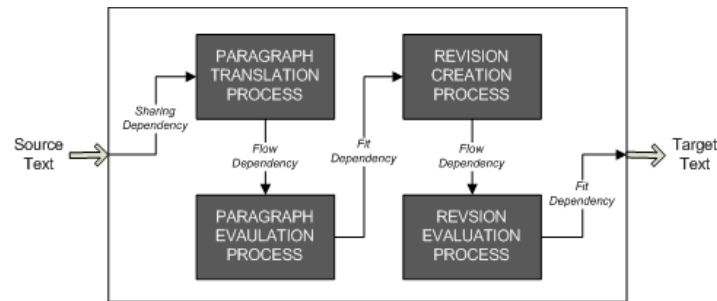
In a system in which there could be "*an unlimited number of translators and revisors*" (Guyon, 2010), many of who would be functioning as non-professional bilinguals with less experience and ability than their professional counterparts, the concept of *task repetition* becomes a mechanism to increase the probability that the participation in this type of system produces a higher quality output. This concept was described in the Ph.D dissertation of Luis von Ahn, entitled "Human Computation" (von Ahn, 2005). As part of his dissertation, von Ahn created a series of games with small, repeatable tasks. From these online games, he used the concept of *task repetition* as a way for large numbers of paired anonymous participants to take guesses at labeling images with keywords. When both anonymously paired participants provided identical keywords, these matches were recorded. To ensure the probability that a set of matched keywords was likely to actually describe the content of the image, the image was sent to multiple pairs of anonymous

participants (von Ahn, 2005). In this context, it is understood then that increases in *task repetition* should increase the likelihood of achieving a better outcome for a given task.

The concept of *task repetition* as a mechanism for increasing the probability of acceptable quality for the problem of translation in an LWPP system appears in the case of Eagle (Eagle, 2009). The txtEagle mobile crowdsourcing service was used to translate English-language terminology into over 15 Kenyan languages through the use of mobile phones (Eagle, 2009). The txtEagle service sent the same translation task to multiple users; this had the effect of increasing quality “as a function of the number of independent users responding to the same task” (Eagle, 2009). Eagle shows that when task repetition approaches eight independent users, the integrated results exceed a 95% quality level (Eagle, 2009).

Garcia indicates that current web-based translation software packages typically deliver only a single segment of a source text to a professional translator at any one time instead of the entire source text (Garcia, 2009). Combining this notion with the concepts of *task repetition* and the previously mentioned distinct processes of translation and revision, we consider the following additional processes are necessary for a Collaborative Online Language Transfer System that embodies the characteristics of an LWPP. These processes are paragraph translation, paragraph evaluation, revision creation and revision evaluation. The process of paragraph translation necessarily involves presenting to each participant the entire source text to translate but only a single paragraph to be translated at a time. Additionally, instructions regarding the language in which to target the translation and the *translation brief* (Nord, 2006) of the source text should also be presented to the participant. As each individual paragraph should be translated multiple times according to the concept of *task repetition*, there must necessarily be a mechanism within the system to aggregate the translated versions of each paragraph into the likely candidate of highest quality. The process of paragraph evaluation therefore culminates in the creation of a ‘rough draft’ of the translation, where each paragraph’s “best” translation is then used as input to the remaining processes of the system. In Figure 3, we see the conceptual model in light of the principles of Lightweight Peer Production.

Figure 3 - Lightweight Peer Production Model



The processes of revision creation and revision evaluation follow a similar pattern as the paragraph processes. In the revision creation process, multiple requests for revisions of the 'rough draft' are provided to the participants. Like the process of paragraph translation, instructions regarding the target language and the *translation brief* (Nord, 2006) of the source text should also be present in order to guide the participants in how to proceed with the revision creation. Since the rough draft would be composed of the work of multiple participants, another goal of this task should be to instruct the participants to revise the content in such a way that it appeared to originate from a single author. The process of revision evaluation allows the participants to consider pairs of revisions and individually determine which revision is better. The participants' diverse, aggregated, decentralized and independent judgments would ultimately result in the target text output of the system.

This view of the system is additionally annotated with some of the ideas from Coordination Theory (Malone, Crowston, & Herman, 2003). Poltrock and Handel state that Coordination Theory is an ostensive framework delineating dependencies in collaborative systems by declaring sharing, flow and fit dependencies (Poltrock & Handel, 2009). "According to Coordination Theory, the purpose of collaboration is to satisfy dependencies, and the business process is constructed for that purpose" (Poltrock & Handel, 2009). The source text and the qualitative goals (*translation brief*) of the translation in this system act as a sharing dependency. Multiple paragraph translations enter into the paragraph evaluation process as a flow dependency, and they exit this process as a fit dependency in the form of the 'rough draft translation'. The revision creation process generates multiple revisions of the 'rough draft translation' which then enter the revision evaluation processes as a flow dependency and then exit as a fit dependency in the form of the target text.

The use of temporal modeling as an approach to performative measurement could be considered for this type of system. This form of performative modeling analysis may be a longitudinally useful way of optimizing participation. By embedding 'predictive qualities' into

the system, this performative modeling approach could help to govern issues related to workload coordination (Poltrock & Handel, 2009).

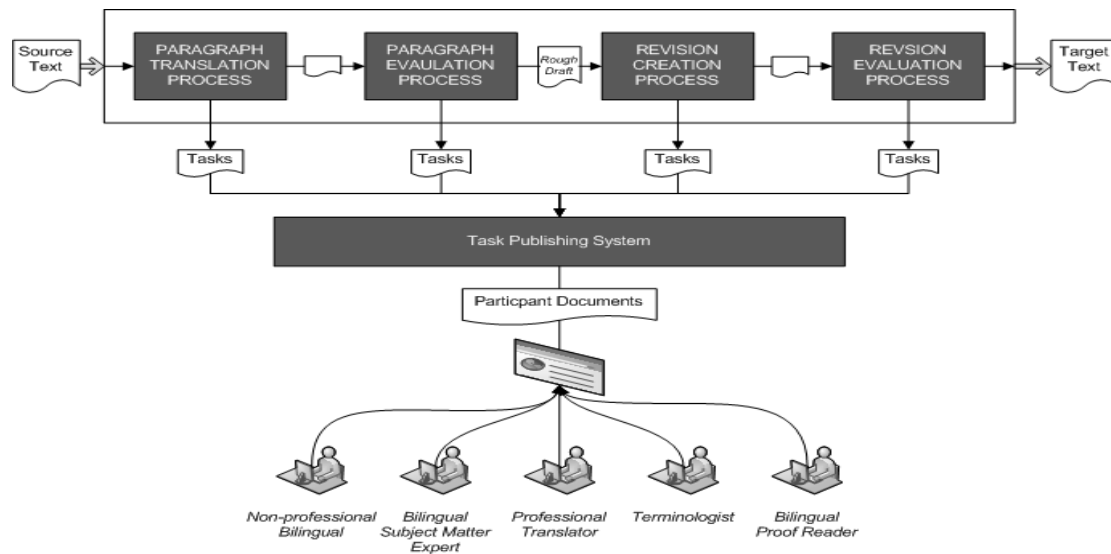
2.4 The Landscape of Document Engineering

"Document Engineering is not intended to be a prescriptive methodology. Rather it presents a landscape in which to place our thinking." (Glushko & McGrath, 2002)

One of the most important ideas from the emerging discipline of Document Engineering for e-Business is the concept that it is the document and the document alone that acts as the 'public interface' between business processes. Document Engineering depicts business processes as loosely coupled architectures, where end-to-end transactions are expressed with discrete, well-defined, unambiguous, and reusable documents (Glushko & McGrath, 2002). The landscape of Document Engineering allows us to consider a Collaborative Online Language Transfer System "as a chain of document exchanges" (Glushko & McGrath, 2002). The Black Box View of the translation process (Figure 2) illustrates this idea with a source text and a target text. Each of the sub-processes illustrated in the Lightweight Peer Production View (Figure 3) can also be understood in this context. Each sub-process should have a beginning and an end with discrete, well-defined and unambiguous characteristics that could be expressed as documents. Furthermore, the characteristics of LWPPs dictate that individual tasks in the system are atomistic in nature with low-effort to participate (Haythornthwaite, 2009). Each individual task created by the sub-processes of Collaborative Online Language Transfer system could ultimately be published to the participants through discrete, well-defined, unambiguous, and reusable documents.

Here we see Figure 3 transformed within the context and landscape of Document Engineering. Figure 4 shows an asynchronous web-based system enabling multiple bilingual individuals to transfer, in parallel, the linguistic content of a source text into a target text.

Figure 4 - Collaborative Online Language Transfer System Model



2.5 Problem Statement

The translation of website content is an essential means of bridging the linguistic digital divide in Canada. There exist a number of obstacles that prevent website content translation from bridging this linguistic digital divide, including a shortage of professional translators, an overabundance of website content, and insufficient quality translations produced by existing alternative methods such as machine translation (OCOL, 2005). We hypothesize that the technique of crowdsourcing may be a possible means of producing an effective translation of website content when compared to machine translation and in some cases when compared to professional translation. Further, no studies have been conducted, to our knowledge, that examine the characteristics for estimating the effectiveness of crowdsourced translations of website content as substitutes for professional translations of website content. Therefore, the general research question of this thesis is:

Can a model be created that can estimate the effectiveness of crowdsourced translations as a means of bridging the Canadian Linguistic Digital Divide?

The concept of a professional translation may be defined as a translation which is performed by one or more professional translators working either individually or collaboratively. The concept of a crowdsourced translation may be defined as a translation performed by a group of people who have some capacity to translate. The concept of an “effective estimation” may be defined by considering the elements of the design of the crowdsourced translation (i.e. how the information was collected and compiled) along the dimensions of time, cost,

and acceptable quality. The following specific questions arise from this general thesis question:

How does crowdsourced translation compare in terms of quality, cost, and speed to alternate methods of translation, such as professional translation and machine translation?

What level of task repetition in crowdsourced translation provides for overall translation quality?

What level of diversity in crowdsourced translation provides for overall translation quality?

Which factors related to a possible bias against crowdsourced translation affect the overall translation quality of crowdsourced translations?

How is the idea of crowdsourced translation accepted as a means of bridging the Canadian Linguistic Digital Divide?

The intent of this research is to examine these questions in detail and provide a set of answers to assess the effectiveness of applying crowdsourced translation techniques to website content as a means of bridging the linguistic digital divide within Canada. To these ends, the study has been divided into two main phases: Phase I (Translation Creation) and Phase II (Translation Evaluation). The chapters that follow present the methodology and analysis for each phase chronologically.

Chapter 3: Phase I (Translation Creation) - Methodology

3.1 Overview

This study received approval from the University of Ottawa Research Ethics Board. For the purposes of this study, the Amazon Mechanical Turk API (Amazon, 2007) was used as the basis for the task publishing system depicted in Figure 4. Amazon Mechanical Turk acted as the instrument used to collect the raw crowdsourced translation data from the participants during this phase. Software, herein referred to as Confluence, was developed and purpose-built for this thesis using the Amazon Mechanical Turk API to allow the source text to be transformed by the process depicted in Figure 4. The Amazon Mechanical Turk API allowed Confluence to record data for most of the performance indicators that are described later in the section on system performance.

Participants in both phases of the study were classified into two groups – Language Professionals (LP) and Non-Language Professionals (NLP). Language Professionals, for the purposes of this study, were defined as individuals who were fluent in both Canadian English and Canadian French and who also possessed translation education, translation training, or professional translation work experience. Non-Language Professionals, for the purposes of this study, were defined as individuals who were fluent in both Canadian English and Canadian French, but who lacked translation education, translation training, and professional translation work experience.

Four experiments were conducted between July 2010 and December 2010 using the system. The same Canadian English source text was selected for use in each of the four experiments. The source text was approximately 360 words long and was formatted into three paragraphs. The source text was a restaurant review for a café in the city of Toronto, Ontario, Canada, and was found on a restaurant blogging website using Creative Commons licensing. Participants were recruited through email; upon qualifying for an experiment, they were asked to translate and revise the source text into Canadian French using qualitative translation goals in accordance with the translation quality model defined by House (House, 1997). Each participant was paid \$1.50 USD for each task that they completed and given a maximum of two hours to complete each task. Amazon collected a 10% processing fee for all work requested through Amazon Mechanical Turk API (Amazon, 2007). Participants were notified both via email and via Twitter when new sets of tasks were available to be worked on.

To test the concept of task repetition as an indicator of quality in crowdsourced translation, each participant group was given two experiments. The first experiment in each participant group used a task repetition value of two, indicating that two unique translations of each paragraph and also two unique revisions of the rough draft were requested. The second experiment in each group used a task repetition value of four, indicating that four unique translations of each paragraph and also four unique revisions of the rough draft were requested. The uniqueness restriction was implemented in the software using the Amazon Mechanical Turk API (Amazon, 2007). For the paragraph evaluation tasks and the revision evaluation tasks depicted in Figure 4, no uniqueness restriction was imposed and all such tasks could have been completed by any participant in the experiment. The four experiments that were conducted are labeled and will be subsequently referred to in the following ways:

CSTLP2 - Crowdsourced Translation, Language Professionals, Task Repetition Value Two;

CSTLP4 - Crowdsourced Translation, Language Professionals, Task Repetition Value Four;

CSTNLP2 - Crowdsourced Translation, Non-Language Professionals, Task Repetition Value Two; and

CSTNLP4 - Crowdsourced Translation, Non-Language Professionals, Task Repetition Value Four.

To compare the quality of crowdsourced translation to professional translation, three Canadian translation companies were hired to translate the source text into Canadian French. The purpose of purchasing three professional translations was to obtain a solid baseline for comparison within the study. Likewise, to compare the quality of crowdsourced translation to machine translation, the source text was translated into French using three online translation websites. It was expected that the nature of the source text would make it difficult for machine translation software to provide translation of adequate quality. While there are commercially available machine translation systems for under \$500 (Bowker, 2008), and while these systems provide better quality than freely available online machine translation websites (Bowker, 2008), because of the wide availability and low cost of online translation systems, it is reasonable to expect that the average Canadian may turn to online machine translation websites instead of to commercially available machine translation systems.

At the conclusion of Phase 1, ten translations of the source text (four CST, three PT and three MT) were presented to a translation PhD student. The translation PhD student was hired to rank the translations in order of quality and to answer short questionnaires about aspects of the translation quality. At no time was the translation PhD student told how any of the translations were created. The highest ranked PT and MT were selected for use along with the four CST in Phase 2 of the study.

3.2 Source Text

The source text that was selected for the study was licensed with a Creative Commons Attribution-Noncommercial-Share Alike 2.0 (Canada) license. This form of Creative Commons licensing allowed for the researcher to copy, distribute and transmit the work, without the express permission of the copyright holder. The following text was selected as the Canadian English source text of the study. It is 360 words long, and has been modified to span three paragraphs of roughly equal length. During this study, this text was translated from Canadian English into Canadian French three times by professional translators, twice by Language Professionals, twice by Non-Language Professionals, and three times by machine translation software, for a total of ten Canadian French translations.

Figure 5 shows the source text that was used in this study. The source text was originally reviewed on June 12, 2010, written by Gadjo Seville and was taken from <http://www.blogto.com/restaurants/blackcamel>. The source text was published under a Creative Commons Attribution-Noncommercial-Share Alike 2.0 (Canada) license, <http://creativecommons.org/licenses/by-nc-sa/2.0/ca/>.

Figure 5 - Source Text

The Black Camel has been slinging sublime sandwiches since it opened in 2004 and their approach is as smart as it is simple. Starting with five basic sandwich choices, they allow you to personalize your meal by choosing from a range of high-quality in house prepared sauces and toppings. Options include slow roasted beef brisket, pulled pork shoulder, seared steak, roasted pulled chicken and roma tomato, red pepper, arugula and eggplant. Prices start at \$7.00 (with one sauce). There are also plenty of sauces and condiments and extra toppings available for less than a buck. So popular are these sandwiches that Black Camel uses their Twitter to let people know when they've sold out which, on most days, is before their scheduled closing time.

Also on the menu is a Black Camel chili (\$3.25 for 6 oz., \$5.25 for 12 oz.) and a variety of multi-roast blend coffee and loose-leaf tea selections from the nearby House of Tea. For breakfast, there's an eggspresso sandwich served until 11 am. Eggs are cooked a little different than usual (they use an espresso machine steam wand) and paired with either pancetta ham, steak or house-cured salmon, sweet pepper and red onion mayo). After lining up for around 5 minutes we order lunch. First is the slow roasted beef brisket with the signature Black Camel BBQ sauce and a slathering of chipotle mayo and creamy horseradish.

The huge and heavy sandwich is loaded with super tender beef brisket (slow-cooked for 12 hours) that is browned with molasses and caramelized onion and pungent with the flavour of the BBQ sauce and the mild zing of creamy horseradish. It's amazing and incredibly filling. We also decide to try the roasted pulled chicken with roasted sweet red pepper. It's moist and equally tasty although lighter than the imposing beef brisket sandwich. Both sandwiches come with generous helpings of meat and sauce and it takes a monumental effort to keep everything together in that large bun. Seating is extremely limited and it can get pretty crowded inside so it's best to try and get seating outside or find a shady spot on any of the nearby park benches.

Prior to the beginning of the Translation Creation phase, a group of fifteen bilingual professors and staff at the University of Ottawa were informally asked to estimate on a seven point Likert scale how easy or how difficult they thought it would be for a bilingual person to translate the source text into French. In this simple poll, a value of one represented that the source text would be extremely easy to translate into French, whereas a value of seven represented that the source text would be extremely difficult to translate into French. The purpose of collecting an estimate of difficulty for the source text was to assess the appropriateness of the source text for the purposes of this experiment. Should the source text have been either too difficult or too easy to translate, it may have been unsuitable because a source text at either extreme might have jeopardized the value of the study. Table 1 shows the estimated difficulty of translating the source text into Canadian French.

Table 1 - Estimated Source Text Difficulty

Number of Samples	15
Standard Deviation	1.429286
Confidence Interval @ 0.05 α	0.723
Average	4.9 out of 7 points
Difficulty Likelihood	70%

3.3 Qualitative Translation Goals (*Translation Brief*)

Both the Phase I research participants and the professional translation companies were provided with the *translation brief* for the translation process, herein referred to as the qualitative translation goals. No such qualitative translation goals were provided to any of the online machine translation websites. The qualitative translation goals were specified as follows:

The source text was written for an educated middle class audience, using language that is appropriate in 2010. The source text was written to be read (and not to be spoken aloud), and it is intended evoke action on the part of the reader. The source text was written in a casual style by a professional blogger.

3.4 Amazon Mechanical Turk

Briefly, Amazon Mechanical Turk acts as an online job board whereby small tasks can be published to the web via custom-written software. To use the Amazon Mechanical Turk website, participants were first required to register for an account with Amazon.com. Once the participants registered for an account with Amazon.com, they proceeded with user registration on the Amazon Mechanical Turk website (<http://www.mturk.com>). Amazon collected information regarding the participants' full names, addresses, cities, provinces, postal codes and phone numbers. None of the information that Amazon collected during the user registration processes was available to the researcher at any time during the study.

Participants used an invitation qualification code provided by the researcher to access the crowdsourced translation tasks on the Amazon Mechanical Turk website for the experiment in which they had been invited to participate. The invitation qualification code reasonably ensured that only the correct people could participate in the study. If and when an incorrect invitation qualification code was entered, the person was barred from retrying

to qualify for a period of seven days. Once the correct invitation qualification code had been entered, the person was qualified to work on any of the tasks available for the experiment in which they had been invited. Each experiment used a different invitation qualification code. Additionally, since Amazon Mechanical Turk is used by people from many different countries, only people who had registered with Amazon using a Canadian physical address and had qualified with the correct invitation qualification code were able to participate in an experiment. Figure 6 illustrates an example of the participant qualification test that was shown to Phase I participants.

Figure 6 - Participant Qualification Test

The screenshot shows the Amazon Mechanical Turk interface. At the top, the logo reads "amazonmechanicalturk Artificial Intelligence". Navigation tabs include "Your Account", "HITS", and "Qualifications". A notification in the top right corner says "45,910 HITS available now" next to the name "Sobr". Below the navigation, there are links for "All Qualifications", "Qualifications Assigned To You", and "Pending Qualifications". A search bar contains the text "Search for HITS containing that pay at least \$ 0.00 for which you are qualified" with a "GO" button. A timer shows "00:00:33 of 10 minutes". Below the timer are two buttons: "Submit" and "Cancel". A box displays the test details: "Research Participant Qualification @ 16/06/2010 7:44:16 PM", "Author: S O'BRIEN", "Retake Delay: 7 days", and "Qualification Value: 0". The main question is "What is the invitation code?" followed by a list of radio button options: benches, blend, available, seating, served, know, five, horseradish, the, prepared, and dogwood.

When tasks were available to be worked on, the participants were notified via Twitter of the URL of each of the crowdsourced translation tasks on the Amazon Mechanical Turk website. The participants were asked follow the URL to the Amazon Mechanical Turk website, where they could then log in. Once logged in, the participants saw a preview of the crowdsourced translation task, which included a statement of informed consent. When informed consent had been granted by the participant, typically a participant would complete a crowdsourced translation task, concluding by clicking on the "Submit HIT" button on the Amazon Mechanical Turk website. Participants were allowed a maximum of two hours to complete each task; participants also had the option of abandoning a task to allow another participant to complete it. Submitting a task provided Confluence with the participant's crowdsourced translation task data. During this process, the only identifying information about the participant that was provided to Confluence was a unique WorkerID

that was generated by the Amazon Mechanical Turk website, ensuring that each participant was able to work anonymously.

3.5 Confluence Software

Confluence is the name of software that was purpose-developed to manage the process illustrated in Figure 4. Confluence was written using Visual Studio 2008, utilizing the C# programming language, Windows Communication Foundation, and a series of XML documents that embody the ideas from Document Engineering. The details of the Confluence software for Collaborative Online Language Transfer are provided in Appendix A – Confluence Software. In the next sections are brief descriptions of the sub-processes referred to in Figure 4. In each of the sub-processes, participants would see at least the following information: the statement of informed consent, the entire source text, and the qualitative translation goals. Other information specific to each sub-process is herein defined. It is also important to note that participants were allowed to complete as many or as few tasks as were available to them.

3.6 Paragraph Translation Process

Confluence begins the Collaborative Online Language Transfer process by dividing the source text into paragraphs. Confluence then publishes a number of paragraph translation tasks equal to the number of paragraphs in the source text multiplied by the *task repetition value* for the experiment. For example, since the source text was composed of three paragraphs, in CSTLP2 and CSTNLP2, six paragraph translation tasks were published, whereas in CSTLP4 and CSTNLP4, twelve paragraph translation tasks were published.

Within a paragraph translation task, participants were shown the specific paragraph to translate and provided a space on the web page in which they could translate the specified paragraph from Canadian English to Canadian French. When publishing paragraph translation tasks, Confluence imposed a uniqueness restriction; participants would be allowed to translate each paragraph only once. This uniqueness restriction ensured a minimum level of diversity within each experiment.

3.7 Paragraph Evaluation Process

Once all the published paragraph translation tasks had been completed, Confluence collected the different versions of each paragraph and then proceeded with these results as

input to the paragraph evaluation process. The goal of the paragraph evaluation process was to determine which of the paragraph translations was the best translation of each source text paragraph. Unlike the paragraph translation process, no uniqueness restriction was imposed; any of the paragraph evaluation tasks could have been completed by any participant.

Since the number of available translations per paragraph was dependent upon the *task repetition value* for each experiment, Confluence published paragraph evaluation tasks to participants according to the *task repetition value* used in each experiment. With a *task repetition value* of two, Confluence published nine paragraph evaluation tasks; this ensured that the winning paragraph translation would have survived a minimum of five binary comparisons. With a *task repetition value* of four, Confluence used a “Balanced Triple Elimination Tournament” (Czapski, 2008) algorithm to determine the winning paragraph translation.

Within each paragraph evaluation task, participants were shown two translations of the same source text paragraph and were asked to select the better translated paragraph. For the translated paragraph that they did not select, participants were asked to indicate any errors related to the qualitative translation goals, errors in grammar, or errors in lexical use. The result of these binary comparisons was a rough draft of the translation. It is the rough draft of the translation that was then used as input to the revision creation process.

3.8 Revision Creation Process

The revision creation process, in a general sense, mimicked the paragraph translation process. Confluence published a number of revision creation tasks equal to the *task repetition value* associated to the experiment. For example in CSTLP2 and CSTNLP2, two revision creation tasks were published; whereas in CSTLP4 and CSTNLP4, four revision creation tasks were published.

Within revision creation tasks, participants were shown the rough draft translation and were provided a space on the web page to revise the rough draft. The participants were also instructed to ensure that the revision appeared to come from a single author; this was used to help smooth out the possible disjointed nature of the rough draft translation.

When publishing revision creation tasks, Confluence imposed a uniqueness restriction; participants would be allowed to revise the rough draft of the translation only

once. This uniqueness restriction ensured a minimum level of diversity within each experiment.

3.9 Revision Evaluation Process

Once all revision creation tasks had been completed, Confluence used those revisions as input for the revision evaluation process. The revision evaluation process, in a general sense, mimicked the paragraph evaluation process. Since the number of revisions available was dependent on the *task repetition value* for the experiment, Confluence published revision evaluation tasks according to the *task repetition value* used in the experiment. With a *task repetition value* of two, Confluence published nine revision evaluation tasks; this ensured that the winning revision survived a minimum of five binary comparisons. With a *task repetition value* of four, Confluence used a “Balanced Triple Elimination Tournament” (Czapski, 2008) algorithm to determine the winning revision.

Within each revision, the participants were shown the source text and two revisions of the rough draft translation. The participants were then asked simply to select the better translation. The result of these binary comparisons was the target text output of the Collaborative Online Language Transfer system. These outputs became labeled as CSTLP2, CSTLP4, CSTNLP2, and CSTNLP4. The next chapter presents the analysis of Phase I (Translation Creation), followed by the methodology and analysis of Phase II (Translation Evaluation).

Chapter 4: Phase I (Translation Creation) - Analysis

As noted earlier, a Collaborative Online Language Transfer System should be judged in its performance primarily in three ways: first by the performance of the translation process itself; second, by the performance of the target text in the context of the source text; and third, by the performance of the target text in response to the qualitative translation goals (*translation brief*). This chapter will focus on the first point, which is addressed by the findings and observations of Phase I (Translation Creation).

4.1 Key Performance Indicators

The key performance indicators for Phase 1 (Translation Creation) can be broken down into the following general categories: Non-Participation, Participation, Time, Cost, Evaluation Selection, Diversity and Approximate Translation Quality. With the exception of the estimated quality and some of the participation data, the key performance indicators described in this section were mostly derived from data collected using the Amazon Mechanical Turk API (Amazon, 2007). This API provides the best glimpse into what participants actually did during the experiments.

4.2 Participation

CSTLP2 had nine participants, CSTLP4 had seven participants, and CSTNLP2 and CSTNLP4 had six participants each. Participants were recruited via email in Phase I (Translation Creation). For CSTLP2 and CSTLP4, emails were sent to certified translators and freelancers; in CSTNLP2 and CSTNLP4, a combination of personal email contacts, LinkedIn, and Kijiji.ca were used to recruit participants. The requirement letter was posted on Kijiji.ca in the following cities: Gatineau, Ottawa, Montreal, Moncton, Edmunston, and Sherbrooke. Communication with the participants was conducted via email and Twitter.

Each task completed by participants included the participant's unique Worker ID, which was generated by Amazon Mechanical Turk. The worker ID assures that the identity of each participant in Phase I (Translation Creation) was anonymous to the researcher; therefore no demographic information about the participants in Phase I (Translation Creation) can be provided in this thesis. The invitation qualification code ensured a high degree of probability that work was completed only by qualified and invited participants.

4.3 Non-Participation

Non-Participation can be broken down into two distinct groups. In the first group are people who declined to participate in any experiment. In the second group are people who were registered to participate in an experiment but who ultimately did not participate (Registered Non-Participants). Table 2 shows the breakdowns for each group across the four experiments.

Table 2 - Non-Participation

Experiments	CSTLP2	CSTLP4	CSTNLP2	CSTNLP4
People Who Declined To Participate	26	19	4	11
Registered Non-Participants	2	6	0	5

Table 3 shows for each of the experiments the tallies of the reasons given by people who declined to participate. For the people who declined to participate in an experiment, their aggregated reasons are shown in Table 3.

Table 3 - Reasons for Declining to Participate (by experiment)

CSTLP2 – Reasons for Declining to Participate	Frequency
Dislike Amazon/Mechanical Turk	2
Not Interested	2
Unavailable	22
Total	26

CSTLP4 – Reasons for Declining to Participate	Frequency
Dislike Amazon/Mechanical Turk	2
Not Interested	2
Threatens Livelihood	2
Unavailable	10
Wrong Language Direction	3
Total	19

CSTNLP2 – Reasons for Declining to Participate	Frequency
Not Fluent in French	1
Professional Translation Work Experience	2
Unavailable	1
Total	4

CSTNLP4 – Reasons for Declining to Participate	Frequency
Dislike Amazon/Mechanical Turk	4
Not Interested	2
Professional Translation Work Experience	2
Unavailable	3
Total	11

In CSTLP4, two persons responded that they would not participate because they felt that this study threatened their livelihood. In the same experiment, three of the respondents indicated that since the translation direction was from English to French, they did not feel qualified to participate as they only translate in the other direction, despite their fluency in both languages.

4.4 Time

Table 4 shows the approximate total duration for each experiment, the total idle time for each experiment in which no work was recorded/performed, and the total time for each experiment in which work was recorded/performed; these measures are also broken down by task types.

Table 4 - Approximate Time Measures

Approximate Time Measurements	CSTLP2	CSTLP4	CSTNLP2	CSTNLP4
Total Experiment Time (in days)	29	54	15.25	47.6
Total Idle Time (in days)	28.5	53.6	15	47.1
Total Work Time (in hours)	13	9.75	5.5	13
% of Idle Time	98.1	99.2	98.5	98.9
% of Total Work Time	1.9	0.8	1.5	1.1
Paragraph Translation (in hours:minutes)	2:43	3:51	3:26	3:20
Paragraph Evaluation (in hours:minutes)	9:45	2:18	1:33	7:25
Revision Creation (in hours:minutes)	0:26	1:32	0:15	1:51
Revision Evaluation (in hours:minutes)	0:6	2:2	0:10	0:22

4.5 Cost

Each participant was paid \$1.50 USD per task completed. Amazon charged a 10% processing fee for using Mechanical Turk. Table 5 shows a comparison of the total costs of each experiment with the average hourly rates for each experiment, broken down by experiment task type (Paragraph Translation, Paragraph Evaluation, Revision Creation, and Revision Evaluation). The average hourly rates were calculated by considering the individual hourly rate for each task and then deriving the average of those rates for each translation experiment and task type.

Table 5 - Cost Measurements

Cost Measurements	CSTLP2	CSTLP4	CSTNLP2	CSTNLP4
Total Translation Costs	\$72.60	\$87.45	\$72.60	\$87.45
Paragraph Translation Average Hourly Rate	\$81.34	\$161.37	\$4.85	\$17.40
Paragraph Evaluation Average Hourly Rate	\$7.17	\$27.99	\$127.63	\$7.42
Revision Creation Average Hourly Rate	\$6.85	\$11.25	\$11.82	\$4.05
Revision Evaluation Average Hourly Rate	\$365.31	\$17.49	\$230.84	\$343.91

4.6 Evaluation Selection

As mentioned previously, for the paragraph evaluation and revision evaluation tasks depicted in Figure 4, no uniqueness restriction was imposed and all such tasks could have been completed by any participant in the experiment. The completed data for each

experiment shows that some participants who translated a paragraph or revised a translation, when presented with their own work, sometimes selected for it. Upon further inspection, we define three distinct measures for the selection of Paragraph Evaluation and Revision Evaluation tasks:

Selfish Selection – when presented with two evaluation options, one of which was created by the participant, the participant selects his or her own work as superior.

Unselfish Selection – when presented with two evaluation options, one of which was created by the participant, the participant does not select his or her own work as superior.

Disinterested Selection – when presented with two evaluation options, neither of which was created by the participant.

Table 6 shows the *Selfish*, *Unselfish*, and *Disinterested Selection* measures for each of the crowdsourced translations, broken down by evaluation task types.

Table 6 - Approximate Evaluation Selection

Approximate Evaluation Selection (in percent)	CSTLP2	CSTLP4	CSTNLP2	CSTNLP4
Selfish Selection – Total Evaluation Tasks	28%	46%	17%	27%
Selfish Selection – Paragraph Evaluation	33%	41%	7%	30%
Selfish Selection – Revision Evaluation	11%	60%	44%	20%
Unselfish Selection – Total Evaluation Tasks	17%	11%	3%	8%
Unselfish Selection – Paragraph Evaluation	19%	15%	0%	7%
Unselfish Selection – Revision Evaluation	11%	0%	11%	10%
Disinterested Selection – Total Evaluation Tasks	56%	43%	81%	65%
Disinterested Selection – Paragraph Evaluation	48%	44%	93%	63%
Disinterested Selection – Revision Evaluation	78%	40%	44%	70%

4.7 Diversity

The completed data for each experiment shows that, in some cases, a few participants completed significantly more tasks than other participants, specifically with respect to Paragraph Evaluation and Revision Evaluation tasks. In the cases where a participant did the lion’s share of work, indicating an uneven distribution of tasks completed, what measures can be used to understand how this type of diversity may have affected the quality of the translations?

Simpson's Diversity Index was selected to measure diversity for crowdsourced translation. Widely used in the field of ecology for measuring biodiversity, Simpson's Diversity Index seemed an appropriate fit for the experimental data of crowdsourced translation tasks. Here there are two measures that are appropriate; the formulas for which are shown below. The first is Equation 1 - Simpson's Index, where in the context of the crowdsourced translation experimental data, D is a measure of diversity where n equals the total number of tasks completed by a particular participant, and N equals the total number of tasks completed by all participants. This measure can be calculated for each of the lightweight peer production tasks (Paragraph Translation, Paragraph Evaluation, Revision Creation, and Revision Evaluation) and also for each experiment as a whole.

Equation 1 - Simpson's Index

$$D = \frac{\sum n (n - 1)}{N (N - 1)}$$

The second important equation is Equation 2 - Simpson's Reciprocal Index; this represents the "diversity" or number of participants once the richness and uniqueness of the experimental data is taken into account. The Reciprocal Index Value (*RIV*) derived from these two equations is useful when used together as they take into account the size (number of tasks) of each experiment, not just the number of unique participants within an experiment. In cases where a participant did the lion's share of the work, specifically with respect to evaluation tasks, the Reciprocal Index Value (*RIV*) will show a lower number than it would have had all of the tasks been completed evenly amongst all of the participants. These numbers can then be related in Phase II (Paragraph Evaluation) to the overall measures of translation quality to gain a better understanding of the effect of diversity on the interpretation of the quality of the target texts.

Equation 2 - Simpson's Reciprocal Index

$$RIV = 1 / D$$

Table 7 shows Simpson's Index Measures, which does not take into account the size of each experiment. The values for Revision Creation tasks are zero because the *Task Repetition Value* exactly matched the number of tasks that were published.

Table 7 - Simpson's Index Measures

Simpson's Index	CSTLP2	CSTLP4	CSTNLP2	CSTNLP4
Paragraph Translation - Simpson's Index	40.00%	15.15%	26.67%	18.18%
Paragraph Evaluation - Simpson's Index	14.81%	52.42%	60.40%	16.52%
Revision Creation - Simpson's Index	0.00%	0.00%	0.00%	0.00%
Revision Evaluation - Simpson's Index	61.11%	64.44%	22.22%	28.89%
Total Diversity - Simpson's Index	19.24%	37.74%	27.48%	17.78%

Table 8 shows Simpson's Reciprocal Index measures, which does take into account the size of each experiment. Here the total number of tasks that were completed by the participants is factored into the Simpson's Reciprocal Index measures.

Table 8 - Simpson's Reciprocal Index Measures

Simpson's Reciprocal Index	CSTLP2	CSTLP4	CSTNLP2	CSTNLP4
Paragraph Translation – Simpson's Reciprocal Index	2.50	6.60	3.75	5.50
Paragraph Evaluation - Simpson's Reciprocal Index	6.75	1.91	1.66	6.05
Revision Creation - Simpson's Reciprocal Index	2.00	4.00	2.00	4.00
Revision Evaluation - Simpson's Reciprocal Index	1.64	1.55	4.50	3.46
Total Diversity - Simpson's Reciprocal Index	5.20	2.65	3.64	5.62

In Table 8, the Revision Creation RIV used was artificially set to the *Task Repetition Value* for the experiment since dividing by zero is not suitable.

4.8 Approximate Translation Quality

To rigorously compare crowdsourced translation to professional and machine translation, a translation PhD student was hired to assess the quality of the ten translations of the source text (three professional translations (PT), three machine translation (MT) and four crowdsourced translations (CST) and to rank each of the ten translations in order from best to worst quality. The professional and machine translations ranked highest by the PhD student were then selected to be studied in Phase II along with the four crowdsourced translations. At no time was the translation PhD student told how any of the translations were created.

Table 9 depicts the approximate translation quality of the ten translations at the end of Phase I according to the rankings by the translation PhD student. The best quality translation was given a rank of one, while the worst quality translation was given a rank of ten. The rankings by the PhD translation student are as follows.

Table 9 - Rankings of Approximate Translation Quality

Rank 1 – PT	Rank 6 – PT
Rank 2 – CSTLP2	Rank 7 – CSTNLP4
Rank 3 – PT	Rank 8 – MT
Rank 4 – CSTLP4	Rank 9 – MT
Rank 5 – CSTNLP2	Rank 10 - MT

The translations listed in Table 9 are presented in Appendix B – Translation Creation Data as are the translations that were discarded as part of each Phase I experiment.

Chapter 5: Phase II (Translation Evaluation) - Methodology

The general research question of this thesis is the following: can a model be created that can estimate how effective crowdsourced translations are as a means of bridging the Canadian Linguistic Digital Divide? This chapter describes the major concepts used in the evaluation of the crowdsourced translations and how each concept has been operationalized for the purposes of this study. Within are described both the hypotheses as well as the methodologies that were used to validate the hypotheses. The results of the hypothesis testing and analysis are described in detail in Chapter 6: Phase II (Translation Evaluation) – Analysis.

5.1 Overall Translation Quality

Clear qualitative indications of the target text’s performance can be made by assessing how native speakers in the language of the target text respond after reading the target text. If the quality is acceptable, we posit that a native speaker would likely have positive impressions of the target text. For example, a strong Canadian French translation may increase the likelihood that a Canadian Francophone would contact the Black Camel Café in French. Likewise, if the translation were weak, the same individual might be much less likely to trust the content of the target text, perhaps even being deterred from making purchases of products or services (DePalma, Sargent, & Beninatto, 2006) (Jorgensen, 2009). The concept of Overall Translation Quality within both the source text and target texts has the following operational definitions:

- The quality of this text is acceptable for a Canadian Francophone website.
- Most Canadian Francophones would not be able to tell that this text is a translation.
- Most Canadian Francophones would feel confident trusting the content of this translation.
- Most Canadian Francophone readers of this text would feel confident contacting the company in French.
- Most Canadian Francophone readers of this text would feel confident purchasing products or services from this company.

5.2 Task Repetition

Within the context of this study, the concept of Task Repetition represents unique and repeated paragraph translations or unique and repeated revision creations within a crowdsourced translation experiment. For example, a Task Repetition Value of four indicates that for each paragraph in the source text, four unique translations of each paragraph were created and four unique revisions of the rough draft were created.

5.3 Knowledge of Translation Methods

The concept of Knowledge of Translation Methods is represented by measures of the ways in which participants were biased after being made aware of how the crowdsourced translations were created.

5.4 General Translation Errors

The concept of General Translation Errors is largely defined by the work of House (1997), in which she defines the following qualitative dimensions for translation:

- Dimensions of Language User
 - Text producer's geographical origin
 - Text producer's social class
 - The era in which the source text was written (temporal provenance)
- Dimensions of Language Use
 - Medium
 - Simple – Written to be read
 - Complex – Written to be spoken
 - Participation
 - Simple – Monologue or Dialogue
 - Complex – Elicitation of participation
 - Social Role Relationship – Peer (symmetrical) or Authority (asymmetrical)
 - Social Attitude – Social distance to audience (frozen, formal, consultative, casual, or intimate)
 - Text producer's occupation or profession (province)

In this study, the translation quality model proposed by House (1997) has been loosely interpreted in order to operationalize concepts for the measurement of translation quality. For the purpose of this study, in what House (1997) refers to as Dimensions of Language User are herein referred to as the sub-concept of Author Errors, while the Dimensions of Language Use are herein referred to as the sub-concept of Usage Errors. House (1997) further suggests that the concept of Overt Errors should inform the quality of a target text. Overt Errors represent errors in language mechanics, such as grammar, spelling, punctuation, and large semantic mistakes. Therefore, for the purposes of this study, the concept of Translation Quality is represented by the sub-concepts of Author Errors, Usage Errors, and Overt Errors. The loose interpretation of these concepts for translation quality caters to the characteristics of a recipient evaluation (Bowker, 2009). Since the evaluators of these translations included non-language professionals, it was necessary that translation quality be measured in a way that could be understood by participants who lacked translation education, translation training, translation certification and professional translation work experience.

5.5 Overt Errors

The concept of Overt Errors within both the source and target texts has the following operational definitions:

- This text uses correct terminology.
- This text uses correct grammar.
- This text has an understandable meaning.

5.6 Usage Errors

The concept of Usage Errors within both the source and target texts has the following operational definitions:

- This text is written in a style that calls the reader to action.
- This text is written for an educated middle class audience.
- This text is written in the style of a professional blogger.
- This text is written in a style that is intended to be read instead of a style that is intended to be spoken.

- This text is written in a style that is Frozen, Formal, Consultative, Casual, or intimate.

5.7 Author Errors

The concept of Author Errors within both the source and target texts has the following operational definitions:

- This text appears to be written by a fluent Canadian (Anglophone/Francophone).
- This text appears to be written by someone with a university education.
- This text is written using language that was appropriate in 2010.

5.8 Contextual Translation Errors

House (1997) suggests that the performance of the target text must be understood within the context of the source text. The differential qualities between the source and target texts for the sub-concepts of Overt Errors, Usage Errors, and Author Errors must therefore necessarily be measured. Should a translation improve on the errors measured in the source text, these improvements should inform how the performance of the target text's quality is understood. To these ends, the difference between the quality of the source text and each target text is herein prefaced with the term Contextual, meaning that the performances of the target texts are placed within the context of the performance of the source text. Therefore, these concepts are described as Contextual Overt Translation Errors, Contextual Usage Translation Errors, and Contextual Author Translation Errors, whereby the measurement of the source text is subtracted from the measurement of the target text (e.g., Contextual Overt Translation Errors = Target Text Overt Errors - Source Text Overt Errors).

5.9 Translation Diversity

The concept of *Translation Diversity*, as represented by the values in Table 8, takes into account the size of each experiment. Here, the total number of tasks that were completed by the participants is factored into the Simpson's Reciprocal Index measures. This table describes the diversity within the Paragraph Translation process, the Paragraph Evaluation process, the Revision Creation process, the Revision Evaluation process, and as a measure for the entire Collaborative Online Language Transfer process for each of the four experiments (CSTLP2, CSTLP4, CSTNLP2, and CSTNLP4). The indicators, TDPT –

Translation Diversity Paragraph Translation, TDPE – Translation Diversity Paragraph Evaluation, TDRC – Translation Diversity Revision Creation, and TDRE – Translation Diversity Revision Evaluation, are used.

5.10 Machine Translation

The concept of Machine Translation is represented using the top-ranked machine translation selected as a result of the rankings provided in Table 9 – Rankings of Approximate Translation Quality.

5.11 Professional Translation

The concept of Professional Translation is represented using the top-ranked professional translation selected as a result of the rankings provided in Table 9 – Rankings of Approximate Translation Quality.

5.12 Crowdsourced Translation

The concept of Crowdsourced Translation is represented by the four experiments conducted in Phase I (Translation Creation); c.f. Table 9 – Rankings of Approximate Translation Quality.

5.13 Crowdsourced Translation Acceptance

The concept of Crowdsourced Translation Acceptance is represented as a set of measurements designed to gauge how participants regarded the concept of crowdsourced translation as a means of bridging the Canadian Linguistic Digital Divide. The following operational definitions were designed for this purpose:

- I prefer crowdsourced translation more than machine translation.
- I prefer crowdsourced translation more than professional translation.
- Crowdsourced translation should be performed by non-language professionals.
- Crowdsourced translation should be performed by language professionals.
- Crowdsourced translation should be used to translate website content in Canada.

5.14 Research Hypotheses

From the research questions and major operational concepts, the following research hypotheses are derived.

H₁: *An increase in Task Repetition will have a positive effect on the Overall Translation Quality of crowdsourced translation.*

Based on the work of Eagle (2009) and von Ahn (2005), it is expected that the more a task is repeated among unique individuals, the higher the quality of the result of that task. While Eagle (2009) observed a positive relationship with respect to the translation of terminologies using crowdsourcing, this hypothesis is extended in order to test whether a positive relationship exists with respect to the translation of paragraphs using crowdsourcing. While terminologies are usually one or two words with some level of generally accepted meaning among a particular group (e.g. lawyers, accountants, doctors, etc), a paragraph is longer, more descriptive, and open to wider variations in interpretation.

H₂: *The Knowledge of Translation Methods will have a negative effect on the Overall Translation Quality of crowdsourced translation.*

Based on the work of Bowker (2008), it was observed that some forms of translation were rejected simply based upon how respondents viewed the way in which translations were created. This hypothesis aims to determine if crowdsourced translation will be similarly rejected once participants become aware of how these translations were created. Understanding this type of bias is important in assessing how effective this form of translation may be at bridging the Canadian Linguistic Digital Divide. Should people reject crowdsourced translation out of hand, there would likely be less potential benefit for Canadians.

H₃: *An increase in Translation Diversity will have a positive effect on the Overall Translation Quality of crowdsourced translation.*

Hong and Page (2004) suggest that diversity trumps ability and that a group of non-experts can outperform a group of experts. Essentially, the argument is that the more diverse a group is, the more likely they can produce equivalent or better quality work than that of experts. This hypothesis is proposed to test the validity of this idea with respect to crowdsourced translation.

H₄: *When considering Overall Translation Quality, bilingual Canadians will select machine translation over crowdsourced translation.*

OCOL (2005) indicates that machine translation is of generally poor quality. This null hypothesis tests if there is a significant difference between the general acceptance of machine translation quality and the general acceptance of crowdsourced translation quality.

H₅: *When considering Overall Translation Quality, bilingual Canadians will select professional translation over crowdsourced translation.*

OCOL (2005) indicates that professional translation is widely accepted currently as the method by which content should be translated with respect to the Canadian Linguistic Digital Divide. This null hypothesis tests if there is a significant difference between the general acceptance of professional translation quality and the general acceptance of crowdsourced translation quality.

H₆: *The absence of Contextual Overt Translation Errors will have a positive effect on the Overall Translation Quality of crowdsourced translation.*

House (1997) suggests that when overt translation errors are absent, there is a positive effect on the quality of a translation. This hypothesis is proposed to test whether there is a significant correlation between the absence of contextual overt translation errors and the overall quality of crowdsourced translation. For example, these errors are related to terminology, grammar, and meaning. How did the crowdsourcing translators address these?

H₇: *The absence of Contextual Usage Translation Errors will have a positive effect on the Overall Translation Quality of crowdsourced translation.*

House (1997) suggests that when usage translation errors are absent, there is a positive effect on the quality of a translation. This hypothesis is proposed to test if there is a significant correlation between the absence of contextual usage translation errors and the overall quality of crowdsourced translation. For example, these are errors that typically involve stylistic elements of the source text not being carried through to the target text. In essence, how did the crowdsourcing translators convey the same style as the source text?

H₈: *The absence of Contextual Author Translation Errors will have a positive effect on the Overall Translation Quality of crowdsourced translation.*

House (1997) suggests that when author translation errors are absent there is a positive effect on the quality of a translation. This hypothesis is proposed to test if there is a significant correlation between the absence of contextual author translation errors and the overall quality of crowdsourced translation. For example, these are errors related to the source text author's language use, social station, and the era of the language depicted in the text. In essence, how did the crowdsourcing translators capture these types of elements?

5.15 Hypothesis Testing

Most of the hypotheses tested in this study used Smart PLS 2.0 (C. Ringle, Wende, & Will, 2005). Smart PLS is used when we are trying to understand the relationship between two conceptually different things, such as cognitive diversity and overall translation quality. Smart PLS essentially allows us to understand how significant and predictive the relationship is. Confidence intervals, on the other hand, are useful in cases where we are trying to compare two conceptually similar things, such as the performance of a machine translation and the performance of a crowdsourced translation. As such, for the hypotheses that deal with comparing translation methods or assessing the effect of the Knowledge of Translation Methods (KTM), confidence intervals were used.

5.16 Partial Least Squares – Path Modeling with Smart PLS 2.0

Partial Least Squares Path Modeling (PLS-PM) is known as a “soft modeling” approach that has no strong parametric distribution requirements; it can perform well with respect to smaller sample sizes and mixed measurement scales (Vinzi, Trinchera, & Amato, 2010); (Hulland, Ryan, & Rayner, 2010). PLS-PM also allows for both reflective and formative measurement items, whereas covariance-based structural equation modeling allows for only reflective measurements (Chin, 2010b); (Albers, 2010). Compared to other multivariate methods, especially covariance-based structural equation modeling, PLS-PM is appropriate for prediction, theory building, confirmatory analysis and hypothesis testing (Barroso et al., 2010). PLS-PM is also particularly appropriate in research contexts where there are not yet strong theoretical underpinnings and little previous knowledge (Barroso, Carrión, & Roldán, 2010; Cudeck & Henley, 2003). For those who are not familiar with structural equation modeling techniques using Smart PLS, it is important to note that the software does not report any overall measures for Goodness of Fit, such as chi square, as you might expect from CB-SEM packages like LISREL or AMOS (Henseler, Ringle & Sinkovics, 2009). PLS Path Modeling can be used to model the complex relationships between multiple latent and manifest variables (Vinzi, Chin, Hensler, & Wang, 2010) with relatively constant errors expressed with small sample sizes while supporting higher measurement indicators than covariance-based structural equation modeling techniques (Hulland et al., 2010).

Smart PLS 2.0 generates many useful measures for analyzing data, such as t-values, Average Variance Extracted (AVE), Composite Reliability, R^2 , Cronbach’s Alpha, Communality and Redundancy. It also produces manifest variable scores, latent variable

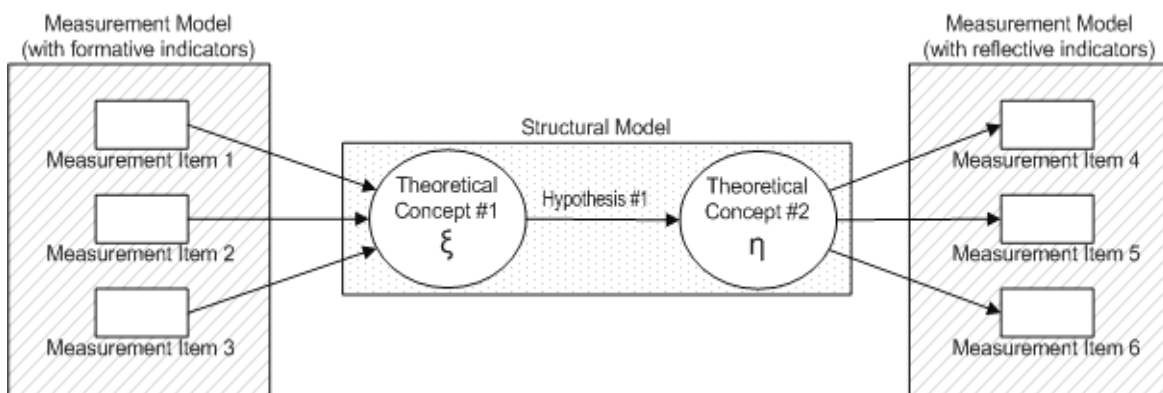
correlations, inner model indicator cross loadings, outer loadings, outer model weights, path coefficients, latent variable scores, manifest variable scores, outer weights, and index values of latent variables. Of particular importance is the bootstrapping resampling method (Temme et al., 2010) which is used to obtain significance values. Bootstrapping works by randomly sampling the data being analyzed, and through a large number of permutations, very reliable t-value statistics are generated. Bootstrapping provides a high level of confidence in the significance measurements of the data.

In light of these qualities of the PLS-PM and the fact that the current study is exploratory nature, PLS-PM is considered an appropriate technique for the analysis. Moreover, the ability of PLS-PM to make predictive statements (Chin, 2010a) about how crowdsourced translation may be accepted as a means of bridging the Canadian Linguistic Digital Divide aligns well with the goals of this study.

5.17 Structural Equation Modeling using PLS-PM

In Structural Equation Modeling (SEM) using Partial Least Squares Path Modeling (PLS-PM), the model has two components, namely, the measurement model and the structural model. The measurement model contains the raw data that was collected for this study; typically this was done using five-point Likert scale questionnaires. Conversely, it is the structural that is used to validate the hypotheses in PLS-PM. In Figure 7, we see a generic Structural Equation Model using PLS-PM.

Figure 7 - Structural Equation Modeling using PLS-PM



Exogenous latent variables (ξ) and endogenous latent variables (η) are used in the illustration of structural models for PLS-PM. Exogenous latent variables are sometimes also referred to as independent variables (X). In PLS-PM, the exogenous latent variable (ξ) has

an effect on an endogenous latent variable (η). Endogenous latent variables (η) are sometimes also referred to as dependent variables (Y).

Generally, a two-step process is used to analyze data and interpret the results using Smart PLS 2.0. The first step involves establishing the reliability and validity of the measurement model. Once a reliable and valid measurement model is established, then the next step involves evaluating the structural model (Fornell & Larcker, 1981). The reliability of the measurement models is assessed to understand how much random error is observed. The validity of the measurement models is assessed to understand how much systemic error is observed. Should the measures for statistical significance, reliability, and validity of the measurement models all be within acceptable thresholds, we can then assess the structural model. To assess the structural model, we consider the direction of the correlated hypothesis, which can be either positive or negative, and how much variance can be explained by the correlated hypothesis.

Table 10 below lists the measures generally used to assess the measurement and structural models, their meanings, and the acceptable thresholds for statistical significance.

Table 10 - Acceptable Thresholds for Interpreting PLS-PM

Measurement	Description	Acceptable Thresholds
T-Value	A measure of statistical significance.	$\geq \sim 1.97$ (Abramowitz, 1965)
P-Value	A measure of statistical significance.	≤ 0.05 (Abramowitz, 1965)
Cronbach's α	A measure of reliability (random error).	≥ 0.7 (Nunnally & Bernstein, 1994).
Composite Reliability	A measure of reliability (random error).	≥ 0.7 (Nunnally & Bernstein, 1994).
AVE	A measure of validity (systemic error).	≥ 0.7 (Fornell & Larcker, 1981)
R^2	A measure of the variance explained.	<ul style="list-style-type: none"> • 0.67 [substantial] • 0.33 [moderate] • 0.19 [weak] (Chin, 1998).

In Figure 7, the left-most measurement model illustrates a set of formative measurement indicators that are represented by arrows pointing toward the first theoretical concept (latent variable). Constructs with formative measurement indicators in PLS-PM allow researchers to gain an understanding of the predictive qualities of the construct (Albers, 2010). When using formative indicators, researchers should avoid measurement items that are correlated (Albers, 2010). Also, adding or dropping formative indicators to a construct may alter the meaning of that construct, so care must be exercised in this regard (Albers, 2010). Moreover, the methods for evaluating the validity of constructs with formative indicators are different from those for evaluating the validity of constructs with reflective indicators (Albers, 2010).

"In the formative model, each manifest variable or each sub-block of manifest variables represents a different dimension of the underlying concept. Therefore, unlike the reflective model, the formative model does not assume monogeneity nor unidimensionality of the block. The latent variable is defined as a linear combination of the corresponding manifest variables, thus each manifest variable is an exogenous variable in the measurement model. These indicators need not to covary: changes in one indicator do not imply changes in the others and internal consistency is no more an issue." (Vinzi, Trinchera et al., 2010, p. 51)

In Figure 7, the right-most measurement model illustrates a set of reflective measurement indicators that are represented by arrows pointing away from the second theoretical concept (latent variable). With reflective measurements, an analogy can be offered that suggests that constructs with reflective indicators are similar to light sources such as the sun. While it may not be possible to directly measure the sun, we can instead measure the amount of the sun's light that is reflected on our surroundings. As such, if the sunlight is stronger or weaker, the amount of reflected light on our surroundings (no matter how different individual measurements may be), should also change accordingly. Reflective measurements, likewise, must necessarily represent interchangeable measurements which could be added or removed at random, without altering the meaning of the associated construct (Albers, 2010; Dijkstra, 2010).

Most of the measured indicators in this study were defined formatively in order to gain a better understanding of the predictive power of the measurement models. Reflective measurement indicators were used only in cases where the measurement items represented general and replaceable measurements. In the case of hypotheses that were tested using confidence intervals, the resultant data was calculated using the CONFIDENCE function in Microsoft Office Excel 2007 with an alpha value of 0.05.

5.18 Study Participants

Participants in both phases of the study were classified into two groups – Language Professionals (LP) and Non-Language Professionals (NLP). Language Professionals, for the purposes of this study, were defined as individuals who are fluent in both Canadian English and Canadian French and who also possessed translation education, translation training, or professional translation work experience. Non-Language Professionals were defined as

individuals who are fluent in both Canadian English and Canadian French, but who lacked translation education, translation training, and professional translation work experience.

Most of the study participants were recruited from the National Capital Region of Canada. The study participants were asked to evaluate the six translations selected from Phase I (Translation Creation): PT, CSTLP2, CSTLP4, CSTNLP2, CSTNLP4, and MT. Thirteen of the participants were Language Professionals and nineteen were Non-Language Professionals. Table 11 shows the aggregated levels of education for both Phase II participant groups.

Table 11 - Phase II Participant Education

Levels of Education	Language Professionals	Non-Language Professionals
Some college/university	Two participants (15.4%)	Seven participants (36.8%)
Bachelor's degree	Seven participants (53.9%)	Six participants (31.6%)
Master's degree	Two participants (15.4%)	Six participants (31.6%)
Doctoral degree	Two participants (15.4%)	No participants

In general, the participants in Phase II were well-educated bilingual individuals. All participants had completed some years of college/university; the majority had completed at least one university degree. Thus the sample is biased towards educated bilingual Canadians.

Table 12 shows the aggregated number of years of professional translation work experience for the Language Professional participants in Phase II.

Table 12 - Phase II Professional Translation Work Experience

Years of Professional Translation Work Experience	
No translation work experience	Two participants (15.4%)
Less than one year	Four participants (30.8%)
Between one year and two years	One participant (7.7%)
Between two years and three years	Two participants (15.4%)
Between three years and four years	Two participants (15.4%)
Between four years and five years	One participant (7.7%)
More than five years	One participant (7.7%)

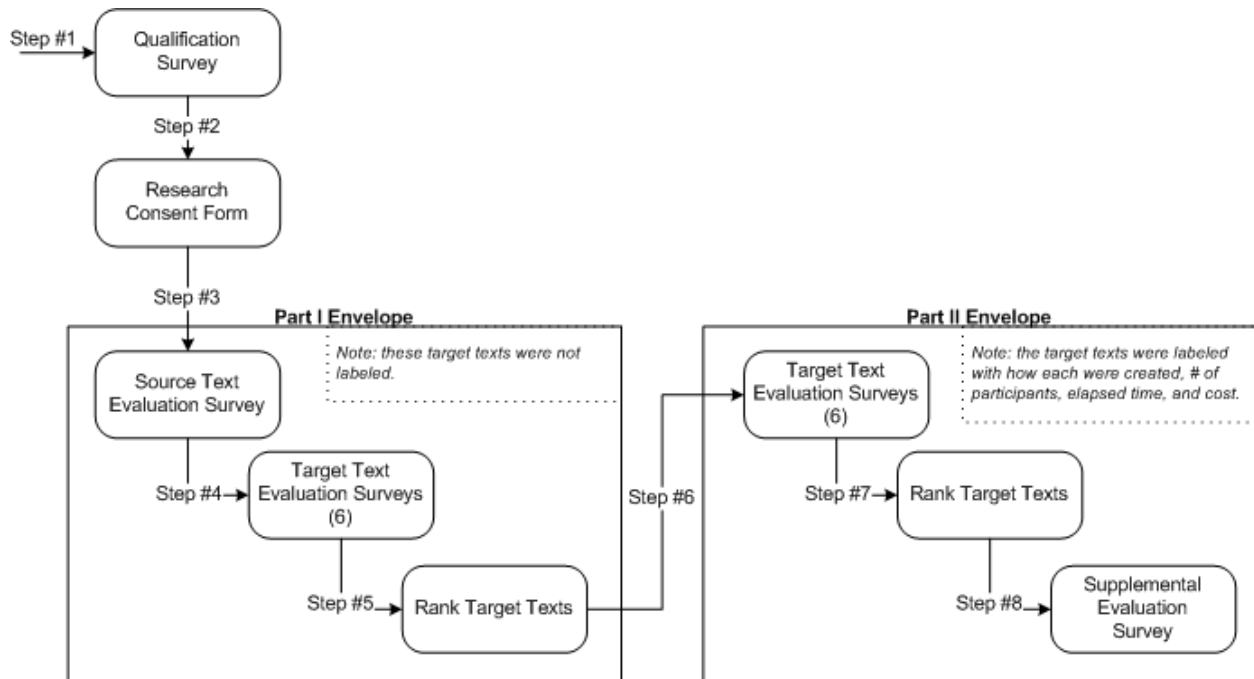
Most of the Language Professional participants in Phase II had professional translation work experience. Most of these participants had completed degrees in Translation and a few had received partial translation education before completing degrees in other fields. None of these participants had any professional translation certifications.

Each research participant was asked to first complete a qualification survey to determine if he or she was qualified either as a Language Professional or Non-Language

Professional; all survey materials are provided in Appendix C – Phase II (Translation Evaluation Data). Next, participants were asked to complete a research consent form and upon completing the form, each was offered forty Canadian dollars for participating in the study. The participants were then provided one large envelope containing two smaller envelopes marked Part I and Part II and asked to return the completed material within two week’s time.

Figure 8 illustrates the work flow process for the phase II.

Figure 8 - Phase II Participant Workflow



The content of the envelope marked Part I was shown to each participant; it contained instructions, the source text, six unlabeled target texts, a source text evaluation survey, and six target text evaluation surveys. The participants were first asked to read all of the material, to complete the source text evaluation survey, and then staple the source text evaluation survey to the source text. Next, the participants were asked to complete a target text evaluation survey for each of the target texts and to staple each completed target text evaluation survey to its associated target text. After completing all of the target text evaluation surveys, the participants were asked to rank the translations from best to worst quality. Finally, the participants were asked to return all of the evaluation survey materials to the envelope marked Part I and to seal the envelope marked Part I before

proceeding to open the envelope marked Part II. The methodological purpose of the two envelopes was to gather data related to **H₂**.

Once the Part I envelope had been sealed and the Part II envelope had been opened, the participants again repeated the process of completing target text evaluation surveys and ranking the stapled target texts. In Part II, each target text had additional information at the top of its page describing the total cost of the translation, the duration of the translation process, the number of people involved in creating the crowdsourced translation, and the participant group of the crowdsourced translation (Language Professional or Non-Language Professional). In the cases of the professional and machine translations, some additional information was provided such as the name of the machine translation website that was used and the translation certifications of the professional translation company. Next, the participants were asked to complete a supplemental evaluation survey, which served as a means of measuring the acceptance of the idea of crowdsourced translation. Finally, the participants were asked to return all of the evaluation survey material to the envelope marked Part II, seal it, and then indicate via email that they had completed their participation; c.f. Appendix C – Phase II (Translation Evaluation Data).

5.19 Instrument and Measurement

With the exception of the qualification survey, the instrument used in Phase II (Translation Evaluation) consisted of a series of five-point Likert scale questions, where 1 = strong disagreement with the statement and 5= strong agreement with the statement. The completed data from Part I created the 128 samples (four crowdsourced translations multiplied by 32 participants) that were used in the analysis and validation of the hypotheses (**H₁**, **H₃**, **H₄**, **H₅**, **H₆**, **H₇** and **H₈**). Data collected from Part II created another 128 samples which were used to calculate the difference between the two sets of data for the validation of **H₂**.

Included within the 128 samples were calculated measures of how the professional translation and machine translation performed relative to each crowdsourced translation, i.e., the PT measurement minus the CST measurement and the MT measurement minus the CST measurement. The ranking information, originally collected with a value of one being highest and six being lowest, was recalculated with an inverted order so that the highest ranked translation also had the highest numerical value; specifically a value of seven minus the measured rank.

In a few cases participants did not complete each measurement in an evaluation survey. In these cases, an imputed value calculated as the average of the data collected from the same translation (i.e. CSTLP2) from across all of the available samples was used (Allan & Seaman, 2010). When done in this way, the calculated average often was not an integer value. Smart PLS standardizes all measurement values before it reports its results, meaning that the mean of the measured values are set at zero by the software prior to the analysis. In the case of the use of confidence intervals for hypothesis testing (**H₂**, **H₄**, and **H₅**), in order to obtain integer counts, values that had been imputed using average measurements from the same translation were rounded to the nearest integer.

Chapter 6: Phase II (Translation Evaluation) - Analysis

6.1 Overview

This chapter provides a detailed analysis for each of the hypotheses. For the hypotheses that were tested using confidence intervals (**H₂**, **H₄**, and **H₅**), the resultant data was calculated using the CONFIDENCE function in Microsoft Office Excel 2007 with an alpha value of 0.05. For the hypotheses that were tested using Smart PLS 2.0 (**H₁**, **H₃**, **H₆**, **H₇**, and **H₈**), please refer to Table 10 - Acceptable Thresholds for Interpreting PLS-PM.

In some cases, sub-hypotheses were generated using different sets of the entire 128 samples. The sub-hypotheses were analyzed to examine whether there were statistical differences between translation creators, between translation evaluators, and between experiment types (task repetition value two and task repetition value four). The following notation is herein used to denote cases where a subset of the entire 128-sample population is being used to test a hypothesis, where the # represents the hypothesis being tested:

- All Phase II Samples:
 - **H_{#A}** - A hypothesis being tested with all 128 samples (thirty-two participants multiplied by four crowdsourced translations)
- Differences between Translation Creators:
 - **H_{#B}** - (TCLP) A hypothesis being tested using only the crowdsourced translations that were created by Language Professionals in Phase I; and
 - **H_{#C}** - (TCNLP) A hypothesis being tested using only the crowdsourced translations that were created by Non-Language Professionals in Phase I.
- Differences between Translation Evaluators:
 - **H_{#D}** - (TELP) A hypothesis being tested using only the results of Language Professional translation evaluators in Phase II; and
 - **H_{#E}** - (TENLP) A hypothesis being tested using only the results of Non-Language Professional translation evaluators in Phase II.

- Differences between Experiment Types:
 - **H_{#F}** - (TRV2) A hypothesis being tested using only the crowdsourced translations that were created using a Task Repetition Value of two; and
 - **H_{#G}** - (TRV4) A hypothesis being tested using only the crowdsourced translations that were created using a Task Repetition Value of four.

Since no strong theoretical knowledge exists for a Collaborative Online Language Transfer system, each hypothesis is first tested in isolation instead of combining multiple hypotheses into a single PLS path model. At the end of this chapter, a more comprehensive PLS path model that includes multiple latent variables is defined and tested. Based upon the findings of this thesis, further research could be conducted to develop a more comprehensive and predictive PLS path model.

In the following sections are diagrams for the hypotheses that are validated using PLS path modeling. In this diagram (PLS notation), ovals indicate latent variables (theoretical concepts) and rectangles represent measurement indicators. The direction of the arrows between a latent variable and its measurement indicators shows whether the measurement indicators are formative or reflective. Arrows leading from the measurement indicators to the latent variable are formative; arrows leading from the latent variable to the measurement indicators are reflective, cf. Figure 7. Tables that describe confidence interval values are shown where applicable.

6.2 Hypothesis (H₁)

An increase in Task Repetition will have a positive effect on the Overall Translation Quality of crowdsourced translation.

Based on the work of Eagle (2009) and von Ahn (2005), it is expected that the more a task is repeated among unique individuals, the more likely there will be an increase in the quality of the result of that task. While Eagle (2009) observed a positive relationship with respect to the translation of terminology using crowdsourcing, this hypothesis is extended in order to test if a positive relationship exists with respect to the translation of paragraphs using crowdsourcing. This hypothesis is tested as **H_{1A}**, with $\alpha \leq 0.05$ level which includes all 128 samples. Table 13 shows the operationalized measurement items that were used to test **H_{1A}**.

Table 13 - H_{1A} (Structural and Measurement Model)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Task Repetition	Formative	TRV	Task Repetition Value (2 or 4)
Overall Translation Quality	Reflective	AQ1	The quality of this text is acceptable for a Canadian Francophone website.
		OTQ11	Most Canadian Francophones would not be able to tell that this text is a translation.
		OTQ21	Most Canadian Francophones would feel confident trusting the content of this translation.
		OTQ31	Most Canadian Francophone readers of this text would feel confident contacting the company in French.
		OTQ41	Most Canadian Francophone readers of this text would feel confident purchasing products or services from this company.

Table 14 shows the bootstrapping results for **H_{1A}**.

Table 14 - H_{1A} (Bootstrapping Results)

BOOTSTRAPPING (128 samples with 500 iterations/degrees of freedom)		T-STATISTIC VALUES
Structural Model	Task Repetition → Overall Translation Quality	3.686792 (significant @ 0.0005)
Measurement Model	AQ1	35.232309
	OTQ11	26.392718
	OTQ21	71.336185
	OTQ31	35.691735
	OTQ41	33.718549

In **H_{1A}**, there is a statistically significant relationship between *Task Repetition* and *Overall Translation Quality* with a two-tailed P-value of 0.0005 (Abramowitz, 1965); (Graphpad,

2011). The bootstrapping algorithm shows that the measurement model is statistically significant. Figure 9 and Table 15 show the PLS algorithm results for for **H_{1A}**.

Figure 9 - H_{1A} (PLS Algorithm Results)

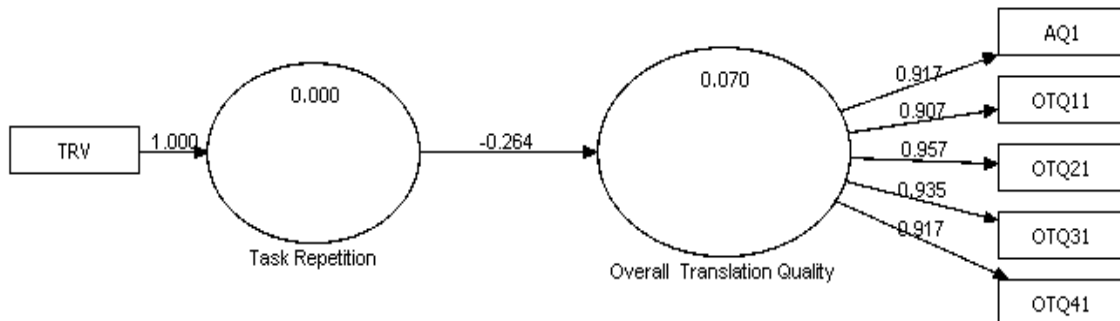


Table 15 - H_{1A} (PLS Algorithm Results)

PLS ALGORITHM	VALUES
Path Coefficient (Task Repetition → Overall Translation Quality)	-0.2644
Cronbach's Alpha	0.9592
Composite Reliability	0.9681
Average Variance Extracted (AVE) for Overall Translation Quality	0.8587
R ²	0.0699

Cronbach's Alpha and Composite Reliability values are above 0.7, the commonly used threshold for reliability (Nunnally & Bernstein, 1994). The Average Variance Extracted (AVE) of 0.85 exceeds the 0.7 rule of thumb for validity (Fornell & Larcker, 1981). The R² value explains about 7% of the variance between *Task Repetition* and *Overall Translation Quality* while the path coefficient shows a negative correlation between these two concepts. Therefore, hypothesis **H_{1A}**, which considers all 128 samples, is rejected with $\alpha \leq 0.0005$.

On the face of it, these results do not seem to support the findings of Eagle (2009) who argues that an increase in task repetition is an indicator for increased task quality. In the case of Eagle (2009), mobile crowdsourcing was used to translate terminology for mobile devices. This finding suggests that paragraph translation tasks and revision creation tasks are more complex than are terminology translation tasks. This finding suggests that simply adding more unique participants does not guarantee a better quality crowdsourced translation.

When referring to Figure 4, it is important to remember that before a target text translation was produced, many of the participants' diverse, aggregated, decentralized and independent judgments (Surowiecki, 2005) were discarded (Czapski, 2008). In the case of a crowdsourced translation using a *Task Repetition* value of four, it means that three unique

translations of each paragraph and three unique revisions of the rough draft translation did not make the cut. This result raises the following questions. How much of an effect did *Task Repetition* have on the differential quality of the rough draft translation and the target text? Stated differently, do increases in *Task Repetition* account for any increases in quality during the Collaborative Online Language Transfer process? These questions suggest an avenue of investigation for future research related to crowdsourced translation using a Collaborative Online Language Transfer System. Perhaps by looking in this direction, more knowledge can be gained which could help to compensate or improve the slight negative correlation observed in **H_{1A}**. The rough draft translations and discarded revisions for each experiment are available in Appendix B.

6.3 Hypothesis (H₂)

The Knowledge of Translation Methods will have a negative effect on the Overall Translation Quality of crowdsourced translation.

Based on the work of Bowker (2008), it was observed that some forms of translation were rejected simply based on how respondents viewed the way in which translations were created. This hypothesis is extended in order to determine if such an effect is present in these methods of crowdsourced translation. This hypothesis is tested as **H_{2A}**, with $\alpha \leq 0.05$ level. It includes all 128 samples, using the difference between the Part II results and the Part I results, cf. Figure 8 - Phase II Participant Workflow.

If the first time the question of acceptable quality was raised (ACQ1), the participant provided a value of one and the second time (ACQ2) they provided a value of three, ACQ4 should equal a value of two, indicating that the *Knowledge of Translation Methods* (KTM) for the Acceptable Quality indicator had a positive change of two points. Table 16 shows the operationalized measurement items that were used to test **H_{2A}**.

Table 16 – H_{2A} (Measurement Indicators)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Overall Translation Quality	Change	AQ4	The quality of this text is acceptable for a Canadian Francophone website.
		OTQ14	Most Canadian Francophones would not be able to tell that this text is a translation.
		OTQ24	Most Canadian Francophones would feel confident trusting the content of this translation.
		OTQ34	Most Canadian Francophone readers of this text would feel confident contacting the company in French.
		OTQ44	Most Canadian Francophone readers of this text would feel confident purchasing products or services from this company.

Table 17 shows how each of the five indicators for *Overall Translation Quality* changed once the Phase II participants were aware of how the crowdsourced translations were created. Referring to Figure 14, the *Knowledge of Translation Methods* were revealed in the second envelope by including information about the elapsed time of each experiment, the total cost of the translation, the number of unique participants, and also the participant group responsible for creating the translation (Language Professionals vs. Non-Language Professionals).

Table 17 - Changes in OTQ after Knowledge of Translation Methods (KTM)

Changes in OTQ after Knowledge of Translation Methods (KTM)					
Indicator	Description	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4
ACQ4	The quality of this text is acceptable for a Canadian Francophone website.	2	5	12	15
OTQ14	Most Canadian Francophones would not be able to tell that this text is a translation.	4	-4	8	5
OTQ24	Most Canadian Francophones would feel confident trusting the content of this translation.	-5	-3	7	7
OTQ34	Most Canadian Francophone readers of this text would feel confident contacting the company in French.	-3	6	5	10
OTQ44	Most Canadian Francophone readers of this text would feel confident purchasing products or services from this company.	-8	-3	2	5
Total		-10	1	34	42

Table 17 shows that, in three out of four of the crowdsourced translation experiments, there was a positive change in how *Overall Translation Quality* was assessed. By far, the strongest positive changes were observed in the crowdsourced translations that used a *Task Repetition Value* of four. The crowdsourced translation experiments using a *Task Repetition Value* of two showed negative changes on multiple indicators for *Overall Translation Quality*, whereas the other experiments showed only positive changes. Here we also see that the crowdsourced translation experiments that were performed by non-language professionals did better than their language professional counterparts after the *Knowledge of Translation Methods (KTM)* were introduced. For the sake of clarity, the *Knowledge of Translation Methods* included on the part II target texts included the duration of the experiment, the cost, the type and number of participants (language professionals or non-language professionals), cf. Figure 14. Taken together, these results indicate that, for crowdsourced translation, there is a positive general effect on *Overall Translation Quality* once people know how the crowdsourced translations are created. This positive general effect seems to be enhanced by increases in *Task Repetition*. The observations about translation evaluation bias made by Bowker (2008) with respect to maximally edited machine translation do not appear to be consistent with biases related to the evaluation of crowdsourced translation enhanced by *Task Repetition*.

Table 18 aggregates the changes measured after *Knowledge of Translation Methods* into positive, no-change, and negative categories.

Table 18 - H_{2A} (Aggregated Changes in Overall Translation Quality after KTM)

Aggregated Changes in Overall Translation Quality after KTM									
Change	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	CSTLP2 %	CSTNLP2 %	CSTLP4 %	CSTNLP4 %	TOTAL
Positive	25	38	43	53	16%	24%	27%	33%	159
No Change	100	85	99	78	63%	53%	62%	49%	362
Negative	35	37	18	29	22%	23%	11%	18%	119
TOTAL	160	160	160	160	100%	100%	100%	100%	640
Confidence Interval using $\alpha = 0.05$					6.3099	4.2500	6.4268	3.7965	

The confidence intervals shown in Table 18 confirm the results presented in Table 17. Here we see that in general, once the *Knowledge of Translation Methods* were presented, the participants viewed the *Overall Translation Quality* of the crowdsourced translations in a more positive light, specifically with respect to the crowdsourced translations enhanced by *Task Repetition*. Here we see a higher positive view of the *Overall Translation Quality* of the crowdsourced translations performed by non-language professionals than of the crowdsourced translations performed by language professionals. Therefore, **H_{2A}**, which includes all 128 samples, is rejected with $\alpha \leq 0.05$.

6.4 Hypothesis (H₃)

An increase in Translation Diversity will have a positive effect on the Overall Translation Quality of crowdsourced translation.

Hong and Page (2004) suggest that diversity trumps ability and that a group of non-experts can outperform a group of experts. This hypothesis is proposed to test the validity of this idea with respect to crowdsourced translation. The measures of diversity used to test this hypothesis are shown in Table 8. This hypothesis is tested as **H_{3A}**, with $\alpha \leq 0.05$ level which includes all 128 samples. Table 19 shows the operationalized measurement items that were used to test **H_{3A}**.

Table 19 – H_{3A} (Structural and Measurement Model)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Translation Diversity	Formative	TDPT	Simpson's Reciprocal Index Value (Paragraph Translation)
		TDPE	Simpson's Reciprocal Index Value (Paragraph Evaluation)
		TDRC	Simpson's Reciprocal Index Value (Revision Creation)
		TDRE	Simpson's Reciprocal Index Value (Revision Evaluation)
Overall Translation Quality	Reflective	AQ1	The quality of this text is acceptable for a Canadian Francophone website.
		OTQ11	Most Canadian Francophones would not be able to tell that this text is a translation.
		OTQ21	Most Canadian Francophones would feel confident trusting the content of this translation.
		OTQ31	Most Canadian Francophone readers of this text would feel confident contacting the company in French.
		OTQ41	Most Canadian Francophone readers of this text would feel confident purchasing products or services from this company.

Table 20 shows the bootstrapping results for **H_{3A}**.

Table 20 – H_{3A} (Bootstrapping Results)

BOOTSTRAPPING (128 samples with 500 iterations/degrees of freedom)		T-STATISTIC VALUES
Structural Model	Translation Diversity → Overall Translation Quality	4.691107 (significant @ 0.0001)
Measurement Model	TDPT	0.174752
	TDPE	0.106400
	TDRC	0.231972
	TDRE	4.184684
	AQ1	65.461778
	OTQ11	54.196157
	OTQ21	83.124210
	OTQ31	66.116474
	OTQ41	49.066368

In **H_{3A}**, there is a statistically significant relationship between *Task Repetition* and *Overall Translation Quality* with a two-tailed P-value of 0.0001 (Abramowitz, 1965); (Graphpad, 2011). The bootstrapping algorithm shows that the measurement model is statistically significant.

Figure 16 and Table 21 show the results of the PLS algorithm for **H_{3A}**.

Figure 10 - H_{3A} (PLS Algorithm Results)

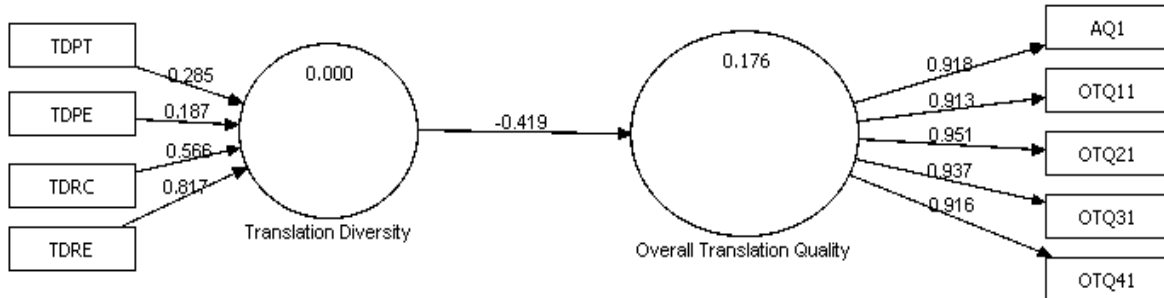


Table 21 – H_{3A} (PLS Algorithm Results)

PLS ALGORITHM	VALUES
Path Coefficient (Translation Diversity → Overall Translation Quality)	-0.418995
Cronbach's Alpha	0.959184
Composite Reliability	0.968315
Average Variance Extracted (AVE) for Overall Translation Quality	0.859423
R ²	0.175557

Cronbach's Alpha and Composite Reliability values are above 0.7, the commonly used threshold for reliability (Nunnally & Bernstein, 1994). The Average Variance Extracted (AVE) of 0.85 exceeds the 0.7 rule of thumb for validity (Fornell & Larcker, 1981). The R² value explains about 17.5% of the variance between *Translation Diversity* and *Overall Translation Quality*, while the path coefficient shows a negative correlation between these two concepts. Therefore, hypothesis **H_{3A}**, which considers all 128 samples, is rejected with $\alpha \leq 0.0001$.

The results indicate that too much diversity in crowdsourced translation has a slight negative impact on the *Overall Translation Quality*. The diversity in the revision evaluation stage of the Collaborative Online Language Transfer process had the greatest negative impact on *Overall Translation Quality*. The theory proposed by Hong and Page (2004) does not seem to be completely consistent with the results in **H_{3A}**. Here these results echo the results of **H_{1A}**; too many voices in the Collaborative Online Language Transfer process seem to slightly and negatively affect the *Overall Translation Quality* of crowdsourced translation.

Perhaps the effect of *Task Repetition* and *Translation Diversity* on *Overall Translation Quality* is not purely linear. Future research may uncover such a relationship.

6.5 Hypothesis (H₄)

When considering Overall Translation Quality, bilingual Canadians will select machine translation over crowdsourced translation.

OCOL (2005) indicates that machine translation is of generally poor quality. This null hypothesis tests if there is a significant difference between the general acceptance of machine translation quality and the general acceptance of crowdsourced translation quality. This hypothesis is tested as **H_{4A}**, with $\alpha \leq 0.05$ level including the Part I sample results; cf. Figure 8 - Phase II Participant Workflow. Table 22 shows the operationalized measurement items that were used to test **H_{4A}**.

Table 22 – H_{4A} (Measurement Indicators)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Overall Translation Quality	Before KTM	AQ1	The quality of this text is acceptable for a Canadian Francophone website.
		OTQ11	Most Canadian Francophones would not be able to tell that this text is a translation.
		OTQ21	Most Canadian Francophones would feel confident trusting the content of this translation.
		OTQ31	Most Canadian Francophone readers of this text would feel confident contacting the company in French.
		OTQ41	Most Canadian Francophone readers of this text would feel confident purchasing products or services from this company.

Table 23 shows a comparison of the highest-ranked machine translation from Phase I and the crowdsourced translation for aggregated measures of *Overall Translation Quality*. The measurements have been placed into positive, neutral, and negative categories.

Table 23 - H_{4A} (Aggregated Overall Translation Quality)

Aggregated Overall Translation Quality				
SUMMARY	CST	MT	CST %	MT %
Positive	349	4	54.5%	2.5%
Neutral	128	8	20.0%	5.0%
Negative	163	148	25.5%	92.5%
TOTAL	640	160	100%	100%
Confidence Interval			9.2029	12.7071

This table shows all five measurement indicators aggregated together and grouped into positive, neutral, and negative categories. Specific breakdowns by indicator are available in Appendix C. Here we see how crowdsourced translation compares in general to machine translation for *Overall Translation Quality*. It is clear from the confidence intervals

shown in Table 23 that crowdsourced translation significantly outperforms machine translation. Therefore, H_{4A} is rejected with $\alpha \leq 0.05$.

This result is important because it informs us as to the applicability of crowdsourced translation as a means of bridging the Canadian linguistic digital divide. This result illustrates the potential of crowdsourced translation as a third and viable option for some types of web content in Canada. Here we see that crowdsourced translation provides better quality translation that does freely available machine translation. The potential benefit here is for Canadian website owners who may be tempted turn to machine translation for their websites. This finding illustrates crowdsourced translation as a strong alternative to machine translation when comparing quality. Table 24 shows how the rankings of the crowdsourced translations and top-ranked machine translation from Phase I compare.

Table 24 - H_4 (MT vs. CST Aggregated Rank)

Before Knowledge of Methods – Aggregated Rank				
Before KTM - Rank				
RANK ORDER	CST	MT	CST %	MT %
1st - 3rd	70	0	55%	0%
4th - 6th	58	32	45%	100%
TOTAL	128	32	100%	100%
Confidence Interval			1.4700	7.8399

The rank order shown in this table aligns closely with the results depicted in Table 23. Again, we see that crowdsourced translation significantly outperforms machine translation. This finding confirms that machine translation, although faster and relatively cheap, does not provide a sufficient level of quality (OCOL, 2005) and that crowdsourced translation, in some cases, does provide a sufficient level of quality. These findings also seem to indicate a positive correlation between *Rank Order* and *Overall Translation Quality*. The analysis of this newly proposed hypothesis is presented at the end of this chapter as H_{9A} .

6.6 Hypothesis (H₅)

When considering Overall Translation Quality, bilingual Canadians will select professional translation over crowdsourced translation.

OCOL (2005) indicates that professional translation is widely accepted as the method by which content should be translated with respect to the Canadian Linguistic Digital Divide. This null hypothesis tests if there is a significant difference between the general acceptance of professional translation quality and the general acceptance of crowdsourced translation quality. This hypothesis is tested as **H_{5A}**, with $\alpha \leq 0.05$ level including the Part I sample results; cf. Figure 8 - Phase II Participant Workflow. Table 25 shows the operationalized measurement items that were used to test **H_{5A}**.

Table 25 – H_{5A} (Measurement Indicators)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Overall Translation Quality	Before KTM	AQ1	The quality of this text is acceptable for a Canadian Francophone website.
		OTQ11	Most Canadian Francophones would not be able to tell that this text is a translation.
		OTQ21	Most Canadian Francophones would feel confident trusting the content of this translation.
		OTQ31	Most Canadian Francophone readers of this text would feel confident contacting the company in French.
		OTQ41	Most Canadian Francophone readers of this text would feel confident purchasing products or services from this company.

Table 26 shows a comparison of the highest-ranked professional translation from Phase I and the crowdsourced translation for aggregated measures of *Overall Translation Quality*.

Table 26 – H_{5A} (Overall Translation Quality by Experiment Type)

Overall Translation Quality - By Experiment Type										
SCALE CATEGORIES	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT
Strongly Agree	53	23	20	8	34	33.1%	14.4%	12.5%	5.0%	21.3%
Agree	77	54	70	44	87	48.1%	33.8%	43.8%	27.5%	54.4%
Neither Agree Nor Disagree	18	40	35	35	19	11.3%	25.0%	21.9%	21.9%	11.9%
Disagree	5	22	22	42	15	3.1%	13.8%	13.8%	26.3%	9.4%
Strongly Disagree	7	21	13	31	5	4.4%	13.1%	8.1%	19.4%	3.1%
TOTAL	160	160	160	160	160	100%	100%	100%	100%	100%
Confidence Interval						9.2111	8.3452	8.6858	8.3394	9.2631

Table 26 was included to illustrate how each of the four crowdsourced translations compared to the top professional translation selected for Phase II; cf. Table 9. Here we see that CSTLP2 outperforms PT in some cases. The remaining crowdsourced translations do not perform as well on *Overall Translation Quality* measures as did the top-ranked professional translation.

Table 27 shows all five measurement indicators aggregated into positive, neutral, and negative categories

Table 27 – H_{5A} (Aggregated Overall Translation Quality)

Aggregated Overall Translation Quality				
SUMMARY	PT	CST	PT %	CST %
Positive	121	349	75.6%	54.5%
Neutral	19	128	11.9%	20.0%
Negative	20	163	12.5%	25.5%
TOTAL	160	640	100%	100%
Confidence Interval			9.0805	9.2029

Specific breakdowns by indicator are available in Appendix C. Here we see how crowdsourced translation compares in general to professional translation for *Overall Translation Quality*. It is clear from the confidence intervals shown in Table 27 that professional translation outperforms crowdsourced translation in most cases. Therefore, H_{5A} is generally accepted with $\alpha \leq 0.05$.

This result is important because it tells us how applicable crowdsourced translation is as a means of bridging the Canadian linguistic digital divide. As was noted in H_{4A}, crowdsourced translation is not a panacea, but that it has potential benefits for Canadians. Specifically, these results show that in some cases, language professionals translating in a diverse, aggregated, decentralized and independent way (Surowiecki, 2005) can outperform traditional professional translation; cf. Table 9. Only one of the three crowdsourced translations was ranked lower than the lowest-ranked professional translation. These findings show that there is some potential for crowdsourced translation by non-language professionals in some cases to perform as well as or better than some commercially available alternatives.

Table 28 shows how the rankings of the crowdsourced translations and top-ranked professional translation from Phase I compare.

Table 28 – H₅ (PT vs. CST Aggregated Rank)

Before Knowledge of Methods – Aggregated Rank				
Before KTM - Rank				
RANK ORDER	CST	PT	CST %	PT %
1st - 3rd	70	26	55%	81%
4th - 6th	58	6	45%	19%
TOTAL	128	32	100%	100%
Confidence Interval			1.4700	4.8999

The rank order shown in Table 28 aligns closely with the results depicted in Table 27. These findings generally confirm that professional translation has superior *Overall Translation Quality* when compared to crowdsourced translation. These findings also seem to indicate a positive correlation between *Rank Order* and *Overall Translation Quality*. The analysis of this newly proposed hypothesis is presented at the end of this chapter as **H_{9A}**.

6.7 Hypothesis (H₆)

The absence of Contextual Overt Translation Errors will have a positive effect on the Overall Translation Quality of crowdsourced translation.

House (1997) suggests that when overt translation errors are absent, there is a positive effect on the quality of a translation. This hypothesis is proposed to test whether there is a significant correlation between the absence of contextual overt translation errors and the overall quality of crowdsourced translation. This hypothesis is tested as **H_{6A}**, with $\alpha \leq 0.05$ level which includes all 128 samples. The results of testing **H_{6D}** and **H_{6E}** are discussed at the end of the section to identify key differences in how Language Professionals and Non-Language Professionals evaluated the crowdsourced translations. Table 29 shows the operationalized measurement items that were used to test **H_{6A}**, **H_{6D}**, and **H_{6E}**.

Table 29 – H_{6A} (Structural and Measurement Model)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Contextual Overt Translation Errors	Formative	COERR11	This text uses correct terminology.
		COERR21	This text uses correct grammar.
		COERR31	This text has an understandable meaning.
Overall Translation Quality	Reflective	AQ1	The quality of this text is acceptable for a Canadian Francophone website.
		OTQ11	Most Canadian Francophones would not be able to tell that this text is a translation.
		OTQ21	Most Canadian Francophones would feel confident trusting the content of this translation.
		OTQ31	Most Canadian Francophone readers of this text would feel confident contacting the company in French.
		OTQ41	Most Canadian Francophone readers of this text would feel confident purchasing products or services from this company.

Table 30 shows the bootstrapping results for **H_{6A}**.

Table 30 – H_{6A} (Bootstrapping Results)

BOOTSTRAPPING (128 samples with 500 iterations/degrees of freedom)		T-STATISTIC VALUES
Structural Model	Contextual Overt Translation Errors → Overall Translation Quality	14.161312 (significant @ 0.0001)
Measurement Model	COERR11	7.616177
	COERR21	4.075129
	COERR31	0.697001
	AQ1	64.502033
	OTQ11	69.194068
	OTQ21	85.611660
	OTQ31	59.288413
	OTQ41	48.687662

In **H_{6A}**, there is a statistically significant relationship between *Task Repetition* and *Overall Translation Quality* with a two-tailed P-value of 0.0001 (Abramowitz, 1965); (Graphpad, 2011). The bootstrapping algorithm shows that the measurement model is statistically significant. The bootstrapping also indicates statistically significant measurement models for the sub-hypotheses **H_{6D}** and **H_{6E}**. Here we see a very low t-statistic value on the question of whether or not the translation had an understandable meaning (COERR31). Since this measurement takes into account whether or not the source text had an understandable meaning, we can interpret this low value in the following way: the crowdsourced translators stuck very closely to the original meaning of the source text and did not significantly improve upon it. Figure 11 and Table 31 shows the PLS algorithm results for **H_{6A}**.

Figure 11 - H_{6A} (PLS Algorithm Results)

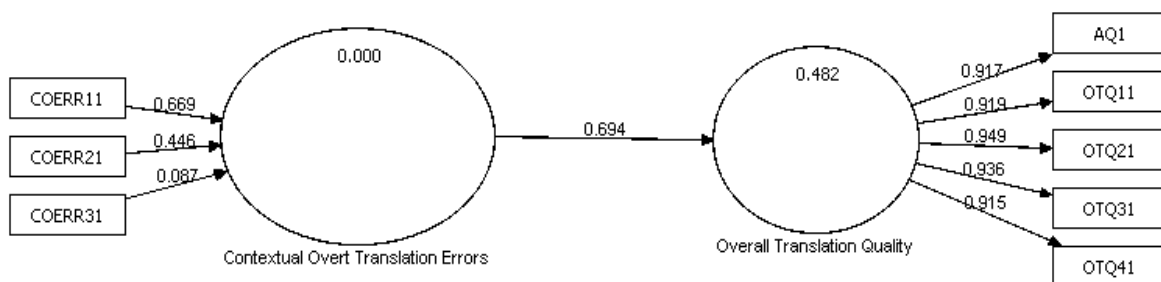


Table 31 – H_{6A} (PLS Algorithm Results)

PLS ALGORITHM	VALUES
Path Coefficient (Contextual Overt Translation Errors) → Overall Translation Quality)	0.694364
Cronbach's Alpha	0.959184
Composite Reliability	0.968358
Average Variance Extracted (AVE) for Overall Translation Quality	0.859589
R ²	0.482141

Cronbach’s Alpha and Composite Reliability values are above 0.7, the commonly used threshold for reliability (Nunnally & Bernstein, 1994). The Average Variance Extracted (AVE) of 0.85 exceeds the 0.7 rule of thumb for validity (Fornell & Larcker, 1981). The R² value explains about 48.2% of the variance between *Contextual Overt Translation Errors* and *Overall Translation Quality*, while the path coefficient shows a positive correlation between these two concepts. Therefore, hypothesis **H_{6A}**, which considers all 128 samples, is accepted with $\alpha \leq 0.0001$. These results support the findings of House (1997) who suggests that the absence of *Contextual Overt Translation Errors* has a positive effect on translation quality.

Here we see that the PLS Algorithm results echo the t-statistic results from the bootstrapping. On the question of whether or not the translation had an understandable meaning, very little of the variance in the measure explains the variance in *Overall Translation Quality*. This implies that both Language Professional and Non-Language Professional crowdsourced translators produced consistently similar quality translations with respect to conveying the meaning of the source text through to the translation. Next, the PLS Algorithm results for **H_{6D}** and **H_{6E}** are shown to identify key differences in how Language Professionals and Non-Language Professionals evaluated the *Contextual Overt Translation Errors* of the crowdsourced translations. Figure 12 and Table 32 shows the PLS algorithm results for the Language Professional translation evaluators (**H_{6D}**).

Figure 12 - H_{6D} (PLS Algorithm Results)

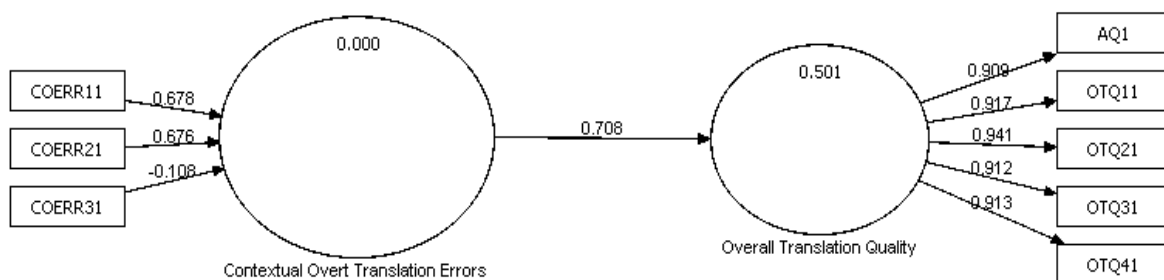


Table 32 – H_{6D} (PLS Algorithm Results)

TRANSLATION EVALUATORS – Language Professionals	
PLS ALGORITHM	VALUES
Path Coefficient (Contextual Overt Translation Errors → Overall Translation Quality)	0.707787
Cronbach’s Alpha	0.953663
Composite Reliability	0.964187
Average Variance Extracted (AVE) for Overall Translation Quality	0.843394
R ²	0.500963

When we consider only the Phase II results that were created by Language Professionals, the findings suggest that the translation training or translation education of Language Professionals leads them to be more critical of the quality of crowdsourced translations with regard to how the source text’s meaning was transferred to the target texts (COERR31). The effect of *Contextual Overt Translation Errors* among Language Professionals is slightly stronger than the effect among the entire sample population, representing just over 50% of the explained variance (up from about 48.2%).

Figure 13 and Table 33 show the PLS algorithm results for the Non-Language Professional translation evaluators (**H_{6E}**).

Figure 13 - H_{6E} (PLS Algorithm Results)

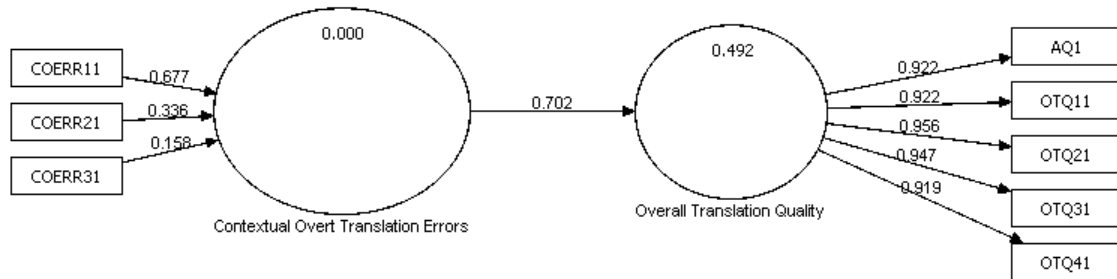


Table 33 – H_{6E} (PLS Algorithm Results)

TRANSLATION EVALUATORS – Non-Language Professionals	
PLS ALGORITHM	VALUES
Path Coefficient (Contextual Overt Translation Errors → Overall Translation Quality)	0.701604
Cronbach’s Alpha	0.962941
Composite Reliability	0.971212
Average Variance Extracted (AVE) for Overall Translation Quality	0.870954
R ²	0.492249

This result indicates that Non-Language Professionals are more accepting and less critical of how well the meaning of the source text is carried through to a crowdsourced translation than are their Language Professional counterparts (COERR31). For both groups, issues related to correct terminology were the most important, followed by issues related to grammar.

6.8 Hypothesis (H₇)

The absence of Contextual Usage Translation Errors will have a positive effect on the Overall Translation Quality of crowdsourced translation.

House (1997) suggests that when usage translation errors are absent, there is a positive effect on the quality of a translation. This hypothesis is proposed to test if there is a significant correlation between the absence of contextual usage translation errors and the overall quality of crowdsourced translation. This hypothesis is tested as **H_{7A}**, with $\alpha \leq 0.05$ level which includes all 128 samples. Table 34 shows the operationalized measurement items that were used to test **H_{7A}**.

Table 34 – H_{7A} (Structural and Measurement Model)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Contextual Usage Translation Errors	Formative	CUERR11	This text is written in a style that calls the reader to action.
		CUERR21	This text is written for an educated middle class audience.
		CUERR31	This text is written in the style of a professional blogger.
		CUERR41	This text is written in a style that is intended to be read instead of a style that is intended to be spoken.
		CUERR51	This text is written in a style that is: Frozen, Formal, Consultative, Casual, or Intimate.
Overall Translation Quality	Reflective	AQ1	The quality of this text is acceptable for a Canadian Francophone website.
		OTQ11	Most Canadian Francophones would not be able to tell that this text is a translation.
		OTQ21	Most Canadian Francophones would feel confident trusting the content of this translation.
		OTQ31	Most Canadian Francophone readers of this text would feel confident contacting the company in French.
		OTQ41	Most Canadian Francophone readers of this text would feel confident purchasing products or services from this company.

Table 35 shows the bootstrapping results for **H_{7A}**.

Table 35 – H_{7A} (Bootstrapping Results)

BOOTSTRAPPING (128 samples with 500 iterations/degrees of freedom)		T-STATISTIC VALUES
Structural Model	Contextual Usage Translation Errors → Overall Translation Quality	11.422574 (significant @ 0.0001)
Measurement Model	CUERR11	2.596188
	CUERR21	3.501232
	CUERR31	3.812360
	CUERR41	3.728369
	CUERR51	0.116423
	AQ1	57.539113
	OTQ11	62.170153
	OTQ21	98.774123
	OTQ31	69.137728
	OTQ41	51.620056

In **H_{7A}**, there is a statistically significant relationship between *Task Repetition* and *Overall Translation Quality* with a two-tailed P-value of 0.0001 (Abramowitz, 1965); (Graphpad, 2011). The bootstrapping algorithm shows that the measurement model is statistically significant. In looking at these results, it is apparent that on the question of the style of the translation (CUERR51) a very low t-statistic value was reported. This implies that the crowdsourced translators were relatively consistent in capturing the style of the source text using the qualitative goals that were presented to them in Phase I. While CUERR51 seems to contribute little to the construct *Contextual Usage Translation Errors*, eliminating it would alter the meaning of the construct (Roberts & Thatcher, 2009) proposed by House (1997).

Figure 14 and Table 36 show the PLS algorithm results for **H_{7A}**.

Figure 14 – H_{7A} (PLS Algorithm Results)

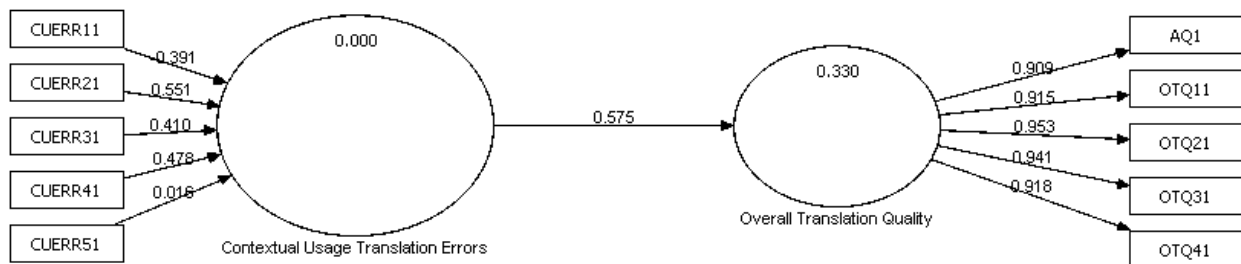


Table 36 – H_{7A} (PLS Algorithm Results)

PLS ALGORITHM	VALUES
Path Coefficient (Contextual Usage Translation Errors) → Overall Translation Quality)	0.574700
Cronbach's Alpha	0.959184
Composite Reliability	0.968411
Average Variance Extracted (AVE) for Overall Translation Quality	0.859814
R ²	0.330280

Cronbach's Alpha and Composite Reliability values are above 0.7, the commonly used threshold for reliability (Nunnally & Bernstein, 1994). The Average Variance Extracted (AVE) of 0.85 exceeds the 0.7 rule of thumb for validity (Fornell & Larcker, 1981). The R^2 value explains about 33% of the variance between *Contextual Usage Translation Errors* and *Overall Translation Quality*, while the path coefficient shows a positive correlation between these two concepts. Therefore, hypothesis **H_{7A}**, which considers all 128 samples, is accepted with $\alpha \leq 0.0001$. These results support the findings of House (1997) who suggests that the absence of *Contextual Usage Translation Errors* has a positive effect on translation quality.

On the question of whether or not the translation called the reader to action (CUERR11), the findings suggest that the crowdsourced translations in general did not have the same kind of impact as the source text. For the source text, this question could be interpreted as "does the target text entice the reader to visit the Black Camel Café and to try one of the delicious sounding sandwiches?" It is interesting to note that from the Phase I data, one of the reasons paragraph translations were discarded was because the paragraph evaluators felt that the food being described in the translation was unappetizing. For example, several instances of translated paragraphs were rejected because a participant had transformed the phrase "a Black Camel chili" into Canadian French with the meaning of "a chili made with the meat of a black camel". This kind of translation error seems to support the sentiment of Jorgensen (2009) who posits that poor translations associated to a company's products or services can negatively affect overall customer satisfaction.

6.9 Hypothesis (H₈)

The absence of Contextual Author Translation Errors will have a positive effect on the Overall Translation Quality of crowdsourced translation.

House (1997) suggests that when author translation errors are absent, there is a positive effect on the quality of a translation. This hypothesis is proposed to test if there is a significant correlation between the absence of contextual author translation errors and the overall quality of crowdsourced translation. This hypothesis is tested as **H_{8A}**, with $\alpha \leq 0.05$ level which includes all 128 samples. Table 37 shows the operationalized measurement items that were used to test **H_{8A}**.

Table 37 – H_{8A} (Structural and Measurement Model)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Contextual Author Translation Errors	Formative	CAERR11	This text appears to be written by a fluent Canadian Anglophone (ST)/Francophone (TT).
		CAERR21	This text appears to be written by someone with a university education.
		CAERR31	This text is written using language that was appropriate in 2010.
Overall Translation Quality	Reflective	AQ1	The quality of this text is acceptable for a Canadian Francophone website.
		OTQ11	Most Canadian Francophones would not be able to tell that this text is a translation.
		OTQ21	Most Canadian Francophones would feel confident trusting the content of this translation.
		OTQ31	Most Canadian Francophone readers of this text would feel confident contacting the company in French.
		OTQ41	Most Canadian Francophone readers of this text would feel confident purchasing products or services from this company.

Table 38 shows the bootstrapping results for **H_{8A}**.

Table 38 – H_{8A} (Bootstrapping Results)

BOOTSTRAPPING (128 samples with 500 iterations/degrees of freedom)		T-STATISTIC VALUES
Structural Model	Contextual Author Translation Errors → Overall Translation Quality	11.907198 (significant @ 0.0001)
Measurement Model	CAERR11	6.469246
	CAERR21	1.198592
	CAERR31	2.404722
	AQ1	57.285442
	OTQ11	71.422398
	OTQ21	92.337753
	OTQ31	77.273425
	OTQ41	58.097928

In **H_{8A}**, there is a statistically significant relationship between *Task Repetition* and *Overall Translation Quality* with a two-tailed P-value of 0.0001 (Abramowitz, 1965); (Graphpad, 2011). In the measurement model, we see issues related to the fluency of the Phase I participants as having the most significance (CAERR11) when determining the *Overall Translation Quality* of a crowdsourced translation.

Figure 15 and Table 39 show the PLS algorithm results for **H_{8A}**.

Figure 15 – H_{8A} (PLS Algorithm Results)

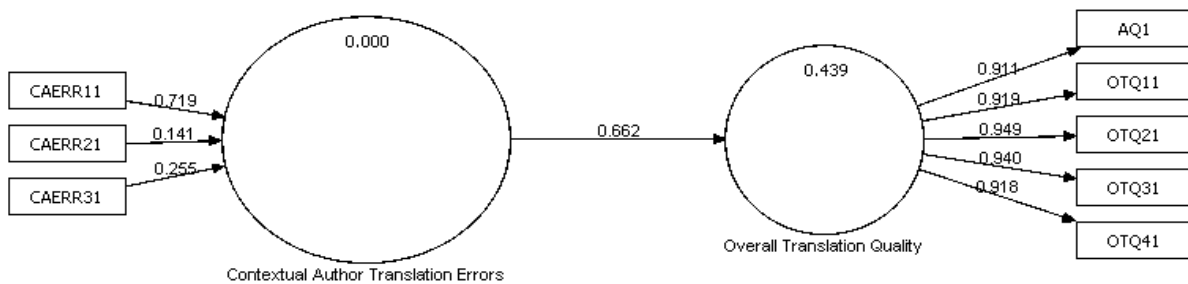


Table 39 – H_{8A} (PLS Algorithm Results)

PLS ALGORITHM	VALUES
Path Coefficient (Contextual Author Translation Errors) → Overall Translation Quality)	0.662289
Cronbach's Alpha	0.959184
Composite Reliability	0.968413
Average Variance Extracted (AVE) for Overall Translation Quality	0.859810
R ²	0.438627

Cronbach's Alpha and Composite Reliability values are above 0.7, the commonly used threshold for reliability (Nunnally & Bernstein, 1994). The Average Variance Extracted (AVE) of 0.85 exceeds the 0.7 rule of thumb for validity (Fornell & Larcker, 1981). The R² value explains about 43.8% of the variance between *Contextual Author Translation Errors* and *Overall Translation Quality*, while the path coefficient shows a positive correlation between these two concepts. Therefore, hypothesis **H_{8A}**, which considers all 128 samples, is accepted with $\alpha \leq 0.0001$. These results support the findings of House (1997) who suggests that the absence of *Contextual Author Translation Errors* has a positive effect on translation quality. No sub-hypotheses were tested since the item loadings did not appear to be unusually small; nor did they appear to negatively correlate to the latent variable, *Contextual Author Translation Errors*.

A loose interpretation of House's translation quality model (1997) is captured by the hypotheses **H_{6A}**, **H_{7A}**, and **H_{8A}**. When looking at the PLS Algorithm results for these hypotheses towards *Overall Translation Quality*, we see the following factors as the most

important in understanding how the quality of crowdsourced translation is interpreted. In all cases, issues relating to trust (OTQ21) and issues related to being able to contact the company in French (OTQ31) stood out as the most important. Issues related to acceptable quality (ACQ1), purchasing products or services (OTQ41), and being able to tell that the text was translated (OTQ11), while significant, were less important to the Phase II translation evaluators. These findings suggest that the quality of a crowdsourced translation can help to establish trust with customers, which in turn may make those customers feel comfortable enough to contact the company in their own language. Conversely, if the translations do not capture the trust of the reader and discourage contact with the company, those customers are more likely to have negative attitudes or opinions of the company's products or services (Jorgensen, 2009) and may be less likely to do business with the company. This seems to lend support to the observation that issues of language quality in translation contribute to whether or not customers will proceed with purchases (DePalma, Sargent, & Beninatto, 2006).

6.9 Crowdsourced Translation Acceptance

How is the idea of crowdsourced translation accepted as a means of bridging the Canadian linguistic digital divide?

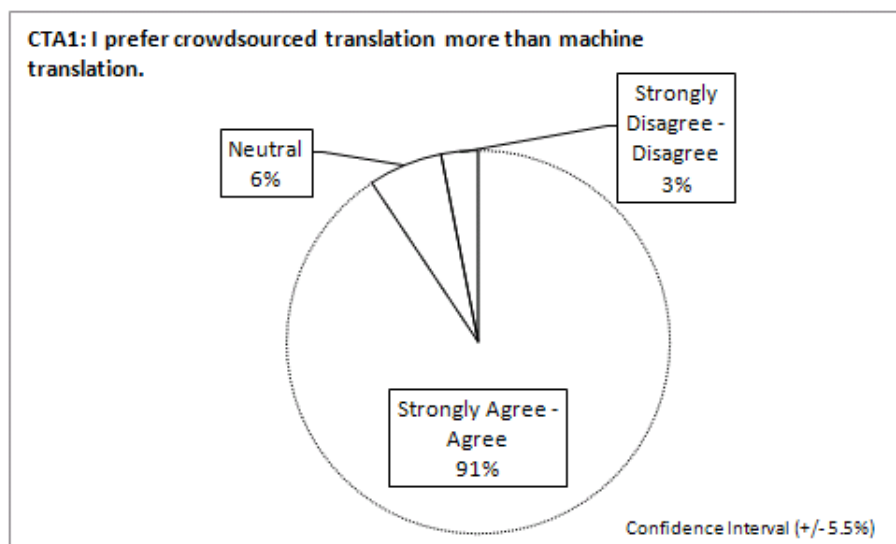
This question was addressed through the use of the supplemental evaluation survey that was completed at the end of the questionnaire process. Next are the definitions of each of the measurement indicators and tables of the values, followed by five charts illustrating each measurement indicator using three-tier aggregations (strongly disagree/disagree, neutral, and agree/strongly agree). Table 40 shows the operationalized measurement items that are reflected by the concept of *Crowdsourced Translation Acceptance*.

Table 40 – Crowdsourced Translation Acceptance (Measurement Indicators)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Crowdsourced Translation Acceptance	After KTM	CTA1	I prefer crowdsourced translation more than machine translation.
		CTA2	I prefer crowdsourced translation more than professional translation.
		CTA3	Crowdsourced translation should be performed by non-language professionals.
		CTA4	Crowdsourced translation should be performed by language professionals.
		CTA5	Crowdsourced translation should be used to translate website content in Canada.

Figure 16 shows the aggregated response of the Phase II participants for indicator CTA1 – *I prefer crowdsourced translation more than machine translation*.

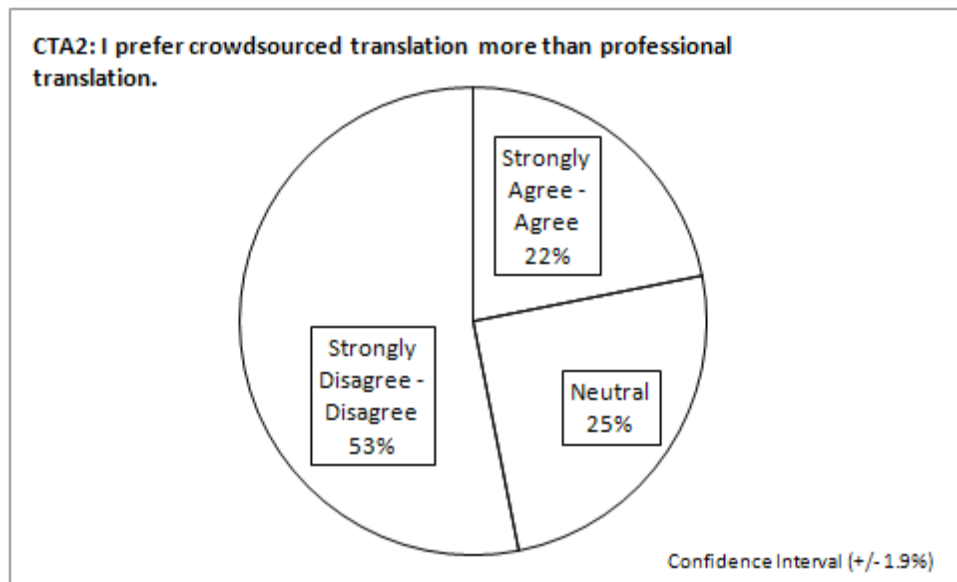
Figure 16 - CTA1 Chart



Here we see overwhelming support for the idea that crowdsourced translation is superior to machine translation (91%). This result supports the findings of **H_{4A}**.

Figure 17 shows the aggregated response of the Phase II participants for indicator CTA2 – *I prefer crowdsourced translation more than professional translation.*

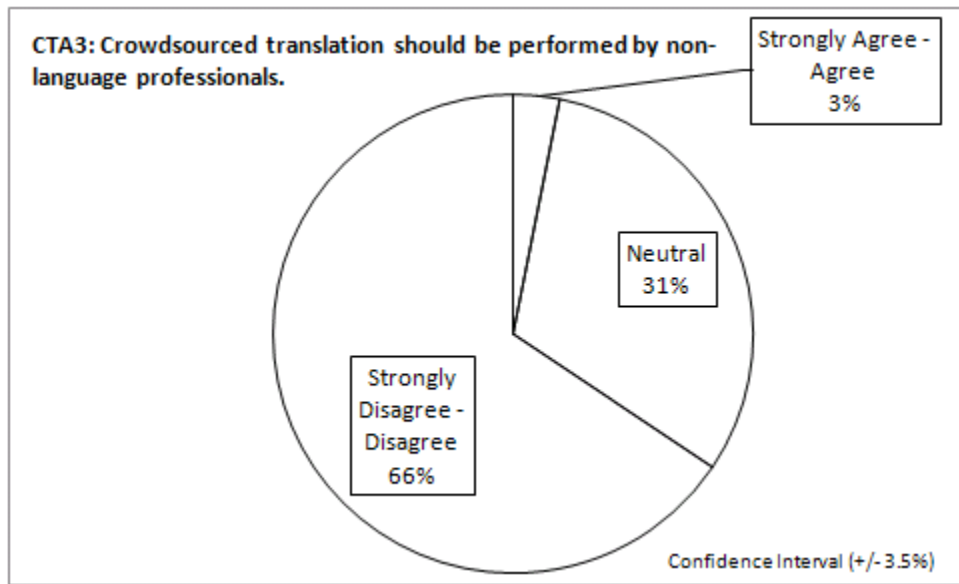
Figure 17 – CTA2 Chart



In Figure 17 we see that professional translation is superior to crowdsourced translation (53%); however, there is a significant minority (22%) that feels that crowdsourced translation is superior to professional translation. These results support the findings of **H_{5A}**. Further research could be done to determine which demographics are more open to crowdsourced translation and which are less open to it.

Figure 18 shows the aggregated response of the Phase II participants for indicator CTA3 – *Crowdsourced translation should be performed by non-language professionals.* This chart summarizes the participants' views about the acceptability of non-language professionals engaging in crowdsourced translation.

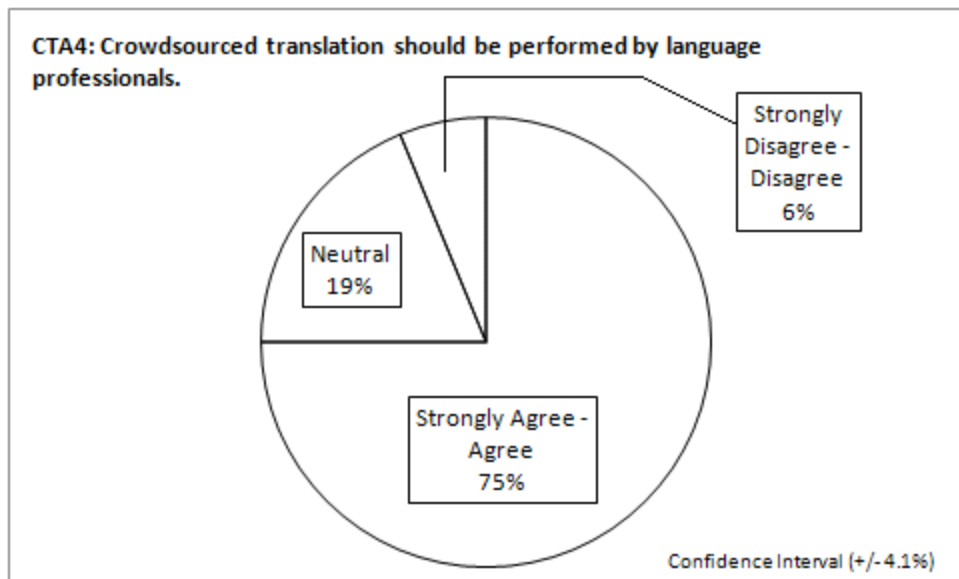
Figure 18 – CTA3 Chart



Here we can see a resistance to the idea of non-professionals performing crowdsourced translation (66%) and extremely limited support this for idea (3%).

Figure 19 shows the aggregated response of the Phase II participants for indicator CTA4 – *Crowdsourced translation should be performed by language professionals*. This chart summarizes the participants' views about the acceptability of language professionals engaging in crowdsourced translation.

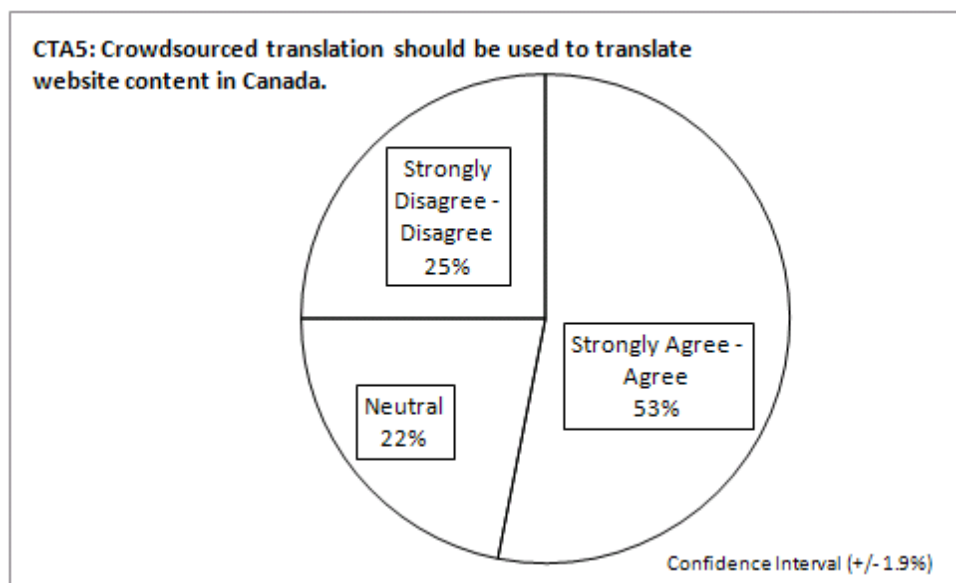
Figure 19 – CTA4 Chart



In Figure 19 we can see a strong acceptance of the idea of crowdsourced translation being performed by Language Professionals (75%). While the preference is given to traditional professional translation methods (CTA2), this result (CTA4) could help inform policy-makers about the general acceptance of Language Professionals engaging in crowdsourced translation. It suggests that crowdsourced translation could be considered as a possible avenue for translation training and education among Language Professionals. Very few respondents disagreed (6%) with the idea that Language Professionals should perform crowdsourced translation.

Figure 20 shows the aggregated response of the Phase II participants for indicator CTA3 – *Crowdsourced translation should be used to translation website content in Canada*. This chart summarizes the participants’ views about the acceptability of crowdsourced translation as a means of bridging the Canadian Linguistic Digital Divide.

Figure 20 – CTA5 Chart



In Figure 20 we can see a general acceptance of crowdsourcing as a means of translating website content in Canada (53%). There is a significant minority that feels that crowdsourced translation should not be used in this way (25%). Table 41 illustrates a more nuanced view of the data and shows some distinct differences between Language Professionals and Non-Language professionals for the concept of *Crowdsourced Translation Acceptance*.

Table 41 - Crowdsourced Translation Acceptance by Participant Type

INDICATOR	DESCRIPTION	LANGUAGE PROFESSIONALS	NON-LANGUAGE PROFESSIONALS
CTA1	I prefer crowdsourced translation more than machine translation.	97%	84%
CTA2	I prefer crowdsourced translation more than professional translation.	48%	53%
CTA3	Crowdsourced translation should be performed by non-language professionals.	42%	45%
CTA4	Crowdsourced translation should be performed by language professionals.	78%	79%
CTA5	Crowdsourced translation should be used to translate website content in Canada.	57%	71%

These results represent the average values for each participant group. With respect to CTA1, we see a very strong reaction against machine translation in support of crowdsourced translation (97%) among Language Professionals. This finding corresponds to how language professionals viewed machine translation in the study by Bowker (2008). In CTA2, we see a slight majority of Non-Language Professional participants who preferred the crowdsourced translations to the professional translation; most Language Professionals do not prefer crowdsourced translation to professional translation. In CTA3, most people in both groups felt that crowdsourced translation should not be performed by non-language professionals, however in both cases, those participants who did made up a significant minority (42-45%). CTA4 showed roughly the same finding for both groups, which is that language professionals should be engaged in crowdsourced translation. Finally, in CTA5 we see distinct differences between how Language Professionals and Non-Language professionals view crowdsourcing as a means of bridging the Canadian Linguistic Digital Divide. There is clearly more support among Non-Language Professionals for crowdsourced translation to be used for this purpose, whereas Language Professionals can be characterized as being supportive but skeptical of the idea. Overall, we see that there is more support among Non-Language Professionals than there is among Language Professionals for using crowdsourcing as a means of bridging the Canadian Linguistic Digital Divide.

6.10 Rank Order Positively Influences Overall Translation Quality

An increase in Rank Order has a positive effect on the Overall Translation Quality of crowdsourced translation.

Based on the observations made in the analysis of **H_{4A}** and **H_{5A}**, this hypothesis is proposed to test the significance of *Rank Order* on the *Overall Translation Quality* of crowdsourced translation. This hypothesis is tested as **H_{9A}**, with $\alpha \leq 0.05$ level which includes all 128 samples. It should be noted that the original rank order measurements were recorded as a rank of one (1) being the best and a rank of six (6) being the worst. The rank order data has been recalculated to subtract the recorded rank order from seven, in order to invert the values. Table 42 shows the operationalized measurement items that were used to test **H_{9A}**.

Table 42 – H_{9A} (Structural and Measurement Model)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Rank Order	Formative	R1CST	Rank Order for the crowdsourced translations.
Overall Translation Quality	Reflective	AQ1	The quality of this text is acceptable for a Canadian Francophone website.
		OTQ11	Most Canadian Francophones would not be able to tell that this text is a translation.
		OTQ21	Most Canadian Francophones would feel confident trusting the content of this translation.
		OTQ31	Most Canadian Francophone readers of this text would feel confident contacting the company in French.
		OTQ41	Most Canadian Francophone readers of this text would feel confident purchasing products or services from this company.

Table 43 shows the bootstrapping results for **H_{9A}**.

Table 43 – H_{9A} (Bootstrapping Results)

BOOTSTRAPPING (128 samples with 500 iterations/degrees of freedom)		T-STATISTIC VALUES
Structural Model	Rank Order → Overall Translation Quality	13.934517 (significant @ 0.0001)
Measurement Model	AQ1	66.887441
	OTQ11	56.106339
	OTQ21	78.889671
	OTQ31	70.314975
	OTQ41	56.436324

In **H_{9A}**, there is a statistically significant relationship between *Rank Order* and *Overall Translation Quality* with a two-tailed P-value of 0.0001 (Abramowitz, 1965); (Graphpad, 2011). The bootstrapping algorithm shows that the measurement model is statistically significant.

Figure 21 and Table 44 show the PLS algorithm results for **H_{9A}**.

Figure 21 – H_{9A} (PLS Algorithm Results)

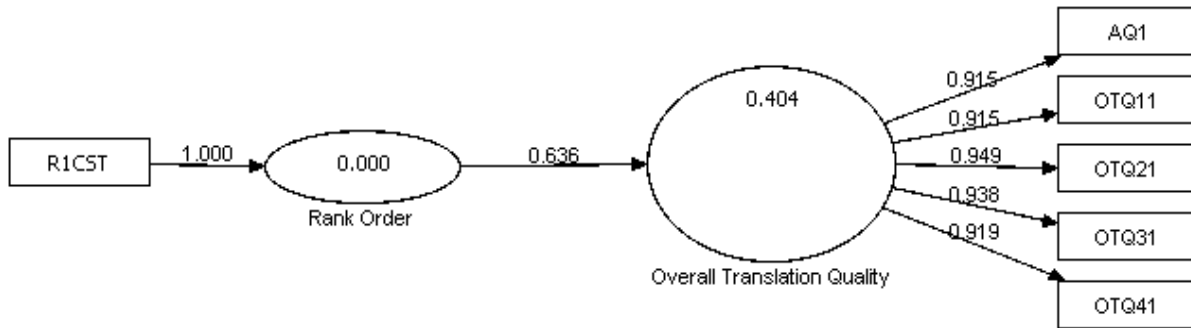


Table 44 – H_{9A} (PLS Algorithm Results)

PLS ALGORITHM	VALUES
Path Coefficient (Rank Order → Overall Translation Quality)	0.635786
Cronbach's Alpha	0.959184
Composite Reliability	0.968390
Average Variance Extracted (AVE) for Overall Translation Quality	0.859720
R ²	0.404223

Cronbach's Alpha and Composite Reliability values are above 0.7, the commonly used threshold for reliability (Nunnally & Bernstein, 1994). The Average Variance Extracted (AVE) of 0.85 exceeds the 0.7 rule of thumb for validity (Fornell & Larcker, 1981). The R² value explains about 40.4% of the variance between *Rank Order* and *Overall Translation Quality*, while the path coefficient shows a negative correlation between these two concepts. Therefore, hypothesis **H_{9A}**, which considers all 128 samples, is accepted with $\alpha \leq 0.0001$.

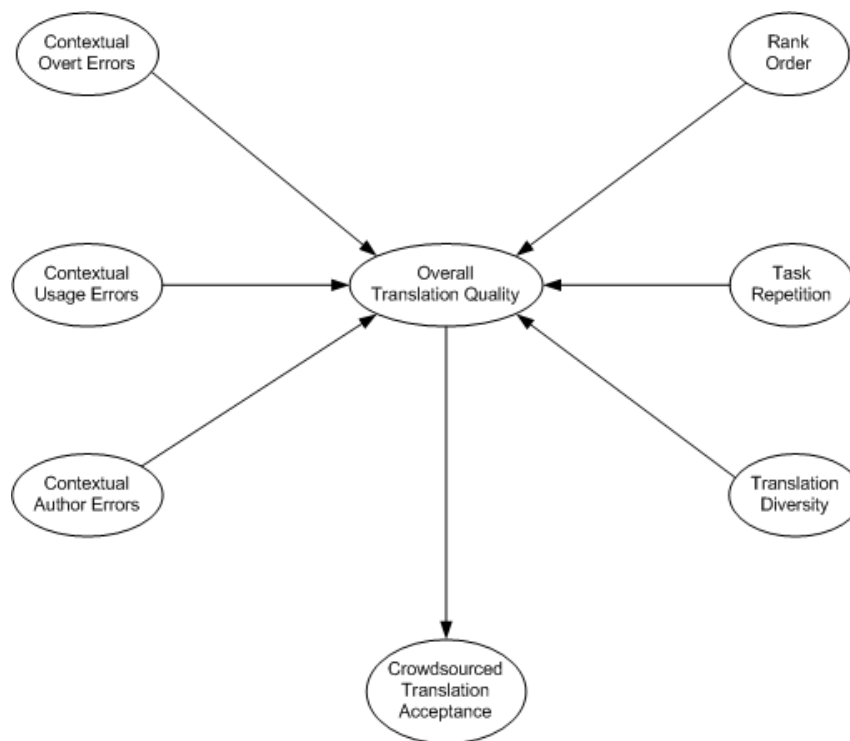
This finding confirms the observations about *Rank Order* in the analysis of **H_{4A}** and **H_{5A}**. *Rank Order* shows predictive significance for the *Overall Translation Quality* of crowdsourced translation. Crowdsourced translations that were ranked higher usually were judged to be of higher quality, but these findings show that rank order by itself only explains about 40.4% of the quality. Future research in this area should use *Rank Order* as an indicator for quality, but should not rely entirely on it to predict crowdsourced translation quality.

6.11 Multivariate Model of Crowdsourced Translation

Can a model be created that can estimate the effectiveness of crowdsourced translations as a means of bridging the Canadian Linguistic Digital Divide?

Here, a more comprehensive PLS path model that includes multiple latent variables is defined and tested. This model is proposed to better understand how each of the major theoretical concepts interacts with concept of the *Crowdsourced Translation Acceptance* (CTA). Although the previous hypotheses were tested in isolation, and shown to be individually significant, this section extends the analysis and describes which theoretical concepts were the most predictive for this study. Figure 28 illustrates the first multivariate PLS path model. Here we also try to test the relationship between *Overall Translation Quality* and *Crowdsourced Translation Acceptance*.

Figure 22 - Multivariate Model of Crowdsourced Translation, Version #1



The multivariate model in Figure 22 focuses primarily on the factors that affect the quality of crowdsourced translation and does not consider other factors such as time or cost. The measurement model indicators for each theoretical concept are not described in this section as they have been previously defined and validated for each hypothesis. Next, the

bootstrapping results and PLS Algorithm results are presented. This first multivariate model is tested with $\alpha \leq 0.05$ level including all 128 samples.

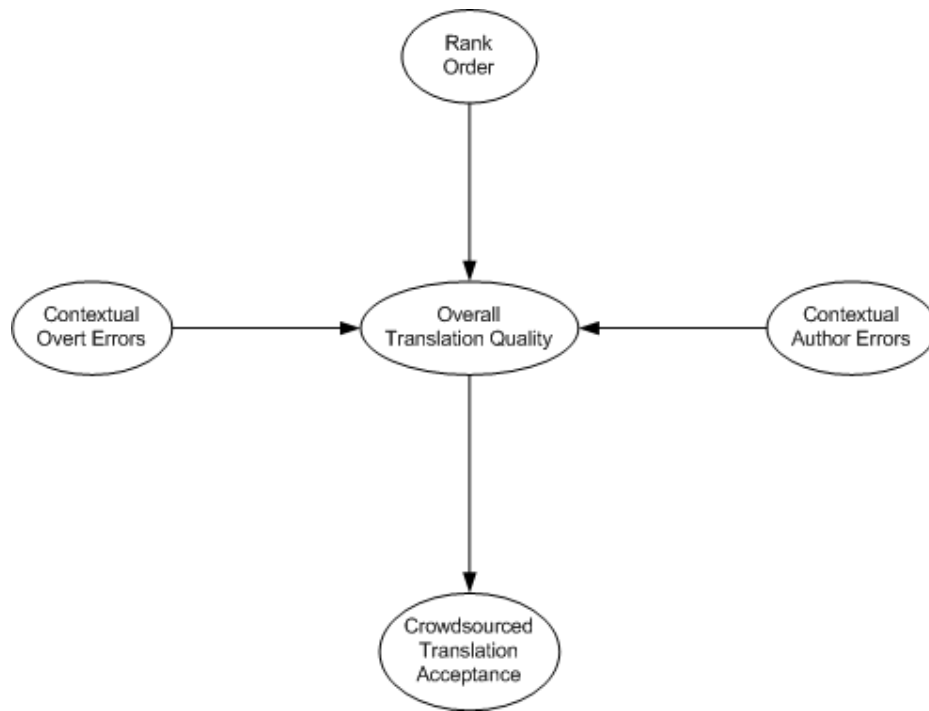
Table 45 shows the bootstrapping results from the first multivariate PLS path model.

Table 45 – Multivariate Model, Version #1 (Bootstrapping Results)

BOOTSTRAPPING (128 samples with 500 iterations/degrees of freedom)		T-STATISTIC VALUES
Structural Model	Contextual Overt Errors → Overall Translation Quality	3.427162 (significant @ 0.001)
	Contextual Usage Errors → Overall Translation Quality	1.810203 (not statistically significant)
	Contextual Author Errors → Overall Translation Quality	2.354090 (significant @ 0.01)
	Rank Order → Overall Translation Quality	2.880417 (significant @ 0.005)
	Task Repetition → Overall Translation Quality	0.728806 (not statistically significant)
	Translation Diversity → Overall Translation Quality	0.653812 (not statistically significant)
	Overall Translation Quality → Crowdsourced Translation Acceptance	4.430085 (significant @ 0.0001)

When running the bootstrapping on the first multivariate model, it becomes clear that some of the theoretical concepts that were tested in isolation no longer have any statistical significance. Specifically, *Contextual Usage Errors*, *Task Repetition*, and *Translation Diversity* no longer significantly influence our understanding of the *Overall Translation Quality* as it pertains to *Crowdsourced Translation Acceptance*. Starting by deleting the least significant theoretical concept and continuing until no insignificant theoretical concepts remained, the following multivariate model is identified. Figure 29 illustrates the second multivariate PLS path model.

Figure 23 - Multivariate Model of Crowdsourced Translation, Version #2



The second multivariate PLS path model is streamlined, showing only the significant theoretical concepts. This model also attempts to understand if there is a significant relationship between *Overall Translation Quality* and *Crowdsourced Translation Acceptance*.

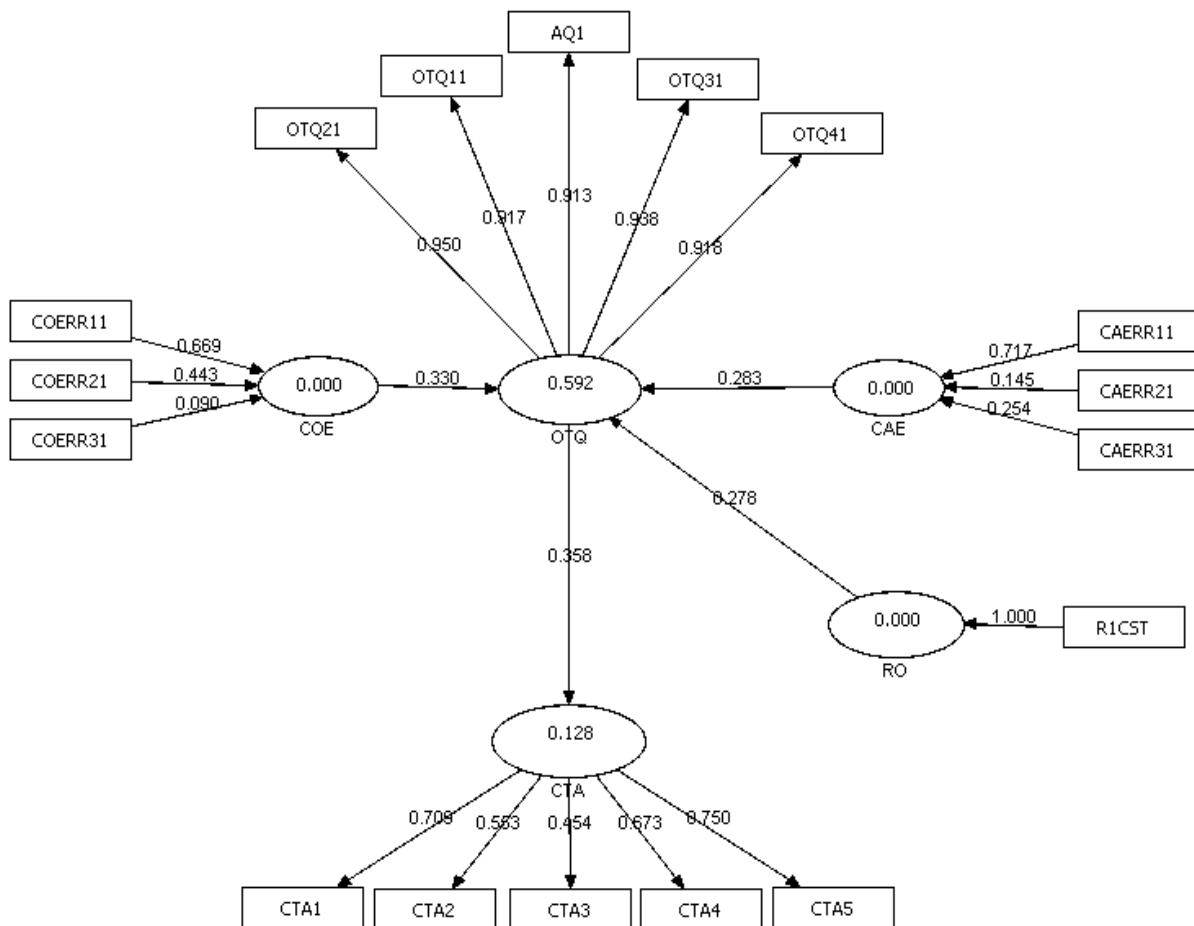
Table 46 shows the bootstrapping results from the second multivariate PLS path model.

Table 46 – Multivariate Model, Version #2 (Bootstrapping Results)

BOOTSTRAPPING (128 samples with 500 iterations/degrees of freedom)		T-STATISTIC VALUES
Structural Model	Contextual Overt Errors → Overall Translation Quality	4.962727 (significant @ 0.0001)
	Contextual Author Errors → Overall Translation Quality	3.812282 (significant @ 0.001)
	Rank Order → Overall Translation Quality	4.091378 (significant @ 0.0001)
	Overall Translation Quality → Crowdsourced Translation Acceptance	5.631862 (significant @ 0.0001)

Here we see statistically significant structural model loadings for these four theoretical concepts. Figure 24 shows the output from the PLS algorithm for the second multivariate PLS path model.

Figure 24 - Multivariate Model, Version #2 (PLS Algorithm Results)



At first glance, we see that this multivariate model explains 59.2% of the variance in *Overall Translation Quality*. This configuration gives us our strongest predictors for *Overall Translation Quality*, and as such, this model should be helpful in informing future models. When we examine how much *Overall Translation Quality* influences *Crowdsourced Translation Acceptance*, we begin to see that this model explains very little of the relationship. Only about 12.8% of the variance between *Overall Translation Quality* and *Crowdsourced Translation Acceptance* is explained by this model. This finding strongly implies that other factors in addition to quality must influence how crowdsourced translations are accepted. Future research could examine the effect of time and cost factors on the acceptance of crowdsourced translation. Future research could also be done to study in more detail the differences in how Language Professionals and Non-Language Professionals accept crowdsourced translation. In this model (cf. Figure 30), quality only explains about 13% of the acceptability of crowdsourced translation, leaving about 87% of the acceptance unexplained.

Table 47 shows the PLS algorithm results for the second multivariate PLS path model.

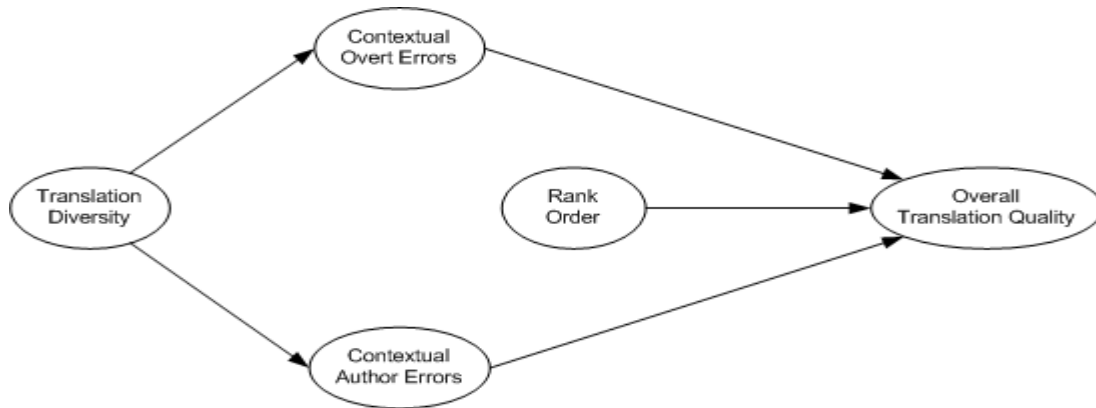
Table 47 – Multivariate Model, Version #2 (PLS Algorithm Results)

PLS ALGORITHM	VALUES
Path Coefficient (Contextual Overt Errors → Overall Translation Quality)	0.329997
Path Coefficient (Contextual Author Errors → Overall Translation Quality)	0.282516
Path Coefficient (Rank Order → Overall Translation Quality)	0.278485
Path Coefficient (Overall Translation Quality → Crowdsourced Translation Acceptance)	0.358453
Overall Translation Quality (OTQ)	
Cronbach's Alpha	0.959184
Composite Reliability	0.968412
Average Variance Extracted (AVE)	0.859804
R ²	0.591962
Crowdsourced Translation Acceptance (CTA)	
Cronbach's Alpha	0.644663
Composite Reliability	0.768738
Average Variance Extracted (AVE)	0.406458
R ²	0.128488

Here we see both reliability and validity issues for *Crowdsourced Translation Acceptance* (CTA). Clearly, the correlation between *Overall Translation Quality* and *Crowdsourced Translation Acceptance* does not sufficiently predict why people may or may not have accepted a crowdsourced translation. While cost may have been a factor here, the long idle times (cf. Table 4 - Approximate Time Measures) may have also been a factor. Testing the significance of these potential effects is beyond the scope of this study.

While *Contextual Usage Errors*, *Task Repetition*, and *Translation Diversity* were no longer significant after the testing of the first multivariate PLS path model (cf. Figure 22), a quick test was conducted to see if any of these dropped theoretical constructs could be applied to the results of the second multivariate PLS path model (cf. Figure 23). Of these three, the only theoretical concept that remained significant with respect to the new model (cf. Figure 24) was *Translation Diversity*. Therefore, Figure 25 illustrates the third multivariate PLS path model.

Figure 25 - Multivariate Model of Crowdsourced Translation, Version #3



The third multivariate PLS path model tests the significance of *Translation Diversity* as an effect upon *Contextual Overt Errors* and *Contextual Author Errors*. Together with *Rank Order*, these theoretical concepts are tested for significance in relation to *Overall Translation Quality*. Table 48 shows the bootstrapping results for the third multivariate PLS path model.

Table 48 – Multivariate Model, Version #3 (Bootstrapping Results)

BOOTSTRAPPING (128 samples with 500 iterations/degrees of freedom)		T-STATISTIC VALUES
Structural Model	Translation Diversity → Contextual Overt Errors	4.525777 (significant @ 0.0001)
	Translation Diversity → Contextual Author Errors	3.469408 (significant @ 0.01)
	Rank Order → Overall Translation Quality	4.174528 (significant @ 0.0001)
	Contextual Overt Errors → Overall Translation Quality	4.884271 (significant @ 0.0001)
	Contextual Author Errors → Overall Translation Quality	3.725835 (significant @ 0.0005)

Here we see statistically significant structural model loadings for these five theoretical concepts. This gives us a sense that this model may be a more accurate fit for our experimental data.

In Figure 26 and Table 49, we look to assess the reliability and validity of the third multivariate PLS path model.

Figure 26 - Multivariate Model, Version #3 (PLS Algorithm Results)

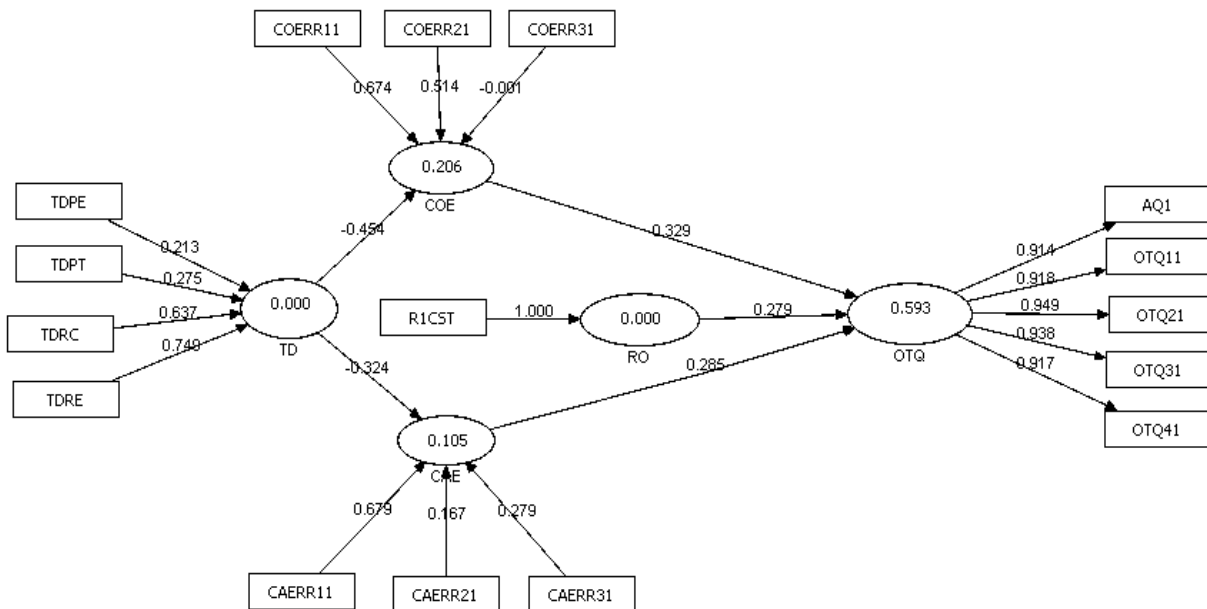


Table 49 – Multivariate Model, Version #3 (PLS Algorithm Results)

PLS ALGORITHM	VALUES
Path Coefficient (Translation Diversity → Contextual Overt Errors)	-0.454411
Path Coefficient (Translation Diversity → Contextual Author Errors)	-0.323763
Path Coefficient (Rank Order → Overall Translation Quality)	0.279223
Path Coefficient (Contextual Overt Errors → Overall Translation Quality)	0.329004
Path Coefficient (Contextual Author Errors → Overall Translation Quality)	0.285159
Overall Translation Quality (OTQ)	
Cronbach's Alpha	0.959184
Composite Reliability	0.968398
Average Variance Extracted (AVE)	0.859750
R ²	0.593215
Contextual Overt Errors (COE) R²	
	0.206490
Contextual Author Errors (CAE) R²	
	0.104823

In the third multivariate PLS path model, we see that it is both reliable and valid. Additionally, we see a similar result as was shown with the testing of **H_{3A}**. Too much *Translation Diversity* has a negative impact on *Contextual Overt Errors* and *Contextual Author Errors*. This model explains about 59.3% of the variance in *Overall Translation Quality*. *Translation Diversity* explains about 20.6% of the variance in *Contextual Overt Errors* and about 10.5% of the variance in *Contextual Author Errors*. The explanatory power of *Translation Diversity* in the third multivariate PLS path model, particularly with respect to *Contextual Author Errors*, is fairly low. In this case, almost 90% of the variance cannot be explained by this model.

What we can understand from testing these hypotheses and multivariate models is that correct terminology, correct grammar, and fluency are the most predictive factors for producing crowdsourced translations when considering *Overall Translation Quality*. *Task Repetition* does not appear to have a significant effect on the quality of crowdsourced translation; however, it appears to positively influence how people judge the quality of a crowdsourced translation once they know how the crowdsourced translations were created. Too much *Translation Diversity* seems to have a negative effect; we suspect that perhaps an inverted U-shaped curvilinear relationship may be present.

The stylistic issues proposed by House (1997) represented by *Contextual Usage Errors* do not appear to be significantly predictive for the *Overall Translation Quality* of crowdsourced translation. This finding is important because it implies that for crowdsourced translations, both participant groups (Language Professionals and Non-Language Professionals) performed roughly the same. This shows that for issues of style in crowdsourced translations, Non-Language Professionals did as well as the Language Professionals. Finally, issues related to trusting the translated content and feeling able to contact the company in one's own language based upon the quality of the translation continue to stand out as the key factors for predicting the quality of crowdsourced translation (Jorgensen, 2009; DePalma, Sargent, & Beninatto, 2006).

Chapter 7: Discussion

This study has several governmental, business, and research implications. From a governmental perspective, while the OCOL (2005) report focuses its view primarily on how the Canadian federal government should deal with the issue of official languages internally, the report also outlines several opportunities for government to support the indirect creation of Canadian French-language digital web content. This study represents a starting point for understanding crowdsourced translation as a means of bridging the Canadian Linguistic Digital Divide. The OCOL (2005) report calls for research and development activities aimed at the creation of technolinguistic tools from which all Canadians can stand to benefit. The results of this study show that this approach to translating content from Canadian English to Canadian French was not only feasible, but in some cases, it produced translations with quality on par with or exceeding existing alternatives; cf. Table 9. This study's implication for government is that more research and development should be done on a larger scale to assess how crowdsourced translation can help reach the government's goal of supporting the indirect creation of translated Canadian web content (OCOL, 2005).

Overall, we see that crowdsourced translation was most effective when performed by Language Professionals. The best quality translation of the ten that were studied was a crowdsourced translation created by Language Professionals. While this result is encouraging, it should be noted that much of what has been written about crowdsourcing in the popular press has gone largely untested. While this research seems to support the effectiveness of crowdsourced translation as a means of bridging the Canadian Linguistic Digital Divide, still more research needs to be done before the benefits of crowdsourced translation can be demonstrated conclusively.

This study has implications for Canadian business. In Phase I of this study, hundreds of certified professional translators and freelancers were contacted via email, yet only fifteen actually participated in producing these crowdsourced translations. This seems to confirm the observations of OCOL (2005) that there is a shortage of professional translators in Canada. In most cases, the Language Professionals declined to participate because they were simply too busy. For businesses, crowdsourced translation should not be seen as a panacea: it is unlikely to be appropriate for all types of Canadian business web content and finding Language Professionals presents a challenge. Poor quality translations, whether they are machine, professional or crowdsourced, represent a risk for online business. Crowdsourced translation, in the scope of this study, took much longer than

existing alternatives. This may be a deterrent for companies looking to have translations done cheaply, quickly and also with outstanding quality. Optimizing quality, time and cost is a difficult task to accomplish even for the best professional translators.

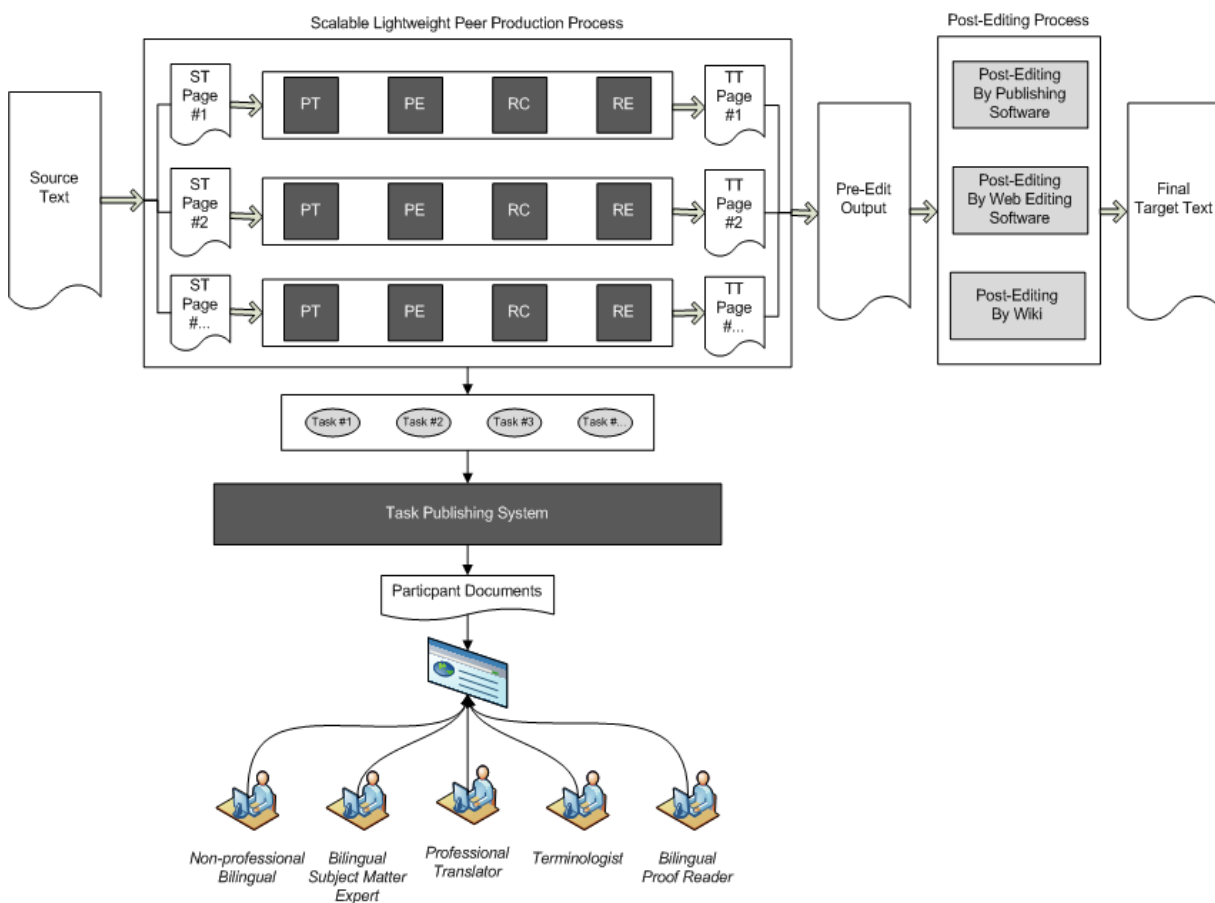
For companies that do not currently translate their website content, crowdsourced translation may be an acceptable alternative to doing nothing at all. This study shows in terms of quality that crowdsourced translation is a much better alternative to running website content through freely available machine translation websites. Additional research could be done with small-to-medium enterprises on a pilot project basis to see what impact using crowdsourced translation to translate their website content had on their business. This study opens the door to new avenues of business research. In this context, Collaborative Online Language Transfer could help to provide many important performance metrics.

This study has implications for research. To know more about the efficacy of crowdsourced translation within a Canadian context, more comprehensive testing could be conducted using a larger set of source texts, varying in style, length, and function. Translation between other Canadian language pairs could also be tested; for example, between French and Inuktitut or English and Cree. It is important to note that there are linguistic digital divides that exist apart from the two official languages of Canada (Désilets, 2007). The Collaborative Online Language Transfer process described in this study makes it technically easy to test these types of scenarios. Scaling the research in these directions is an obvious next step.

From a technical standpoint, while the Confluence software was able to handle a source text of approximately 360 words in length, it is reasonable to assume that the longer the length of the source text, the more arduous the Revision Creation and Revision Evaluation tasks would be. One of the principles of lightweight peer production is of course to have many tasks that are easy to perform and which do not take a long time to complete. To these ends, at the end of this study, a scalable lightweight peer production process is proposed. In essence, while the unit of translation used in this study is the paragraph, the unit of revision creation would need to be limited in accordance with the principles of lightweight peer production. It would be reasonable to set the unit of revision creation at approximately one page worth of content. In the scalable lightweight peer production process, multiple pages of the source text would be divided and the lightweight peer production process would then be run in parallel for each page. At the end of the scalable lightweight peer production process, each translated page would then be

reassembled into a pre-edit output document – in essence, a lightweight peer-produced translated version of the entire source text. Allan (2001) discusses the idea of *post-editing* as part of a translation workflow process that has been applied to machine translation. This idea of *post-editing* could also easily be adapted to the pre-edit output of the scalable lightweight peer production process for crowdsourced translation. Depending on the characteristics of the pre-edit output document, a variety of postediting software could be used to ensure the final quality of the document, from simple word processing, web editing software, or even as input to a collaborative translation wiki for *post-editing* by Language Professionals. Figure 27 depicts the scalable lightweight peer production process with post-editing.

Figure 27 - Scalable Lightweight Peer Production with Post-Editing



In this study, Amazon Mechanical Turk did not appear to be optimal within a Canadian context as a task publishing system. Part of this stems from the intrusive nature of the registration process for the participants. Part of it stems from participants only being able to be compensated in “Amazon dollars”. It should be noted that the nature of

Amazon Mechanical Turk is such that it is currently only possible to request work through the website when the requester of the work can provide United States tax information. As such, this represents a major barrier to entry for it to be used as a task publishing system within a Canadian context. Because the researcher enjoys both Canadian and American citizenship, this barrier was surmountable; for other Canadians wishing to duplicate the experiments described in this thesis, it presents a large technical limitation. More research should be done to identify or develop a suitable alternative to Amazon Mechanical Turk as a task publishing system for a broader Canadian audience. Lastly, another technical limitation of Amazon Mechanical Turk is the lack of an API to support worker applications. Currently, Amazon Mechanical Turk only supports an API for submitting work, but not for building thin clients with the purpose of enabling participants to perform work. This limitation means that in order for participants to do work, they have to use the Amazon Mechanical Turk website. If Amazon Mechanical Turk were to be replaced with a different task publishing system, the new system should include web services that allow platform-independent and language-independent thin clients to be developed. This could enable a richer and easier user experience for the participants and may be a driver for increased participation, likely leading to decreased idle times.

Determining how to increase participation in a Collaborative Online Language Transfer System should be a goal for future research. Haythornthwaite (2009) discusses *Recognition*, *Reputation*, and *Reward* as mechanisms for increasing participation in systems developed using a Lightweight Peer-Production model. The system developed for this study did not take advantage of these aspects (with the exception of the \$1.50 paid per task completed) and as such, this may have contributed to the lengthy idle times observed in the experiments. *Recognition* (Haythornthwaite, 2009) in the Collaborative Online Language Transfer system could be accomplished quantitatively by presenting to the participants the percentage of their work that has contributed to the production of target text. *Recognition* measures could be further broken down by task type (paragraph translation, paragraph evaluation, revision creation, and revision evaluation). The system could recognize other quantitative measures such as the amount of time that participants spent working in the system for a particular target text. *Reputation* (Haythornthwaite, 2009) in the Collaborative Online Language Transfer system could be accomplished by publishing lists of the best-performing participants among each of the different task types. For example, if a participant created a paragraph translation which was then evaluated as superior by peers, this information could be shown publically. Likewise, this idea could be applied to any of the other tasks. For example, if a participant shows good judgment when selecting different

revisions of the rough draft, that participant's reputation in this matter could be publically shown. *Reward* (Haythornthwaite, 2009) in the Collaborative Online Language Transfer system could be accomplished by taking into account a participant's reputation, whereby the system could grant to the participant additional privileges, such as having a broader selection of texts to work on or even allowing the participant to work on texts for specific causes or organizations that are more interesting. It is also easy to conceive of mechanisms in which well-performing participants are provided increased monetary rewards. Conversely, should a participant's reputation be poor for some types of tasks, perhaps the system could limit the participant's participation in ways that could decrease the diversity for those types of tasks; cf. Figure 10 - H3A (PLS Algorithm Results).

Hong and Page (2004) ask, "Can a functionally diverse group whose members have less ability outperform a group of people with high ability who may themselves be diverse?" Overall, the experiments performed in this study suffered slightly from too much diversity. However, when *Translation Diversity* was taken together with the other theoretical concepts, this observed effect became insignificant (cf. Table 45). Looking at how the crowdsourced translations did, it is clear that the Language Professionals outperformed the Non-Language Professionals both in terms of quality and ranking. The best quality and highest ranked translation of the ten that were studied was a crowdsourced translation that was created by Language Professionals. Perhaps coincidentally, this translation (CSTLP2) also had nine participants, which was higher than any of the other experiments. CSTLP4 had seven participants while CSTNLP2 and CSTNLP4 had six participants each. What stands out about the diversity of CSTLP2 is that it had lower Paragraph Translation Diversity and higher Paragraph Evaluation diversity than the other experiments (cf. Table 8 - Simpson's Reciprocal Index Measures). This finding suggests that the effect of diversity in crowdsourced translation should be understood in the context of each task type in the lightweight peer-production process. In future research, more data could be collected to test if lower diversity in Paragraph Translation and higher diversity in Paragraph Evaluation culminate in higher quality translations. In general, however, when viewing the ideas proposed by Hong and Page (2004) within the context of this study, it is clear that simply allowing for uncontrolled diversity is not a guarantee of a higher quality crowdsourced translation, nor is it a guarantee that a crowdsourced translation will be completed more quickly (cf. Table 4 - Approximate Time Measures) or with lower cost (cf. Table 5 - Cost Measurements).

Having anonymous sets of eyes engaged in peer-review of the crowdsourced translations during the Collaborative Online Language Transfer process seems to have eliminated some types of critical errors that would have adversely affected the quality of crowdsourced translation. While the specifics of the Phase I data were not discussed in detail, there were cases in which some of the participants had produced information that was clearly unusable. An example of this was within CSTNLP2 in which the same participant translated the wrong paragraph in three different instances. When it came time to move into the Paragraph Evaluation phase, the participants were able to quickly identify the problem and eliminate it. As was mentioned previously, paragraph translations that described the food at the Black Camel Café as unappetizing (“a chili made with the meat of a black camel”) were generally removed. Also, some of the discarded revised translations that included major punctuation mistakes were eliminated during the revision creation phase. These types of error would be critical were it not for the group’s *diverse, independent, decentralized, and aggregated* collective private judgments (Surowiecki, 2005). In summary, for the question of the effect of *Task Repetition* (Eagle, 2009; von Ahn, 2005), we do not see it as a significant predictor of the quality of a crowdsourced translation. However, *Task Repetition* does seem to be an important mechanism for eliminating critical translation and revision errors by multiple anonymous peer-reviewing participants.

An interesting observation that was made during the recruitment for this research relates to instances of non-language professionals sometimes engaging in translation. There is a “grey area” in which some bilingual Canadians are asked to translate during the course of their normal job duties by their employers because of their bilingualism. While these bilingual Canadians lack any translation education or translation training, they help to meet an otherwise unmet demand for translation. This phenomenon reflects the shortage of professional translators within Canada (OCOL, 2005). The need to bridge the Canadian Linguistic Digital Divide is strong and this observation about Non-Language Professionals engaged in translation service is indicative of a healthy and literate bilingual Canadian society. While the results of the study do not support the idea of Non-Language Professionals performing crowdsourced translation, the process of transferring written content between languages has long been in the hands of non-professional bilingual individuals (Garcia, 2009) despite the professionalization of translation in Canada.

7.1 Limitations and Future Research

Caution must be used when interpreting the results of this study. Most of the translation evaluators in Phase II were recruited from the National Capital Region of Canada

(Ottawa/Gatineau). While there is a high density of bilingual people in the National Capital Region, this study does not attempt to convey the idea that the beliefs and views of English to French Canadian translation represented by this study accurately reflect all bilingual Canadian's opinions about translation quality. Additionally, because only one source text was selected for study, this naturally restricts the way in which the results can be generalized to different types of texts or between different language pairs and directions. As such, this restriction presents a great opportunity for future research. For translation researchers, this study may allow them to take a more nuanced look at the dynamics of group translation, specifically with respect to the principles of lightweight peer production.

As this study shows strong support for the idea of having language professionals engaging in crowdsourced translation, at the conclusion of this study, an intriguing idea is to further develop the technology into a tool to help train translators. As a pedagogy tool used in translation schools, it may allow for distance learning in translation. Peer-reviewed and anonymously collected feedback could be provided to students in near-real time, and through practice, those students might benefit from the collective intelligence of other Language Professionals evaluating their work. Using this technology as a pedagogy tool may even help translation schools graduate higher numbers of trained professional translators, ultimately helping to reduce the shortage of professional translators in Canada and to create more bilingual content for the web. Whether used for pedagogy or for pure translation purposes, crowdsourced translation seems to offer much potential for helping to bridge the Canadian Linguistic Digital Divide.

7.2 Conclusion

Web 2.0 technologies have been catalysts for the emergence of online collaboration and peer production. These disruptive technologies are among the drivers for business model changes in many sectors, including the language service industry. This thesis described a framework in which the techniques of Lightweight Peer Production can be joined with the principles of Document Engineering for e-Business to enable a Collaborative Online Language Transfer System for crowdsourced translation. It is within the context of the Canadian Linguistic Digital Divide that this research is anchored.

This study attempted to shed light on the efficacy of crowdsourcing as a means of translating web content in Canada. Within, we sought to explore and understand if a model can be created that can estimate the effectiveness of crowdsourced translation as a means of bridging the Canadian Linguistic Digital Divide. To test our hypotheses and models, we

used structural equation modeling techniques coupled with confidence intervals for comparing experimental crowdsourced translation to both professional and machine translation baselines. Furthermore, we explored a variety of factors which influenced the quality of the experimental translations, how those translations performed in the context of their source text, and the ways in which the views of the quality of the experimental translations were measured before and after participants were made aware of how the experimental translations were created.

This study considered the work of both Language Professionals (translators) and Non-Language Professional (bilinguals). We found a wide acceptance of the idea of crowdsourced translation being performed by Language Professionals as a means of bridging the Canadian Linguistic Digital Divide. Furthermore, our study showed that this method of crowdsourced translation performed relatively well with respect to translation quality when compared with traditional translation methods and online machine translation methods. We found that Non-Language Professional translation evaluators are more accepting of crowdsourced translation than are their Language Professional counterparts. We hope that this research is a useful stepping stone in the area of lightweight peer produced translation and the Canadian Linguistic Digital Divide.

References

- Abramowitz, M. & Stegun, I. (1965). *Handbook of Mathematical Functions : with Formulas, Graphs and Mathematical Tables*, Dover Publications, Inc. New York.
- Albers, S. (2010). PLS and success factor studies in marketing. In V. E. Vinzi, W. W. Chin, J. Hensler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 409-425). Berlin: Springer. doi:10.1007/978-3-540-32827-8
- Allan, I., & Seaman, J. (2010). *Imputation Explanation : Find the best way to handle missing data in surveys*, American Society for Quality website, Retrieved 03/2011 from <http://asq.org/quality-progress/2010/07/statistics-roundtable/imputation-explanation.html>
- Allan, J. (2001). Postediting: an integrated part of a translation software program. *Language Interational magazine*, April 2001. (pp. 26-29)
- Alonso, O., Rose, D. E., & Stewart, B. (2008). Crowdsourcing for relevance evaluation. *SIGIR Forum*, 42(2), 9-15.
- Amazon Web Services LLC. (2007). *Amazon mechanical turk SDK for .NET*
- Barroso, C., Carrión, G. C., & Roldán, J. L. (2010). Applying maximum likelihood and PLS on different sample sizes: Studies on SERVQUAL model and employee behavior model. In V. E. Vinzi, W. W. Chin, J. Hensler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 427-447). Berlin: Springer. doi:10.1007/978-3-540-32827-8
- Bey, Y., Kageura, K., Boitet, C., & Marzari, F. (2008). Translating the DEMGOL etymological dictionary of Greek Mythology with the BEYTrans wiki. WikiSym '08, Porto, Portugal. (pp 1-12). Doi:10.1145/1822258.1822275
- Bowker, L. (2008). Official language minority communities, machine translation, and translator education: Reflections on the status quo and considerations for the future. *TTR: traduction, terminologie, rédaction*, 21(2), 15-61.
- Bowker, L. (2009). Can Machine Translation Meet the Needs of Official Language Minority Communities in Canada? A Recipient Evaluation. *Linguistica Antverpiensia* 8, 123-155.

Brabham, D. C. (2008). Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence: The International Journal of Research into New Media Technologies*, Vol.14, no.1, Pp.75-90, 2008, 14(1), 75.

Chin, W. W. (1998). Issues and Opinion on Structural Equation Modeling, *MIS Quarterly*, (22), pp vii-xvi.

Chin, W. W. (2010a). Bootstrap cross-validation indices for PLS path model assessment. In V. E. Vinzi, W. W. Chin, J. Hensler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 83-97). Berlin: Springer. doi:10.1007/978-3-540-32827-8

Chin, W. W. (2010b). How to write up and report PLS analyses. In V. E. Vinzi, W. W. Chin, J. Hensler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 655-690). Berlin: Springer. doi:doi:10.1007/978-3-540-32827-8

Clavet, A. (2002). French on the Internet: Key to the Canadian Identity and the Knowledge Economy. Follow-Up Study by the Commissioner of Official Languages. Ottawa, Minister of Public Works and Government Services Canada. Retrieved 05/2011 from http://www.ocolclo.gc.ca/docs/e/fr_Internet_id_can-2002_e.pdf

Crenna, D., & Diller, J. (2010). *TweetSharp*. Retrieved from <https://github.com/danielcrenna/tweetsharp> on 05/2010

Cudeck, R., & Henley, J. J. (2003). A realistic perspective on pattern representation in growth data: Comment on bauer and curran (2003). *Psychological Methods*, 8, 378-383.

Czapski, J. (2008). *Balanced triple elimination tournament - 4 players*. Retrieved 10/23, 2009, from <http://www.tournamentdesign.org/td/czap3e4pv2.pdf>

DePalma, D. A., Sargent, B. B. & Beninatto, R. S. (2006). *Can't read, won't buy: Why language matters on global websites*. Retrieved 06/22, 2009, from http://www.lionbridge.com/lionbridge/en-US/kc/globalization/cant_read_wont_buy.htm

Désilets, A. (2007). Translation Wikified: How will Massive Online Collaboration Impact the World of Translation? Opening keynote for Translating and the Computer 29, Nov 29-30, 2007. London, UK.

Dijkstra, T. K. (2010). Latent variables and indices: Herman wold's basic design and partial least squares. In V. E. Vinzi, W. W. Chin, J. Hensler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 23-46). Berlin: Springer. doi:10.1007/978-3-540-32827-8

dotSUB LLC. (2009). *dotSub | any video, any language*. Retrieved 09/06, 2009, from <http://dotsub.com/>

Eagle, N. (2009). Txteagle: Mobile crowdsourcing. *HCI (14)*, 447-456.

Fornell, C. & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement error, *Journal of Marketing Research*, 18, pp. 39-50.

Garcia, I. (2009). Beyond translation memory: Computers and the professional translator. *Journal of Specialised Translation*, (12), 199-214.

Glushko, R. J., & McGrath, T. (2002). Document engineering for e-business. *DocEng '02: Proceedings of the 2002 ACM Symposium on Document Engineering*, McLean, Virginia, USA. 42-48.

GraphPad Software (2011), P Value Calculator, Retrieved 02/12, 2011 from <http://www.graphpad.com/quickcalcs/PValue1.cfm>,

Guyon, A. (2010). *The ups and downs of online collaborative translation*. Retrieved 05/25, 2010, from <http://www.btb.gc.ca/btb.php?lang=eng&cont=1536>

Haythornthwaite, C. (2009). Crowds and communities: Light and heavyweight models of peer production. *Hawaii International Conference on System Sciences*, 0, 1-10.

Hensler, J., & Fassot, G. (2010). Testing moderating effects in PLS path models: An illustration of available procedures. In V. E. Vinzi, W. W. Chin, J. Hensler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 713-735). Berlin: Springer. doi:doi:10.1007/978-3-540-32827-8

Henseler, J., Ringle, C., & Sinkovics, R. (2009). *The Use of Partial Least Squares Path Modeling in International Marketing*. *New Challenges to International Marketing Advances in International Marketing*, Volume 20, 277-319, ISSN: 1474-7979/doi:10.1108/S1474-7979(2009)0000020014

- Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, 101(46), 16385-16389.
- House, J. (1977). A model for assessing translation quality. *Meta: Translators' Journal*, 22(2), 103-109.
- House, J. (1997). *Translation Quality Assessment: A Model Revisited*. Tübingen: Gunter Narr Verlag.
- Howe, J. (2008). The wisdom of the crowd resides in how the crowd is used. *Nieman Reports*, 62(4), 47.
- Huberdeau, L.-P., Désilets, A., & Paquet, S. (2008). *The Cross-Lingual Wiki Engine: Enabling Collaboration Across Language Barriers*. National Research Council of Canada.
- Huberman, B. A., Romero, D. M., & Wu, F. (2008). Crowdsourcing, attention and productivity. *SSRN eLibrary*.
- Hulland, J., Ryan, M. J., & Rayner, R. K. (2010). Modeling customer satisfaction: A comparative performance evaluation of covariance structure analysis versus partial least squares. In V. E. Vinzi, W. W. Chin, J. Hensler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 307-325). Berlin: Springer. doi:10.1007/978-3-540-32827-8
- Hutchins, J. (2001). Machine translation and human translation: In competition or in complementation? *International Journal of Translation*, 13(1-2), 5-20.
- IBM. (2010). *IBM center for social software: Projects: N.fluent*. Retrieved 11/24, 2009, from http://www.research.ibm.com/social/projects_nfluent.html
- Jorgensen, J. (2009). Talk about a tough sell. *World Trade*, 22(8), 47.
- Kaiser, M., & Lowe, J. (2008). Creating a research collection of question-answer sentence pairs with Amazon's mechanical turk.
- Kittur, A., Chi, E. H., & Suh, B. (2008). Crowdsourcing user studies with mechanical turk. *CHI '08: Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems*, Florence, Italy. 453-456.

Kocbek, A. (2006) Language and culture in international legal communication, *Managing Global Transitions: International Research Journal* 4(3), 231- 247.

Lakhani, K. R., & Panetta, J. A. (2007). The principles of distributed innovation. *Innovations: Technology, Governance, Globalization Summer*, 2(3), 2009-06-21. Retrieved from <http://ssrn.com.proxy.bib.uottawa.ca/paper=1021034>

Ledlie, J., Odero, B., Minkov, E., Kiss, I., & Polifroni, J. (2009). Crowd translator : On building localized speech recognizers through micropayments. *Nokia Research Center*

Lee, A. S., & Nickerson, J. V. (2010). Theory as a case of design: Lessons for design from the philosophy of science. *Hawaii International Conference on System Sciences*, 0, 1-8.

Lord, B. (2008). Report on the Government of Canada's Consultations on Linguistic Duality and Official Languages. Report submitted to the Minister of Canadian Heritage, Status of Women and Official Languages. Retrieved on 05/2011 from http://www.pch.gc.ca/pc-ch/conslttn/lo-ol_2008/lord/index-eng.cfm

Malone, T. W., Crowston, K., & Herman, G. A. (2003). *Organizing business knowledge: The MIT process handbook* (1st ed.) The MIT Press.

Nord, C. (2006). Loyalty and Fidelity in Specialized Translation. *Journal of Scientific and Technical Translation*, (4), 29-41.

Nunnally, J. C., & Bernstein. I. (1994). *Psychometric Theory*. 3rd ed. New York, NY: McGraw Hill.

OCOL. (2005). *Bridging the digital divide: Official languages on the internet* No. SF31-78/2005). Ottawa, Ontario: Office of the Commissioner of Official Languages, Minister of Public Works and Government Services Canada.

Poltrock, S., & Handel, M. (2009). Modeling collaborative behavior: Foundations for collaboration technologies. *Hawaii International Conference on System Sciences*, 0, 1-10.

Porter, M. E. (2008). In *Harvard Business Review* (Ed.), *On competition: Updated and expanded edition* (11th ed.). Boston, Massachusetts: Harvard Business School Publishing Corporation.

Prieto Ramos, F. (2002). "Beyond the Confines of Literality: A Functionalist Approach to the Sworn Translation of Legal Documents.", Retrieved on 04/29/2011 from <http://www.ugr.es/~greti/puentes/puentes2/04-articulo.pdf>

Roberts, N., & Thatcher, J. (2009). Conceptualizing and testing formative constructs: Tutorial and annotated example. *SIGMIS Database*, 40(3), 9-39. doi:<http://doi.acm.org.proxy.bib.uottawa.ca/10.1145/1592401.1592405>

Ringle, C., Wende, S., & Will, S. (2005). *Smart PLS 2.0 (M3) beta*

Snow, R., O'Connor, B., Jurafsky, D., & Ng, A. (2008). Cheap and fast -- but is it good? evaluating non-expert annotations for natural language tasks. *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, 254-263.

Su, Q., Pavlov, D., Chow, J., & Baker, W. C. (2007). Internet-scale collection of human-reviewed data. *WWW '07: Proceedings of the 16th International Conference on World Wide Web*, Banff, Alberta, Canada. 231-240.

Sun Microsystems. (2008). *The OpenOffice localization project*. Retrieved 09/06, 2009, from <http://l10n.openoffice.org/>

Surowiecki, J. (2005). *The wisdom of crowds : Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations* (First ed.). New York: Anchor Books.

TED Conferences, L. (2009). *The TED open translation project*. Retrieved 09/06, 2009, from <http://www.ted.com/index.php/OpenTranslationProject>

Temme, D., Kreis, H., & Hildebrandt, L. (2010). A comparison of current PLS path modeling software: Features, ease-of-use, and performance. In V. E. Vinzi, W. W. Chin, J. Hensler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 737-756). Berlin: Springer. doi:10.1007/978-3-540-32827-8

txteagle, L. (2009). *Txteagle | mobile crowdsourcing*. Retrieved 09/06, 2009, from <http://www.txteagle.com/>

Van Den Ende, J. (2008). *Organising innovation*. ERIM Report Series Reference No. EIA-2008-034-ORG. Available at SSRN: <http://ssrn.com/abstract=1376146>

Vinzi, V. E., Chin, W. W., Hensler, J., & Wang, H. (2010). Editorial: Perspectives on partial least squares. In V. E. Vinzi, W. W. Chin, J. Hensler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 1-20). Berlin: Springer. doi:10.1007/978-3-540-32827-8

Vinzi, V. E., Trinchera, L., & Amato, S. (2010). PLS path modeling: From foundations to recent developments and open issues for model assessment and improvement. In V. E. Vinzi, W. W. Chin, J. Hensler & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods and applications* (pp. 47-82). Berlin: Springer. doi:10.1007/978-3-540-32827-8

von Ahn, L. (2005). Human computation. Ph.D Dissertation in Computer Science, Carnegie Mellon University.

Williams, M. (2004). *Translation quality assessment: An argumentation-centered approach*. Ottawa, Ontario: University of Ottawa Press.

Appendix A – Confluence Software

This section describes, at a high-level, the software tool purpose-developed for this thesis. This section describes the major components of the software, provides screenshots of the participant screens, describes the types of raw data that are collected, the integration with Twitter, and the ways in which various levels of task repetition are handled. The software tool is referred to as *Confluence*.

A.1 Major Components

The *Confluence* software tool is subdivided into two separate applications: *Confluence Client* and *Confluence Host*. The purpose of this division is simply to separate the overall Collaborative Online Language Transfer process from the *Confluence Client* application. Separating the software in this way means a separate user interface using different technologies could easily be created in the future (say using HTML5, Silverlight or Flash). Both the client and the host software were written using C#, the 3.5 .NET Framework, and Visual Studio 2008. Windows Communication Foundation was used as the communications technology that connects the two applications. Two freely available third-party APIs were used in the application, most notably the Amazon Mechanical Turk API for .NET (Amazon Web Services LLC, 2007) and the TweetSharp API for Twitter (Crenna & Diller, 2010). Persistence, which for non-software engineers refers to the means by which data is saved, was achieved in the application by using XML Serialization – this approach conforms to the general principles of Document Engineering for E-Business and is consistent with Figure 4.

Additionally, the system uses a series of XML template documents to present the work that the participants would see on the Amazon Mechanical Turk website. This approach allows *Confluence* to separate the presentation of the tasks from the underlying Collaborative Online Language Transfer process. In short, should the tool be used in the future for other research projects, the XML template documents could be rewritten to support other language pairs or usability/content requirements without affecting the way *Confluence* works. More details regarding the XML templates and specific types of raw XML data generated by *Confluence* are described later in this appendix.

A.2 Confluence Host

The *Confluence Host* application is generally devoid of visible features: its main purpose is to separate the overall Collaborative Online Language Transfer process from the

Confluence Client application. Among the visible features in the *Confluence Host* application are the ability to start and stop communication between the *Confluence Host* application and the *Confluence Client* application as well as a small screen to facilitate configuration and communication with Twitter. Using technical vernacular, the *Confluence Host* provides self-hosted WCF services which are then used by the *Confluence Client*. This approach makes it very lightweight to deploy. More details regarding Twitter and what participants would see is described later in this appendix.

A.3 Confluence Client

The *Confluence Client* application is composed of five main screens, each depicting a particular aspect of the overall Collaborative Online Language Transfer process. The screens are *Instruction Stage*, *Paragraph Translation Stage*, *Paragraph Evaluation Stage*, *Revision Creation Stage*, and *Revision Evaluation Stage*. With the exception of the *Instruction Stage*, the aspects of the overall Collaborative Online Language Transfer process have been depicted in Figure 4.

The *Confluence Client* application allows the user to set a task repetition value of two, four, or eight, and to set a monetary reward per task. A unique document ID is provided in editable form; this document ID is stored with the raw XML data generated by the *Confluence Host* application.

The content of the source text to translate can be pasted into the source text box. Doing so provides some statistics in the Translation Settings pane, such as the number of words, paragraphs, an estimated number of unique tasks (# of HITs), the estimated Amazon Mechanical Turk fee, an estimated total cost for the translation and an estimated future balance. The Budget and Future Balance portions of the screen are calculated using the prepaid balance on Amazon Mechanical Turk. An invitation code could be entered in the software and also a checkbox which could be used to limit participation to only those whom have registered with a Canadian address through Amazon.com and the Mechanical Turk website. When the overall Collaborative Online Language Transfer process begins, before participants can engage in work, they must first identify the participant code on the Amazon Mechanical Turk website as a qualification. In this qualification screen, *Confluence* creates 20 total possible choices, with the invitation code being one possibility out of twenty. These mechanisms attempt to limit participants to people living in Canada with prior knowledge of the invitation code. Lastly, before starting the translation, a language pair and package can be selected; this allows *Confluence* to use a specific set of XML templates for presenting the

work to the participants. The Document Engineering approach to the development of Confluence is what provides for its future flexibility.

Statistics about the overall process, such as the total time of the process, the number of tasks completed, the average time participants worked on each task and the raw translation costs up to that point were available for display within the client software. At the conclusion of the Paragraph Evaluation phase, a rough draft of the translated source text is presented to the user. At the conclusion of the Revision Creation phase, a final draft of the translated source text is presented to the user as the Target Text. The target text is the final output of the Collaborative Online Language transfer process. In addition to the target text, Confluence generates many different types of raw data during the Collaborative Online Language Transfer process.

A.4 Types of Raw Data Generated by Confluence

Confluence generates eleven unique XML raw data files as part of the overall Collaborative Online Language Transfer process. In accordance with the Research Ethics Board guidelines, the raw XML data generated from the experiments is not included in this thesis because all results may only be published in an aggregate manner in order to protect the identities of the participants.

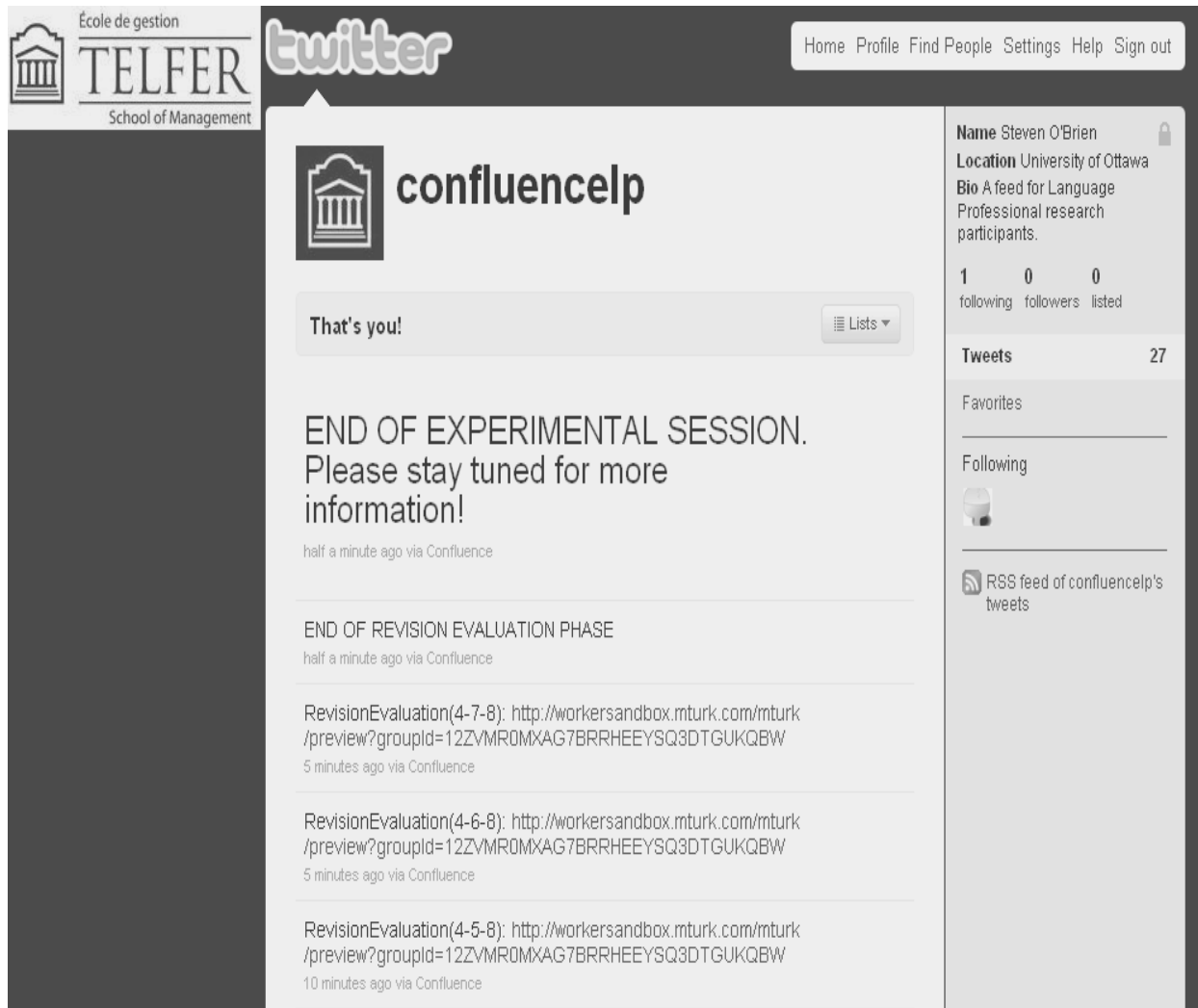
Figure 28 - Confluence Data Files

DATA FILE	DESCRIPTION
TRRECIPE.xml	Contains all of the initial instructions and configuration for the overall translation process.
PTBATTER.xml	Represents the input document to the Paragraph Translation process.
PTDISH.xml	Represents the output document from the Paragraph Translation process.
PEBATTER.xml	Represents the input document to the Paragraph Evaluation process.
PEDISH.xml	Represents the output document from the Paragraph Evaluation process.
ROUGHDraft.xml	Represents the rough draft of the translation.
RCBATTER.xml	Represents the input document to the Revision Creation process.
RCDISH.xml	Represents the output document from the Revision Creation process.
REBATTER.xml	Represents the input document to the Revision Evaluation process.
REDISH.xml	Represents the output document from the Revision Evaluation process.
TARGETTEXT.xml	Represents the final output of the Translation process.

A.5 Twitter Integration

As previously indicated, *Confluence* used Twitter to notify participants when tasks are made available to the Amazon Mechanical Turk website. *Confluence* also used Twitter to report on the status of the overall Collaborative Online Language Transfer process. For the purposes of this study, two different Twitter accounts were established to support the informational needs of each participant group in the Phase I (Translation Creation). Upon qualifying for a research participant group, the participants were asked to follow these Twitter accounts for details about when to participate. Participants were also notified via email when tasks were available.

Figure 29 - Sample Twitter Feed



The image shows a screenshot of a Twitter profile for the account 'confluencelp'. The profile header includes the account name, a bio stating 'A feed for Language Professional research participants', and statistics: 1 following, 0 followers, and 0 listed. The main content area shows a tweet from 'half a minute ago via Confluence' that reads: 'END OF EXPERIMENTAL SESSION. Please stay tuned for more information!'. Below this are three more tweets, each starting with 'END OF REVISION EVALUATION PHASE' and followed by a URL: 'RevisionEvaluation(4-7-8): http://workersandbox.mturk.com/mturk/preview?groupId=12ZVMR0MXAG7BRRHEEYSQ3DTGUKQBW'. The right sidebar shows the user's name 'Steven O'Brien', location 'University of Ottawa', and a link to an RSS feed for the account's tweets.

Upon viewing the research participant twitter feed, the participants were able to simply click on the provided link to go directly to the Amazon Mechanical Turk website where they could perform the work.

A.6 Amazon Mechanical Turk Screens

Here are examples of the main screens that the participants would see during the Translation Creation phase. The first screen represents the invitation by Amazon Mechanical Turk for the participant to qualify to work on the tasks.

Figure 30 - Pre-Qualification Screen

The screenshot displays the Amazon Mechanical Turk interface. At the top left is the logo for Amazon Mechanical Turk, with the text 'Artificial Intelligence' below it. To the right of the logo are three navigation tabs: 'Your Account', 'HITS', and 'Qualifications'. Further right, it says '45,916 HITS available now'. Below the navigation is a search bar with the text 'Search for Qualifications containing confluence' and a 'GO' button. A section titled 'What is a Qualification?' explains that some HITs require specific qualifications and that users can obtain them by browsing or searching. Below this is a section titled 'Qualifications containing 'confluence'' showing '1-1 of 1 Results'. The search results are sorted by 'Qualification Name (A-Z)'. The first result is 'Research Participant Qualification @ 16/06/2010 7:44:16 PM' with a 'Take the Qualification test' button. The details for this qualification are: Author: S OBRIEN, Qualified Users: 0, Description: 'If you have been invited to participate in this research, select the correct invitation code from the available options. If you make a mistake, you will be able to retry in one week.', and Retake Delay: 7 days.

Figure 31 - Participant Qualification Acceptance

amazonmechanical turk Artificial Intelligence Your Account HITS Qualifications 45,918 HITS available now

All Qualifications | Qualifications Assigned To You | Pending Qualifications

Search for containing | that pay at least \$ for which you are qualified

Research Participant Qualification @ 16/06/2010 7:44:16 PM
 Author: S OBRIEN
 Retake Delay: 7 days Qualification Value: 0

You are required to complete a test to obtain this qualification. Click on the continue link to proceed with the test. You will be notified by email when your qualification test has been graded by the Requester.

Figure 32 - Confluence Search Example

amazonmechanical turk Artificial Intelligence Your Account HITS Qualifications 45,881 HITS available now

All HITS | HITS Available To You | HITS Assigned To You

Search for containing that pay at least \$ for which you are qualified

HITS containing 'confluence'

1-1 of 1 Results

Sort by: Show all details | Hide all details

[Translate a paragraph \(en-CA --> fr-CA\)](#) [View a HIT in this group](#)

Requester: [S OBRIEN](#) HIT Expiration Date: Jun 19, 2010 (2 days 23 hours) Reward: \$0.01
 Time Allotted: 2 hours HITS Available: 3

Description: You will translate a paragraph into Canadian French. You are not allowed to submit/use machine translation. You have two hours to complete this task.

Keywords: [sobri094](#), [confluence](#), [translate](#)

Qualifications Required:

	Your Value	
Research Participant Qualification @ 16/06/2010 7:44:16 PM is 100	100	You meet this qualification requirement
Location is CA	CA	You meet this qualification requirement Contact the Requester of this HIT

Figure 33 - Paragraph Translation Task Example

Timer: 00:00:24 of 2 hours

Finished with this HIT? Let someone else do it?

Submit HIT

Return HIT

Total Earned: \$41.95
Total HITs Submitted: 448

Automatically accept the next HIT

Translate a paragraph (en-CA --> fr-CA).

Requester: S OBRIEN

Reward: \$0.01 per HIT

HITs Available: 3

Duration: 2 hours

Qualifications Required: Research Participant Qualification @ 16/06/2010 7:44:16 PM is 100, Location is CA

DESCRIPTION: For this task, you are asked to translate a Canadian English paragraph into Canadian French.

Statement of Informed Consent

Purpose of research study: To assess the use of Web 2.0 technologies to harness collective intelligence for translation as a means of bridging the Canadian Linguistic Digital Divide.

Benefits: Aside from monetary compensation, you will be given a chance to see the results of the study once they are published. Canadian society stands to benefit from this research because it may lead to new ways of bridging the Canadian linguistic digital divide.

Risks: There are no risks for participating in this study.

Voluntary participation: You may stop participating at any time without penalty by clicking on the "Return HIT" button, or by closing your browser window.

Confidentiality: The only identifying information kept about you will be a unique WorkerID and your IP address. This information shall not be disclosed.

Questions/concerns: You may e-mail questions to the principle investigator, Steven O'Brien or his thesis supervisor, Dr. Ajax Persaud. If you have any questions regarding the ethical conduct of the study, you may contact the Protocol Officer for Ethics in Research, University of Ottawa.

Clicking on the "Accept HIT" button indicates that you understand the information in this consent form. You have not waived any legal rights you otherwise would have as a participant in a research study.

Source Paragraph to Translate:

Also on the menu is a Black Camel chili (\$3.25 for 6 oz., \$5.25 for 12 oz.) and a variety of multi-roast blend coffee and loose-leaf tea selections from the nearby House of Tea. For breakfast, there's an eggspresso sandwich served until 11 am. Eggs are cooked a little different than usual (they use an espresso machine steam wand) and paired with either pancetta ham, steak or house-cured salmon, sweet pepper and red onion mayo). After lining up for around 5 minutes we order lunch. First is the slow roasted beef brisket with the signature Black Camel BBQ sauce and a slathering of chipotle mayo and creamy horseradish.

Source Text

The Black Camel has been slingin' sublime sandwiches since it opened in 2004 and their approach is as smart as it is simple. Starting with five basic sandwich choices, they allow you to personalize your meal by choosing from a range of high-quality in house prepared sauces and toppings. Options include slow roasted beef brisket, pulled pork shoulder, seared steak, roasted pulled chicken and roma tomato, red pepper, arugula and eggplant. Prices start at \$7.00 (with one sauce). There are also plenty of sauces and condiments and extra toppings available for less than a buck. So popular are these sandwiches that Black Camel uses their Twitter to let people know when they've sold out which, on most days, is before their scheduled closing time.

Also on the menu is a Black Camel chili (\$3.25 for 6 oz., \$5.25 for 12 oz.) and a variety of multi-roast blend coffee and loose-leaf tea selections from the nearby House of Tea. For breakfast, there's an eggspresso sandwich served until 11 am. Eggs are cooked a little different than usual (they use an espresso machine steam wand) and paired with either pancetta ham, steak or house-cured salmon, sweet pepper and red onion mayo). After lining up for around 5 minutes we order lunch. First is the slow roasted beef brisket with the signature Black Camel BBQ sauce and a slathering of chipotle mayo and creamy horseradish.

The huge and heavy sandwich is loaded with super tender beef brisket (slow-cooked for 12 hours) that is browned with molasses and caramelized onion and pungent with the flavour of the BBQ sauce and the mild zing of creamy horseradish. It's amazing and incredibly filling. We also decide to try the roasted pulled chicken with roasted sweet red pepper. It's moist and equally tasty although lighter than the imposing beef brisket sandwich. Both sandwiches come with generous helpings of meat and sauce and it takes a monumental effort to keep everything together in that large bun. Seating is extremely limited and it can get pretty crowded inside so it's best to try and get seating outside or find a shady spot on any of the nearby park benches.

QUALITATIVE GOALS:

The source text was written for an educated middle class audience, using language that is appropriate in 2010. The source text was written to be read (and not to be spoken aloud), and it is intended to evoke action on the part of the reader. The source text was written in a casual style by a professional blogger.

S'il vous plaît traduire le paragraphe précédent en français canadien. Do not use machine translation!

Figure 34 - Paragraph Evaluation Task Example (abbreviated)

QUALITATIVE GOALS:

The source text was written for an educated middle class audience, using language that is appropriate in 2010. The source text was written to be read (and not to be spoken aloud), and it is intended evoke action on the part of the reader. The source text was written in a casual style by a professional blogger.

Paragraph 1

Également sur le menu est un piment noir de chameau (\$3.25 pour 6 onces., \$5.25 pour 12 onces.) et une série de multi-rôtissez les choix de thé de café et de feuille de mélange à partir de la Chambre voisine du thé. Pour le déjeuner, there' ; s qu'un sandwich à eggspresso a servi jusqu'à ce que 11 AM. Les oeufs sont un peu différents cuit que d'habitude (ils utilisent une baguette magique de vapeur de machine de café express) et appareillé avec du l'un ou l'autre jambon de pancetta, bifteck ou saumons maison-traités, poivron doux et oignon rouge mayo). Après l'alignement pendant environ 5 minutes nous commandons le déjeuner. Est d'abord la poitrine de boeuf rôtie lente avec de la sauce à BBQ de chameau de noir de signature et slathering du chipotle mayo et du raifort crémeux.

Source Paragraph:

Also on the menu is a Black Camel chili (\$3.25 for 6 oz., \$5.25 for 12 oz.) and a variety of multi-roast blend coffee and loose-leaf tea selections from the nearby House of Tea. For breakfast, there's an eggspresso sandwich served until 11 am. Eggs are cooked a little different than usual (they use an espresso machine steam wand) and paired with either pancetta ham, steak or house-cured salmon, sweet pepper and red onion mayo). After lining up for around 5 minutes we order lunch. First is the slow roasted beef brisket with the signature Black Camel BBQ sauce and a slathering of chipotle mayo and creamy horseradish.

Paragraph 2

Également dans le menu est un chili Camel noir (3.25 \$ pour 6 oz, 5.25 \$ pour 12 onces) et une variété de multi-roast café et thé feuillets mobiles sélections de la maison de thé à proximité de la fusion. Pour le petit déjeuner, il y a un sandwich eggspresso servi jusqu'à 11 h. Les oeufs sont cuites un peu différent de l'habituel (ils utilisent une baguette espresso machine à vapeur) et appairée avec soit pancetta jambon, steak ou maison-cured saumon, poivrons doux et mayo oignon rouge). Après la doublure pour environ 5 minutes nous commander déjeuner. La première est la poitrine de boeuf rôti lent avec la sauce barbecue de Camel noir de signature et une slathering de mayo chipotle et crémeux raifort.

Lequel de ces paragraphes est le mieux traduit? S.v.p., gardez en tête les <<Qualitative Goals>>.

- Le paragraphe 1 est la meilleure
- Le paragraphe 2 est la meilleure

QUALITATIVE GOALS: In English, indicate what was *wrong* with the paragraph that you **did not select**.

GRAMMATICAL ERRORS: In English, indicate what was *wrong* with the paragraph that you **did not select**.

LEXICAL ERRORS (spelling, punctuation, etc.): In English, indicate what was *wrong* with the paragraph that you **did not select**.

Finished with this HIT? Let someone else do it?

Submit HIT

Return HIT

Figure 35 - Revision Creation Task Example (abbreviated)

<p>Rough Draft Translation</p> <p>Le Camel noir a été élingage sandwichs sublime puisqu'il a ouvert en 2004 et leur approche est aussi intelligent que c'est simple. À partir de cinq choix de base sandwich, ils permettent de personnaliser votre repas en choisissant parmi une gamme de sauces préparées en maison de haute qualité et garnitures. Les options comprennent la poitrine de bœuf rôti lente, extraites d'épaule de porc, steak poêlé, poulet rôti extraites et tomates roma, poivrons rouges, arugula et aubergine. Prix de départ \$ 7.00 (avec une sauce). Il existe également beaucoup de sauces et condiments et Topping supplémentaire pour votre argent, moins d'un puisse. Si populaires sont ces sandwichs que Camel noir utilise leur Twitter pour informer les gens lorsqu'ils ont vendu qui, sur la plupart des jours, est avant leur heure de fermeture.</p> <p>Également dans le menu est un chili Camel noir (3.25 \$ pour 6 oz, 5,25 \$ pour 12 onces) et une variété de multi-roast café et thé feuilletés mobiles sélections de la maison de thé à proximité de la fusion. Pour le petit déjeuner, il y a un sandwich eggspresso servi jusqu'à 11 h. Les oeufs sont cuites un peu différent de l'habituel (ils utilisent une baguette espresso machine à vapeur) et appariée avec soit pancetta jambon, steak ou maison-cured saumon, poivrons doux et mayo oignon rouge). Après la doublure pour environ 5 minutes nous commander déjeuner. La première est la poitrine de bœuf rôti lent avec la sauce barbecue de Camel noir de signature et une slathering de mayo chipotle et crémeux raifort.</p> <p>Le sandwich énorme et lourd est chargé avec la poitrine de bœuf super offres (slow-cuit pendant 12 heures) et qui est browned avec MELASSE et oignon caramélisée et âcre avec la saveur de la sauce barbecue et le doux zing de raifort crémeux. Il est incroyable et incroyablement de remplissage. Nous avons également décider d'essayer le rôti extraites de poulet avec rôti poivrons rouges doux. Il est humide et tout aussi savoureux, bien que plus léger que l'imposant sandwich de poitrine de bœuf. Les deux sandwichs sont livrés avec helpings généreuses de la viande et sauce et il faut un effort monumental pour conserver tout ensemble dans cette grande bun. Assise est extrêmement limitée et il peut obtenir assez entassés à l'intérieur afin qu'il soit mieux d'essayer d'obtenir des places à l'extérieur ou de trouver un endroit ombragé sur n'importe quel les bancs de parc à proximité.</p>	<p>Source Text</p> <p>The Black Camel has been slinging sublime sandwiches since it opened in 2004 and their approach is as smart as it is simple. Starting with five basic sandwich choices, they allow you to personalize your meal by choosing from a range of high-quality in house prepared sauces and toppings. Options include slow roasted beef brisket, pulled pork shoulder, seared steak, roasted pulled chicken and roma tomato, red pepper, arugula and eggplant. Prices start at \$7.00 (with one sauce). There are also plenty of sauces and condiments and extra toppings available for less than a buck. So popular are these sandwiches that Black Camel uses their Twitter to let people know when they've sold out which, on most days, is before their scheduled closing time.</p> <p>Also on the menu is a Black Camel chili (\$3.25 for 6 oz., \$5.25 for 12 oz.) and a variety of multi-roast blend coffee and loose-leaf tea selections from the nearby House of Tea. For breakfast, there's an eggspresso sandwich served until 11 am. Eggs are cooked a little different than usual (they use an espresso machine steam wand) and paired with either pancetta ham, steak or house-cured salmon, sweet pepper and red onion mayo). After lining up for around 5 minutes we order lunch. First is the slow roasted beef brisket with the signature Black Camel BBQ sauce and a slathering of chipotle mayo and creamy horseradish.</p> <p>The huge and heavy sandwich is loaded with super tender beef brisket (slow-cooked for 12 hours) that is browned with molasses and caramelized onion and pungent with the flavour of the BBQ sauce and the mild zing of creamy horseradish. It's amazing and incredibly filling. We also decide to try the roasted pulled chicken with roasted sweet red pepper. It's moist and equally tasty although lighter than the imposing beef brisket sandwich. Both sandwiches come with generous helpings of meat and sauce and it takes a monumental effort to keep everything together in that large bun. Seating is extremely limited and it can get pretty crowded inside so it's best to try and get seating outside or find a shady spot on any of the nearby park benches.</p>
<p>QUALITATIVE GOALS: <i>(ensure text appears to come from single author)</i></p> <p>The source text was written for an educated middle class audience, using language that is appropriate in 2010. The source text was written to be read (and not to be spoken aloud), and it is intended evoke action on the part of the reader. The source text was written in a casual style by a professional blogger.</p>	
<p>Veillez réviser la traduction. Do not use machine translation!</p>	
<div style="border: 1px solid black; padding: 5px;"> <p>Le Camel noir a été élingage sandwichs sublime puisqu'il a ouvert en 2004 et leur approche est aussi intelligent que c'est simple. À partir de cinq choix de base sandwich, ils permettent de personnaliser votre repas en choisissant parmi une gamme de sauces préparées en maison de haute qualité et garnitures. Les options comprennent la poitrine de bœuf rôti lente, extraites d'épaule de porc, steak poêlé, poulet rôti extraites et tomates roma, poivrons rouges, arugula et aubergine. Prix de départ \$ 7.00 (avec une sauce). Il existe également beaucoup de sauces et condiments et Topping supplémentaire pour votre argent, moins d'un puisse. Si populaires sont ces sandwichs que Camel noir utilise leur Twitter pour informer les gens lorsqu'ils ont vendu qui, sur la plupart des jours, est avant leur heure de fermeture.</p> <p>Également dans le menu est un chili Camel noir (3.25 \$ pour 6 oz, 5,25 \$ pour 12 onces) et une variété de multi-roast café et thé feuilletés mobiles sélections de la maison de thé à proximité de la fusion. Pour le petit déjeuner, il y a un sandwich eggspresso servi jusqu'à 11 h. Les oeufs sont cuites un peu différent de l'habituel (ils utilisent une baguette espresso machine à vapeur) et appariée avec soit pancetta jambon, steak ou maison-cured saumon, poivrons doux et mayo oignon rouge). Après la doublure pour environ 5 minutes nous commander déjeuner. La première est la poitrine de bœuf rôti lent avec la sauce barbecue de Camel noir de signature et une slathering de mayo chipotle et crémeux raifort.</p> </div>	
<p>Finished with this HIT? Let someone else do it?</p> <p style="text-align: center;"> <input type="button" value="Submit HIT"/> <input type="button" value="Return HIT"/> </p>	

Figure 36 - Revision Evaluation Task Example (abbreviated)

QUALITATIVE GOALS:

The source text was written for an educated middle class audience, using language that is appropriate in 2010. The source text was written to be read (and not to be spoken aloud), and it is intended evoke action on the part of the reader. The source text was written in a casual style by a professional blogger.

Translation 1

Le Camel noir a été élingage sandwiches sublime puisqu'il a ouvert en 2004 et leur approche est aussi intelligent que c'est simple. À partir de cinq choix de base sandwich, ils permettent de personnaliser votre repas en choisissant parmi une gamme de sauces préparées en maison de haute qualité et garnitures. Les options comprennent la poitrine de bœuf rôti lente, extraites d'épaule de porc, steak poêlé, poulet rôti extraites et tomates roma, poivrons rouges, arugula et aubergine. Prix de départ \$ 7.00 (avec une sauce). Il existe également beaucoup de sauces et condiments et Topping supplémentaire pour votre argent, moins d'un puisque. Si populaires sont ces sandwiches que Camel noir utilise leur Twitter pour informer les gens lorsqu'ils ont vendu qui, sur la plupart des jours, est avant leur heure de fermeture.

Également dans le menu est un chili Camel noir (3,25 \$ pour 6 oz, 5,25 \$ pour 12 onces) et une variété de multi-roast café et thé feuilletés sélections de la maison de thé à proximité de la fusion. Pour le petit déjeuner, il y a un sandwich eggspresso servi jusqu'à 11 h. Les oeufs sont cuites un peu différent de l'habituel (ils utilisent une baguette espresso machine à vapeur) et appairée avec soit pancetta jambon, steak ou maison-cured saumon, poivrons doux et mayo oignon rouge). Après la doublure pour environ 5 minutes nous commander déjeuner. La première est la poitrine de bœuf rôti lent avec la sauce barbecue de Camel noir de signature et une slathering de mayo chipotle et crémeux raifort.

Le sandwich énorme et lourd est chargé avec la poitrine de boeuf super offres (slow-cuit pendant 12 heures) et qui est browned avec MELASSE et oignon caramélisée et âcre avec la saveur de la sauce barbecue et le doux zing de raifort crémeux. Il est incroyable et incroyablement de remplissage. Nous avons également décider d'essayer le rôti extraites de poulet avec rôti poivrons rouges doux. Il est humide et tout aussi savoureux, bien que plus léger que l'imposant sandwich de poitrine de boeuf. Les deux sandwiches sont livrés avec helpings généreuses de la viande et sauce et il faut un effort monumental pour conserver tout ensemble dans cette grande bun. Assise est extrêmement limitée et il peut obtenir assez entassés à l'intérieur afin qu'il soit mieux d'essayer d'obtenir des places à l'extérieur ou de trouver un endroit ombragé sur n'importe quel les bancs de parc à proximité.

Source Text

The Black Camel has been slinging sublime sandwiches since it opened in 2004 and their approach is as smart as it is simple. Starting with five basic sandwich choices, they allow you to personalize your meal by choosing from a range of high-quality in house prepared sauces and toppings. Options include slow roasted beef brisket, pulled pork shoulder, seared steak, roasted pulled chicken and roma tomato, red pepper, arugula and eggplant. Prices start at \$7.00 (with one sauce). There are also plenty of sauces and condiments and extra toppings available for less than a buck. So popular are these sandwiches that Black Camel uses their Twitter to let people know when they've sold out which, on most days, is before their scheduled closing time.

Also on the menu is a Black Camel chili (\$3.25 for 6 oz., \$5.25 for 12 oz.) and a variety of multi-roast blend coffee and loose-leaf tea selections from the nearby House of Tea. For breakfast, there's an eggspresso sandwich served until 11 am. Eggs are cooked a little different than usual (they use an espresso machine steam wand) and paired with either pancetta ham, steak or house-cured salmon, sweet pepper and red onion mayo). After lining up for around 5 minutes we order lunch. First is the slow roasted beef brisket with the signature Black Camel BBQ sauce and a slathering of chipotle mayo and creamy horseradish.

The huge and heavy sandwich is loaded with super tender beef brisket (slow-cooked for 12 hours) that is browned with molasses and caramelized onion and pungent with the flavour of the BBQ sauce and the mild zing of creamy horseradish. It's amazing and incredibly filling. We also decide to try the roasted pulled chicken with roasted sweet red pepper. It's moist and equally tasty although lighter than the imposing beef brisket sandwich. Both sandwiches come with generous helpings of meat and sauce and it takes a monumental effort to keep everything together in that large bun. Seating is extremely limited and it can get pretty crowded inside so it's best to try and get seating outside or find a shady spot on any of the nearby park benches.

Translation 2

Le Camel noir a été élingage sandwiches sublime puisqu'il a ouvert en 2004 et leur approche est aussi intelligent que c'est simple. À partir de cinq choix de base sandwich, ils permettent de personnaliser votre repas en choisissant parmi une gamme de sauces préparées en maison de haute qualité et garnitures. Les options comprennent la poitrine de bœuf rôti lente, extraites d'épaule de porc, steak poêlé, poulet rôti extraites et tomates roma, poivrons rouges, arugula et aubergine. Prix de départ \$ 7.00 (avec une sauce). Il existe également beaucoup de sauces et condiments et Topping supplémentaire pour votre argent, moins d'un puisque. Si populaires sont ces sandwiches que Camel noir utilise leur Twitter pour informer les gens lorsqu'ils ont vendu qui, sur la plupart des jours, est avant leur heure de fermeture.

Également dans le menu est un chili Camel noir (3,25 \$ pour 6 oz, 5,25 \$ pour 12 onces) et une variété de multi-roast café et thé feuilletés sélections de la maison de thé à proximité de la fusion. Pour le petit déjeuner, il y a un sandwich eggspresso servi jusqu'à 11 h. Les oeufs sont cuites un peu différent de l'habituel (ils utilisent une baguette espresso machine à vapeur) et appairée avec soit pancetta jambon, steak ou maison-cured saumon, poivrons doux et mayo oignon rouge). Après la doublure pour environ 5 minutes nous commander déjeuner. La première est la poitrine de bœuf rôti lent avec la sauce barbecue de Camel noir de signature et une slathering de mayo chipotle et crémeux raifort.

Le sandwich énorme et lourd est chargé avec la poitrine de boeuf super offres (slow-cuit pendant 12 heures) et qui est browned avec MELASSE et oignon caramélisée et âcre avec la saveur de la sauce barbecue et le doux zing de raifort crémeux. Il est incroyable et incroyablement de remplissage. Nous avons également décider d'essayer le rôti extraites de poulet avec rôti poivrons rouges doux. Il est humide et tout aussi savoureux, bien que plus léger que l'imposant sandwich de poitrine de boeuf. Les deux sandwiches sont livrés avec helpings généreuses de la viande et sauce et il faut un effort monumental pour conserver tout ensemble dans cette grande bun. Assise est extrêmement limitée et il peut obtenir assez entassés à l'intérieur afin qu'il soit mieux d'essayer d'obtenir des places à l'extérieur ou de trouver un endroit ombragé sur n'importe quel les bancs de parc à proximité.

Quelle traduction est la meilleure?

Le revision 1 est la meilleure

Le revision 2 est la meilleure

A.7 Template Files

Figure 37 - Hit Type Template

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <HITTypeTemplate xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:xsd="http://www.w3.org/2001/XMLSchema">
3   <PTTitle>Translate a paragraph (en-CA --&gt; fr-CA).</PTTitle>
4   <PTDescription>You will translate a paragraph into Canadian French. You are not allowed to submit/use machine translation. You have two hours to complete this task.</PTDescription>
5   <PTKeywords>sobri094 confluence translate</PTKeywords>
6   <PETitle>Evaluate two translated paragraphs (en-CA --&gt; fr-CA).</PETitle>
7   <PEDescription>You will evaluate two translated paragraphs. You must use your own judgement to select the best translation. You have two hours to complete this task.</PEDescription>
8   <PEKeywords>sobri094 confluence translate</PEKeywords>
9   <RCTitle>Revise a translation (en-CA --&gt; fr-CA).</RCTitle>
10  <RCDescription>You will revise a translation into Canadian French. You are not allowed to use submit/use machine translation. You have two hours to complete this task.</RCDescription>
11  <RCKeywords>sobri094 confluence translate</RCKeywords>
12  <RETTitle>Evaluate two translations (en-CA --&gt; fr-CA).</RETTitle>
13  <REDescription>You will evaluate two translations. You must use your own judgement to select the best translation. You have two hours to complete this task.</REDescription>
14  <REKeywords>sobri094 confluence translate</REKeywords>
15 </HITTypeTemplate>
16
```

Figure 38 - Answer Key Template

```
1 <AnswerKey xmlns="http://mechanicalturk.amazonaws.com/AWSMechanicalTurkDataSchemas/2005-10-01/AnswerKey.xsd">
2   <Question>
3     <QuestionIdentifier>question1</QuestionIdentifier>
4     <AnswerOption>
5       <SelectionIdentifier><l-INVITATION_CODE--></SelectionIdentifier>
6       <AnswerScore>100</AnswerScore>
7     </AnswerOption>
8   </Question>
9 </AnswerKey>
```

Figure 39 - Invitation Template

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <QuestionForm xmlns="http://mechanicalturk.amazonaws.com/AWSMechanicalTurkDataSchemas/2005-10-01/QuestionForm.xsd">
3   <Question>
4     <QuestionIdentifier>question1</QuestionIdentifier>
5     <QuestionContent>
6       <Text>What is the invitation code?</Text>
7     </QuestionContent>
8     <AnswerSpecification>
9       <SelectionAnswer>
10        <StyleSuggestion>radiobutton</StyleSuggestion>
11        <Selections>
12          <l-SELECTION_TEMPLATE-->
13        </Selections>
14      </SelectionAnswer>
15    </AnswerSpecification>
16  </Question>
17 </QuestionForm>
```

Figure 40 – Paragraph Translation Task Template

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <QuestionForm xmlns="http://mechanicalturk.amazonaws.com/AWSMechanicalTurkDataSchemas/2005-10-01/QuestionForm.xsd">
3  <Question>
4  <QuestionIdentifier>targetparagraph</QuestionIdentifier>
5  <IsRequired>true</IsRequired>
6  <QuestionContent>
7  <FormattedContent>
8  <![CDATA[
9  <-START_TEMPLATE-->
10 <table cellpadding="10" cellspacing="0">
11 <tr>
12 <td colspan="2">
13 <font color="Firebrick" size="2" face="Verdana">
14 <b>DESCRIPTION: </b> For this task, you are asked to translate a Canadian English paragraph into Canadian French.
15 </font>
16 </td>
17 </tr>
18 <tr>
19 <td bgcolor="Gainsboro" width="50%" valign="top">
20 <font face="Verdana" size="1">
21 <b><u>Statement of Informed Consent</u></b><br><br>
22 <b>Purpose of research study: </b> To assess the use of Web 2.0 technologies to harness collective intelligence for translation as a means of bridging the Canadian Linguistic Digital
23 Divide.<br><br>
24 <b>Benefits: </b> Aside from monetary compensation, you will be given a chance to see the results of the study once they are published. Canadian society stands to benefit from this
25 research because it may lead to new ways of bridging the Canadian linguistic digital divide.<br><br>
26 <b>Risks: </b> There are no risks for participating in this study.<br><br>
27 <b>Voluntary participation: </b> You may stop participating at any time without penalty by clicking on the "Return HIT" button, or by closing your browser window.<br><br>
28 <b>Confidentiality: </b> The only identifying information kept about you will be a unique WorkerID. This information shall not be disclosed.<br><br>
29 <b>Questions/concerns: </b> You may e-mail questions to the principle investigator, Steven O'Brien or his thesis supervisor, Dr. Ajax Persaud. If you have any questions regarding
30 the ethical conduct of the study, you may contact the Protocol Officer for Ethics in Research, University of Ottawa.<br><br>
31 <b>Clicking on the "Accept HIT" button</b> indicates that you understand the information in this consent form. You have not waived any legal rights you otherwise would have as a
32 participant in a research study.
33 </font>
34 </td>
35 <td bgcolor="Linen" width="50%" valign="top">
36 <font face="Verdana" size="1">
37 <b><u>Source Text</u></b><br><br>
38 <-SOURCE_TEXT-->
39 </font>
40 </td>
41 </tr>
42 <tr>
43 <td colspan="2">
44 <font color="Firebrick" size="1" face="Verdana">
45 <b>QUALITATIVE GOALS: </b><br><br>
46 </font>
47 </td>
48 </tr>
49 <tr>
50 <td colspan="2">
51 <font face="Verdana" size="2" color="DarkGreen">
52 <b>
53 S'il vous plaît traduire le paragraphe précédent en français canadien. Do not use machine translation!
54 </b>
55 </font>
56 </td>
57 </tr>
58 </table>
59 <-END_TEMPLATE-->
60 ]]>
61 </FormattedContent>
62 </QuestionContent>
63 <AnswerSpecification>
64 | <FreeTextAnswer />
65 </AnswerSpecification>
66 </Question>
67 </QuestionForm>

```

Figure 41 - Paragraph Evaluation Task Template (parts 1 and 2)

```

1  <?xml version="1.0" encoding="UTF-8"?>
2  <QuestionForm xmlns="http://mechanicalturk.amazonaws.com/AWSMechanicalTurkDataSchemas/2005-10-01/QuestionForm.xsd">
3  <Question>
4  <QuestionIdentifier>evaluate_paragraph</QuestionIdentifier>
5  <IsRequired>true</IsRequired>
6  <QuestionContent>
7  <FormattedContent><![CDATA[
8  <table cellpadding="10" cellspacing="0">
9  <tr>
10 <td colspan="3">
11 <font color="Firebrick" size="2" face="Verdana">
12 <b>DESCRIPTION: </b>
13 <br> * For this task, you are asked to choose the best translated paragraph.
14 <br> * Also, indicate what was <i>wrong</i> with the paragraph that you <b>did not select</b>.
15 </font>
16 </td>
17 </tr>
18 <tr>
19 <td valign="top" colspan="3">
20 <table cellpadding="2" cellspacing="0">
21 <tr>
22 <td valign="top" bgcolor="Gainsboro" width="50%">
23 <font face="Verdana" size="1">
24 <b><u>Statement of Informed Consent</u></b><br><br>
25 <b>Purpose of research study:</b> To assess the use of Web 2.0 technologies to harness collective intelligence for translation as a means of bridging the Canadian Linguistic
26 Digital Divide.<br><br>
27 <b>Benefits:</b> Aside from monetary compensation, you will be given a chance to see the results of the study once they are published. Canadian society stands to benefit
28 from this research because it may lead to new ways of bridging the Canadian linguistic digital divide.<br><br>
29 <b>Risks:</b> There are no risks for participating in this study.<br><br>
30 <b>Voluntary participation:</b> You may stop participating at any time without penalty by clicking on the "Return HIT" button, or by closing your browser window.<br><br>
31 <b>Confidentiality:</b> The only identifying information kept about you will be a unique WorkerID. This information shall not be disclosed.<br><br>
32 <b>Questions/concerns:</b> You may e-mail questions to the principle investigator, Steven O'Brien or his thesis supervisor, Dr. Ajax Persaud. If you have any questions
33 regarding the ethical conduct of the study, you may contact the Protocol Officer for Ethics in Research, University of Ottawa.<br><br>
34 <b>Clicking on the "Accept HIT" button</b> indicates that you understand the information in this consent form. You have not waived any legal rights you otherwise would have
35 as a participant in a research study.<br><br>
36 </font>
37 </td>
38 <td colspan="2">
39 <font face="Verdana" size="1">
40 <b><u>Source Text</u></b><br><br>
41 <!--SOURCE_TEXT-->
42 </font>
43 </td>
44 </tr>
45 </table>
46 </td>
47 </tr>
48 <tr>
49 <td colspan="3">
50 <font color="Firebrick" size="1" face="Verdana">
51 <b>QUALITATIVE GOALS:</b><br><!--GOALS-->
52 </font>
53 </td>
54 </tr>
55 <tr>
56 <td colspan="3">
57 <table border="0" style="width:100%; border-collapse: collapse;">
58 <tr>
59 <td style="width:33%; vertical-align: top;">
60 <font size="2" face="Verdana"><b>Paragraph 1</b></font><br>
61 <!--LEFTSEED-->
62 </td>
63 <td style="width:34%; vertical-align: top;">
64 <font size="2" face="Verdana"><b>Source Paragraph:</b></font><br>
65 <!--SOURCE_PARAGRAPH-->
66 </td>
67 <td style="width:33%; vertical-align: top;">
68 <font size="2" face="Verdana"><b>Paragraph 2</b></font><br>
69 <!--RIGHTSEED-->
70 </td>
71 </tr>
72 </table>
73 </td>
74 </tr>
75 <tr>
76 <td colspan="3">
77 <!--TRANSLATION_INSTRUCTION-->
78 <br>
79 <font face="Verdana" size="2" color="DarkGreen">
80 <b>Lequel de ces paragraphes est le mieux traduit? S.v.p., gardez en tête les &lt;&lt;Qualitative Goals&gt;&gt;.</b>
81 </font>
82 </td>
83 </tr>
84 <tr>
85 <td colspan="3">
86 <!--ANSWER_SPECIFICATION-->
87 <!--SELECTIONS-->
88 <!--SELECTIONS-->

```

Figure 42 - Paragraph Evaluation Task Template (part 3)

```

89 | | </SelectionAnswer>
90 | | </AnswerSpecification>
91 | </Question>
92 | <Question>
93 | | <QuestionIdentifier>qualitative_errors</QuestionIdentifier>
94 | | <IsRequired>true</IsRequired>
95 | | <QuestionContent>
96 | | | <FormattedContent><![CDATA[
97 | | | <!-- START_TEMPLATE-->
98 | | | <font face="Verdana" size="2" color="DarkGreen">
99 | | | | <b>QUALITATIVE GOALS:</b> In English, indicate what was <i>wrong</i> with the paragraph that you <b>did not select</b>.
100 | | | </font>
101 | | | <!-- END_TEMPLATE-->
102 | | | ]]></FormattedContent>
103 | | </QuestionContent>
104 | | <AnswerSpecification>
105 | | | <FreeTextAnswer/>
106 | | </AnswerSpecification>
107 | </Question>
108 | <Question>
109 | | <QuestionIdentifier>grammatical_errors</QuestionIdentifier>
110 | | <IsRequired>true</IsRequired>
111 | | <QuestionContent>
112 | | | <FormattedContent><![CDATA[
113 | | | <!-- START_TEMPLATE-->
114 | | | <font face="Verdana" size="2" color="DarkGreen">
115 | | | | <b>GRAMMATICAL ERRORS:</b> In English, indicate what was <i>wrong</i> with the paragraph that you <b>did not select</b>.
116 | | | </font>
117 | | | <!-- END_TEMPLATE-->
118 | | | ]]></FormattedContent>
119 | | </QuestionContent>
120 | | <AnswerSpecification>
121 | | | <FreeTextAnswer/>
122 | | </AnswerSpecification>
123 | </Question>
124 | <Question>
125 | | <QuestionIdentifier>lexical_errors</QuestionIdentifier>
126 | | <IsRequired>true</IsRequired>
127 | | <QuestionContent>
128 | | | <FormattedContent><![CDATA[
129 | | | <!-- START_TEMPLATE-->
130 | | | <font face="Verdana" size="2" color="DarkGreen">
131 | | | | <b>LEXICAL ERRORS (spelling, punctuation, etc.)</b> In English, indicate what was <i>wrong</i> with the paragraph that you <b>did not select</b>.
132 | | | </font>
133 | | | <!-- END_TEMPLATE-->
134 | | | ]]></FormattedContent>
135 | | </QuestionContent>
136 | | <AnswerSpecification>
137 | | | <FreeTextAnswer/>
138 | | </AnswerSpecification>
139 | </Question>
140 | </QuestionForm>

```

Figure 43 - Revision Creation Task Template

```

1 <?xml version="1.0" encoding="UTF-8"?>
2 <QuestionForm xmlns="http://mechanicalturk.amazonaws.com/AWSMechanicalTurkDataSchemas/2005-10-01/QuestionForm.xsd">
3 <Question>
4 <QuestionIdentifier>revision_creation</QuestionIdentifier>
5 <IsRequired>true</IsRequired>
6 <QuestionContent>
7 <FormattedContent>
8 <![CDATA[
9 <!--START_TEMPLATE-->
10 <table cellpadding="10" cellspacing="0">
11 <tr>
12 <td colspan="2">
13 <font color="Firebrick" size="2" face="Verdana">
14 <b>DESCRIPTION: </b>
15 <br> * For this task, you are asked to revise a translation.
16 <br> * When revising, please keep in mind the Qualitative Goals.
17 <br> * When revising, please ensure that the text appears to come from a single author.
18 </font>
19 </td>
20 </tr>
21 <tr>
22 <td bgcolor="Gainsboro" valign="top" colspan="2">
23 <font face="Verdana" size="1">
24 <b><u>Statement of Informed Consent</u></b><br><br>
25 <b>Purpose of research study:</b> <b> To assess the use of Web 2.0 technologies to harness collective intelligence for translation as a means of bridging the Canadian Linguistic Digital
26 Divide <br><br>
27 <b>Benefits:</b> <b> Aside from monetary compensation, you will be given a chance to see the results of the study once they are published. Canadian society stands to benefit from this
28 research because it may lead to new ways of bridging the Canadian linguistic digital divide.<br><br>
29 <b>Risks:</b> <b> There are no risks for participating in this study.<br><br>
30 <b>Voluntary participation:</b> <b> You may stop participating at any time without penalty by clicking on the "Return HIT" button, or by closing your browser window.<br><br>
31 <b>Confidentiality:</b> <b> The only identifying information kept about you will be a unique WorkerID. This information shall not be disclosed.<br><br>
32 <b>Questions/concerns:</b> <b> You may e-mail questions to the principle investigator, Steven O'Brien or his thesis supervisor, Dr. Ajax Persaud. If you have any questions regarding
33 the ethical
34 conduct of the study, you may contact the Protocol Officer for Ethics in Research, University of Ottawa.<br><br>
35 <b>Clicking on the "Accept HIT" button</b><b> indicates that you understand the information in this consent form. You have not waived any legal rights you otherwise would have as a
36 participant in a research study.<br><br>
37 </font>
38 </td>
39 </tr>
40 <tr>
41 <td bgcolor="Pink" width="50%" valign="top">
42 <font size="1" face="Verdana">
43 <b><u>Rough Draft Translation</u></b><br>
44 <!--ROUGH_DRAFT-->
45 </font>
46 </td>
47 <td bgcolor="Linen" width="50%" valign="top">
48 <font face="Verdana" size="1">
49 <b><u>Source Text</u></b><br>
50 <!--SOURCE_TEXT-->
51 </font>
52 </td>
53 </tr>
54 <tr>
55 <td colspan="2" valign="top">
56 <font color="Firebrick" size="2" face="Verdana">
57 <b>QUALITATIVE GOALS:</b> <b> <i>(ensure text appears to come from single author)</i></b><br><!--GOALS-->
58 </font>
59 </td>
60 </tr>
61 </table>
62 <!--TRANSLATION_INSTRUCTION-->
63 <br>
64 <font face="Verdana" size="2" color="DarkGreen">
65 <b>
66 Veuillez réviser la traduction. Do not use machine translation!
67 </b>
68 </font>
69 <!--END_TEMPLATE-->
70 ]]>
71 </FormattedContent>
72 </QuestionContent>
73 <AnswerSpecification>
74 <FreeTextAnswer>
75 <DefaultText><!--DEFAULT_TEXT--></DefaultText>
76 <NumberOfLinesSuggestion>20</NumberOfLinesSuggestion>
77 </FreeTextAnswer>
78 </AnswerSpecification>
79 </Question>
80 </QuestionForm>

```


Appendix B – Phase I (Translation Creation) Data

This appendix includes each of the six translations used in Phase II – Translation Evaluation, and also for each experiment the rough draft translation and the revisions that were rejected during the revision evaluation process. It also includes the two professional and two machine translations that were not selected to be studied as part of Phase II (Translation Evaluation). It should be noted for clarity that the researcher is not fluent in French and has no ability to assess the quality of these translations himself.

In future research, it may be possible to measure if increases in Task Repetition had any effect on the qualitative differences between the rough draft translation and the final target text. Testing this new hypothesis is well beyond the scope of this thesis.

B.1 Professional Translation (PT)

Le Black Camel crée de sublimes sandwichs depuis son ouverture en 2004 et son approche est aussi brillante que simple. À partir des cinq choix de sandwich de base, ils vous permettent de personnaliser votre repas en choisissant parmi une gamme de sauces et de garnitures de première qualité et faites maison. Parmi les choix offerts, on retrouve la pointe de poitrine de boeuf rôtie, l'épaule pique-nique effilochée, le bifteck grillé et le poulet rôti effiloché, ainsi que la tomate Roma, les piments rouges, la roquette et l'aubergine. Le prix le plus bas est de 7 \$ (comprend une sauce). Il y a également plusieurs sauces, condiments et garnitures en supplément pour moins d'un dollar. Ces sandwichs sont si populaires que Black Camel se sert de Twitter pour informer les gens lorsqu'ils ont épuisé leurs stocks, ce qui arrive, la plupart des jours, avant l'heure de fermeture.

On retrouve également sur le menu le chili Black Camel (3,25 \$ pour une portion de 6 oz et 5,25 \$ pour 12 oz), ainsi qu'une variété de mélanges de cafés à torréfaction multiples et de feuilles de thé provenant de la House of Tea, située à proximité. Pour le déjeuner, un sandwich oeuf-presso est offert jusqu'à 11 h. Les oeufs sont cuits d'une manière un peu différente qu'à l'habitude (en effet, ils se servent du bec à vapeur d'une machine à expresso) et sont accompagnés soit de jambon pancetta, de bifteck ou de saumon fumé maison, de mayonnaise aux piments doux et aux oignons rouges. Après avoir fait la file pendant près de cinq minutes, nous commandons notre repas. Nous choisissons d'abord la pointe de poitrine de boeuf rôtie lentement, avec la sauce BBQ signature de Black Camel, une grosse portion de mayonnaise chipotle et de sauce crémeuse au raifort.

L'énorme sandwich est rempli de pointe de poitrine de boeuf extraordinairement tendre (mijoté pendant 12 heures) bruni avec de la mélasse et des oignons caramélisés et aromatisé avec la saveur de la sauce BBQ et le léger piquant de la sauce crémeuse au raifort. Il est formidable et incroyablement remplissant. Nous avons également choisi d'essayer le poulet rôti effiloché avec des piments rouges rôtis. Il est aussi juteux que savoureux, quoique plus léger que l'imposant sandwich à la pointe de poitrine de boeuf. Les deux sandwichs sont servis avec des portions généreuses de viande et de sauce et il faut déployer des efforts monumentaux pour tout garder dans le grand pain. Le nombre de sièges est extrêmement restreint et il arrive que ce soit plutôt bondé à l'intérieur, alors il est préférable d'essayer d'obtenir une place à l'extérieur ou de trouver un endroit à l'ombre sur l'un des bancs de parc situés à proximité.

B.2 Crowdsourced Translation – Language Professionals – Task Repetition Value Two (CSTLP2)

Black Camel sert ses sublimes sandwiches depuis son ouverture en 2004 en se basant sur une approche aussi astucieuse que simple. À partir de cinq choix de sandwich de base, le client peut personnaliser son repas en ajoutant des sauces et des garnitures maison parmi l'éventail d'options offertes incluant : pointe de poitrine de boeuf cuite lentement, épaule de porc effilochée, steak grillé, poulet rôti effiloché avec tomates Roma, poivrons rouges, roquette et aubergine. Les prix commencent à 7 \$ (incluant une sauce). Plusieurs sauces, condiments et garnitures supplémentaires sont disponibles pour moins de 1 \$. Ces sandwiches sont si convoités que Black Camel utilise Twitter pour aviser ses clients lorsque l'offre est épuisée pour la journée ce qui, la plupart du temps, se produit avant l'heure de fermeture annoncée.

Également au menu, le chili Black Camel (3,25 \$ pour 6 oz, 5,25 \$ pour 12 oz) et un assortiment de mélanges de café torréfiés et une sélection de thés en feuille de la boutique House of Tea. Pour le déjeuner, un sandwich « eggspresso » est servi jusqu'à 11 h. Les oeufs sont cuits un peu différemment (ils utilisent le bec à vapeur d'une machine à espresso) et sont servis avec un choix de pancetta, steak ou saumon fumé, poivrons et mayonnaise aux oignons rouges. Nous faisons la file environ 5 minutes avant de commander notre dîner. Nous commençons par la pointe de poitrine de boeuf rôtie lentement avec la sauce BBQ emblématique de Black Camel recouverte de mayonnaise chipotle et de raifort crémeux.

Le lourd et volumineux sandwich est chargé de pointe de poitrine de boeuf très tendre (cuite lentement pendant 12 heures) qui est dorée avec de la mélasse et des oignons caramélisés et relevée par la saveur de la sauce BBQ et du léger piquant de raifort crémeux. C'est incroyable et très substantiel. Nous décidons d'essayer également le poulet rôti effiloché avec poivrons rouges doux rôtis. Il est tendre et tout aussi délicieux quoique plus léger que l'imposant sandwich de pointe de poitrine de boeuf. Les deux sandwiches sont servis avec une généreuse portion de viande et de sauce et requièrent un important effort pour garder les ingrédients à l'intérieur du grand pain. Le nombre de places est très limité et le restaurant peut devenir assez bondé. Il est alors préférable de rechercher une place assise à l'extérieur ou trouver un coin ombragé sur un des bancs de parc tout près.

B.3 Crowdsourced Translation – Language Professionals – Task Repetition Value Four (CSTLP4)

Le Black Camel propose ses merveilleux sandwiches depuis son ouverture en 2004, et son approche est simple et intelligente. On vous propose cinq sandwiches de base au choix, que vous pouvez personnaliser en choisissant parmi une gamme de sauces et de garnitures maison de haute qualité. Parmi ces options figurent de la poitrine de boeuf rôtie lentement, de l'épaule de porc effilée, du steak grillé, du poulet rôti effilé, et des tomates prune, des poivrons rouges, de la roquette et des aubergines. Les premiers prix débutent à 7,00 \$ (avec une sauce). Il y a aussi de nombreuses sauces, condiments et garnitures supplémentaires disponibles pour moins d'un dollar. Ces sandwiches sont si appréciés que le Black Camel se sert de Twitter pour dire aux gens qu'il ne reste plus de sandwiches, ce qui, la plupart du temps, arrive avant l'heure de fermeture prévue.

Au menu, vous trouverez le chili Black Camel (3,25 \$ pour 180 ml ou 5,25 \$ pour 360 ml) et une variété de café torréfié et de thé frais provenant du magasin House of Tea qui est tout près. Le déjeuner, un sandwich oeufpresso, est servi jusqu'à 11 h. Les oeufs sont cuits d'une manière unique (à la vapeur de la machine espresso) et sont servis avec du jambon pancetta, du steak ou du saumon mariné maison et une sauce à la mayonnaise, au poivron et à l'oignon rouge. Après avoir fait la queue pendant environ cinq minutes, nous avons commandé notre dîner. D'abord, il y avait la pointe de boeuf cuite lentement et servie avec sa sauce BBQ signature du Black Camel, et une bonne couche de mayonnaise à la chipotle et de raifort crémeux.

Le gros sandwich épais est garni de poitrine de boeuf tendre extra tendre (cuite lentement pendant 12 heures), dorée avec de la mélasse et des oignons caramélisés, relevé par une sauce barbecue piquante et une pointe de raifort crémeux. Il incroyablement bon, et vous remplit l'estomac. Nous avons également essayé le poulet rôti avec des poivrons rouges grillés. Il est moelleux et tout aussi savoureux que l'impressionnant sandwich de poitrine de boeuf, mais plus léger. Les deux sandwiches sont servis avec de généreuses portions de viande et de sauce, et c'est loin d'être facile de faire tout tenir dans le pain. Le nombre de places est extrêmement limité, et le restaurant peut être plein à l'intérieur, il est donc préférable de s'asseoir dehors ou de trouver une petite place à l'ombre sur l'un des bancs du parc tout proche.

B.4 Crowdsourced Translation – Non-Language Professionals – Task Repetition Value Two (CSTNLP2)

Le restaurant "Black Camel" sert des sandwiches exquis depuis son ouverture en 2004. Leur approche est à la fois simple et intelligente. Partant d'un choix de cinq sandwiches de base, vous pouvez personnaliser votre repas avec des sauces et des condiments de haute qualité préparés sur place. Les options comprennent de la poitrine de boeuf à la mijoteuse, de la viande de porc effilée, du steak ou du poulet grillé ainsi que des tomates roma, des poivrons rouges, de la roquette et de l'aubergine. Les prix commencent à 7.00\$, incluant une sauce. De plus, plusieurs sauces et condiments sont disponibles pour moins d'un dollar. Leurs sandwiches sont d'une telle popularité que le "Black Camel" utilise Twitter pour informer du moment où une sorte de sandwich n'est plus disponible, ce qui arrive fréquemment avant l'heure de fermeture.

Également au menu est un chili Black Camel (3,25 \$ pour 6 onces, \$ 5,25 pour 12 onces) et une variété de torrification de café ainsi qu'une sélection de thés à feuilles de la House of Tea à proximité. Pour le petit déjeuner, il y a un sandwich "Eggspresso" servi jusqu'à 11 heures. Les oeufs sont cuits un peu différent qu'à l'habitude (ils utilisent une buse à vapeur de machine à expresso) et combinés à un choix de jambon pancetta, de steak ou de saumon fumé maison, avec une mayo aux poivrons et oignons rouges. Après avoir patienté en ligne pendant environ 5 minutes, nous commandons le repas. En premier vient la poitrine de boeuf rôti lentement avec la sauce signature Black Camel BBQ et une couche de mayonnaise au chipotle et au raifort crémeux.

L'immense et ardue sandwich est chargé de poitrine de boeuf super tendre (à la cuisson lente pendant 12 heures) qui est dorée avec de la mélasse et des oignons caramélisés et piquante à la saveur de la sauce barbecue et le léger goût du raifort crémeux. C'est stupéfiant et incroyablement satiant. Nous décidons aussi d'essayer le poulet rôti tiré avec poivrons rouges rôtis. C'est tendre et tout aussi savoureux quoique bien plus léger que l'imposant sandwich de poitrine de boeuf. Les deux sandwiches viennent avec de généreuses portions de viande et de sauce et il faut un effort monumental pour garder le tout ensemble dans ce grand pain. Le nombre de places est extrêmement limité et ça peut devenir assez encombré à l'intérieur, il est donc préférable d'essayer d'obtenir des sièges à l'extérieur ou de trouver un endroit ombragé sur l'un des bancs de parc à proximité.

B.5 Crowdsourced Translation – Non-Language Professionals – Task Repetition Value Four (CSTNLP4)

Au Black Camel, ils vendent des sandwichs sublimes depuis leur ouverture en 2004 et leur approche est aussi brillante que simple. À partir de cinq choix de sandwich de base, ils vous permettent de personnaliser votre repas en choisissant parmi une gamme de garnitures de qualité et de sauces faites maison. Les options incluent poitrine de bœuf lentement rôtie, épaule de porc, steak grillé, poulet rôti avec des tomates roma, poivron rouge, roquette et aubergine. Les prix commencent à 7,00 \$ (avec une sauce, uniquement). Il-y-à aussi plein d'autres sauces, condiments et garnitures disponibles pour moins d'un dollar. Ces sandwichs sont si populaires que Black Camel utilise Twitter pour informer les gens quand ils n'en ont plus, ce qui est souvent le cas avant leur heure de fermeture.

Aussi au menu se trouve un chili de Black Camel (3,25 \$ pour 6 oz, 5,25 \$ pour 12 oz), une variété de mélanges de café grillé et des sélections de thé provenant du House of Tea d'à coté. Pour le petit déjeuner, il-y-à un sandwich "eggspresso" servi jusqu'à 11h. Les œufs sont cuits un peu différemment que d'habitude (ils utilisent la vapeur de la machine expresso) et sont accompagnés soit de jambon pancetta, soit du steak ou un saumon mariné de style maison, aussi que du poivron et d'oignon rouge avec mayonnaise. Après une attente d'environ 5 minutes, nous passons nore commande. On commence par la poitrine de bœuf rôtie et accompagnée de la fameuse sauce barbecue Black Camel, d'une couche épaisse de mayonnaise chipotle et de raifort crémeux.

Cet énorme sandwich est rempli de tendre boeuf rôti (cuit lentement pendant 12 heures), bruni avec de la mélasse, des oignons caramélisés et parfumé de sauce barbeque et un petit piquant de raifort crémeux. C'est incroyablement copieux. Nous décidions d'essayer aussi le poulet déchiqueté rôti aux poivrons doux rouges. C'est très moelleux et savoureux mais plus léger que le gros sandwich de boeuf. Les deux sandwichs sont accompagnés de viande et sauce en quantité généreuse et c'est tout un défi à garder tous les ingrédients à l'intérieur de ce grand pain. Il n'y a pas beaucoup de places pour y manger et l'endroit peut devenir assez congestionné, alors c'est mieux de manger à l'extérieur ou de trouvez un petit endroit à l'ombre sur un des nombreux bancs de parc à proximité.

B.6 Machine Translation (MT)

Le chameau noir a été élingage sandwiches sublime depuis son ouverture en 2004 et leur approche est aussi intelligent que c'est simple. À partir de cinq choix de sandwich de base, ils vous permettent de personnaliser votre repas en choisissant parmi une gamme de haute qualité dans des sauces préparées maison et garnitures. Les options incluent lente poitrine de boeuf rôti, tiré d'épaule de porc, steak grillé, poulet rôti tiré et roma tomate, poivron rouge, roquette et aubergines. Les prix commencent à 7,00 \$ (avec une sauce). Il ya aussi beaucoup de sauces et de condiments et garnitures supplémentaires seront disponibles pour moins de un dollar. Alors populaires sont ces sandwiches que Black Camel utilise son Twitter que les gens sachent quand ils ont épuisé qui, la plupart des jours, est avant leur heure de fermeture prévue.

Également au menu est un Noir Camel Chili (3,25 \$ pour 6 onces, \$ 5.25 pour 12 oz.) Et une variété de multi-de torrédaction de café et sélection de thés à feuilles mobiles à proximité de la maison de thé. Pour le petit déjeuner, il ya un sandwich eggspresso servi jusqu'à 11 heures. Les oeufs sont cuits un peu différent que d'habitude (ils utilisent une buse vapeur machine à expresso) et jumelé avec soit du jambon pancetta, steak ou saumon mariné maison, le poivron et l'oignon rouge mayo). Après la queue pendant environ 5 minutes pour le dîner pour nous. La première est la poitrine de boeuf rôti lentement avec la sauce BBQ signature Black Camel et un slathering de mayonnaise au chipotle et au raifort crémeux.

Le sandwich est immense et ardue est chargé de poitrine de boeuf tendre super (cuisson lente pendant 12 heures) qui est dorée avec de la mélasse et l'oignon caramélisé et piquante à la saveur de la sauce barbecue et le zing légère crème de raifort. C'est incroyable et incroyablement remplissage. Nous avons également décider d'essayer le poulet rôti tira de poivron rouge grillé sucré. C'est humide et tout aussi savoureux, bien plus léger que le sandwich de poitrine de boeuf imposant. Les sandwiches sont livrés avec de généreuses portions de viande et la sauce et il faut un effort monumental pour que tout ensemble dans ce grand petit pain. Nombre de places est extrêmement limité et il peut se très fréquentée à l'intérieur il est donc préférable d'essayer d'obtenir en dehors des sièges ou de trouver un endroit ombragé sur l'un des bancs de parc à proximité.

B.7 CSTLP2 Rough Draft

Black Camel sert ses sublimes sandwiches depuis son ouverture en 2004 en se basant sur une approche aussi astucieuse que simple. À partir de cinq choix de sandwich de base, le client peut personnaliser son repas en ajoutant des sauces et des garnitures maison parmi l'éventail d'options offertes incluant : pointe de poitrine de bœuf cuite lentement, épaule de porc effilochée, steak grillé, poulet rôti effiloché avec tomates Roma, poivrons rouges, roquette et aubergine. Les prix commencent à 7 \$ (incluant une sauce). Plusieurs sauces, condiments et garnitures supplémentaires sont disponibles pour moins de 1 \$. Ces sandwiches sont si convoités que Black Camel utilise Twitter pour aviser ses clients lorsque l'offre est épuisée pour la journée ce qui, la plupart du temps, se produit avant l'heure de fermeture annoncée.

Également au menu, le chili Black Camel (3,25 \$ pour 6 oz, 5,25 \$ pour 12 oz) et un assortiment de mélanges de café torréfiés et une sélection de thés en feuille de la boutique House of Tea. Pour le déjeuner, un sandwich « eggspresso » est servi jusqu'à 11 h. Les œufs sont cuits un peu différemment (ils utilisent le bec à vapeur d'une machine à espresso) et sont servis avec un choix de pancetta, steak ou saumon fumé, poivrons et mayonnaise aux oignons rouges. Nous faisons la file environ 5 minutes avant de commander notre dîner. Nous commençons par la pointe de poitrine de bœuf rôtie lentement avec la sauce BBQ emblématique de Black Camel recouverte de mayonnaise chipotle et de raifort crémeux.

Le lourd et volumineux sandwich est chargé de pointe de poitrine de bœuf très tendre (cuite lentement pendant 12 heures) qui est dorée avec de la mélasse et des oignons caramélisés et relevée par la saveur de la sauce BBQ et du léger piquant de raifort crémeux. C'est incroyable et très substantiel. Nous décidons d'essayer également le poulet rôti effiloché avec poivrons rouges doux rôtis. Il est tendre et tout aussi délicieux quoique plus léger que l'imposant sandwich de pointe de poitrine de bœuf. Les deux sandwiches sont servis avec une généreuse portion de viande et de sauce et requièrent un important effort pour garder les ingrédients à l'intérieur du grand pain. Le nombre de places est très limité et le restaurant peut devenir assez bondé. Il est alors préférable de rechercher une place assise à l'extérieur ou trouver un coin ombragé sur un des bancs de parc tout près.

B.8 CSTLP2 Discarded Revision #1

Black Camel sert ses sublimes sandwiches depuis son ouverture en 2004 en se basant sur une approche aussi astucieuse que simple. À partir de cinq sandwiches de base, le client peut personnaliser son repas en ajoutant des sauces et des garnitures maison parmi l'éventail d'options offertes incluant : pointe de poitrine de bœuf cuite lentement, épaule de porc effilochée, steak grillé, poulet rôti effiloché avec tomates Roma, poivrons rouges, roquette et aubergine. Les prix commencent à 7 \$ (incluant une sauce). Plusieurs sauces, condiments et garnitures supplémentaires sont disponibles à moins de 1 \$. Ces sandwiches sont si convoités que Black Camel utilise Twitter pour aviser ses clients lorsque l'offre est épuisée pour la journée ce qui, la plupart du temps, se produit avant l'heure de fermeture annoncée.

Également au menu, le chili Black Camel (3,25 \$ pour 6 oz, 5,25 \$ pour 12 oz), une sélection de mélanges de café torréfiés et de thés en feuille de la boutique avoisinante House of Tea. Pour le déjeuner, un sandwich « eggspresso » est servi jusqu'à 11 h. Les œufs sont cuits un peu différemment (ils utilisent le bec à vapeur d'une machine à espresso) et sont servis avec un choix de pancetta, steak ou saumon fumé, poivrons et mayonnaise aux oignons rouges. Nous faisons la file environ 5 minutes avant de commander notre dîner. Nous commençons par la pointe de poitrine de bœuf rôtie lentement avec la sauce BBQ emblématique de Black Camel, recouverte de mayonnaise chipotle et de raifort crémeux.

Le lourd et volumineux sandwich est chargé de pointe de poitrine de bœuf très tendre (cuite lentement pendant 12 heures) qui est dorée avec de la mélasse et des oignons caramélisés et relevée par la saveur de la sauce BBQ et du léger piquant de raifort crémeux. C'est incroyable et très substantiel. Nous décidons d'essayer également le poulet rôti effiloché avec poivrons rouges rôtis. Il est tendre et tout aussi délicieux quoique plus léger que l'imposant sandwich de pointe de poitrine de bœuf. Les deux sandwiches sont servis avec une généreuse portion de viande et de sauce et demandent un important effort pour maintenir les ingrédients dans le grand pain. Le nombre de places est très limité et le restaurant peut devenir assez bondé. Il est alors préférable de rechercher une place assise à l'extérieur ou trouver un coin ombragé sur un des bancs de parc tout près.

B.9 CSTLP4 Rough Draft

Le Black Camel propose ses merveilleux sandwiches depuis son ouverture en 2004, et son approche est simple et intelligente. On vous propose cinq sandwiches de base au choix, que vous pouvez personnaliser en choisissant parmi une gamme de sauces et garnitures maison de haute qualité. Parmi ces options figurent de la poitrine de bœuf rôtie lentement, de l'épaule de porc effilée, du steak grillé, poulet rôtie effilé, et des tomates prune, des poivrons rouges, de la roquette et des aubergines. Les premiers prix débutent à 7,00 \$ (avec une sauce). Il y a aussi de nombreuses sauces, condiments et garnitures supplémentaires disponibles pour moins d'un dollar. Ces sandwiches sont si appréciés que le Black Camel se sert de Twitter pour dire aux gens qu'il ne reste plus de sandwiches, ce qui, la plupart du temps, arrive avant l'heure de fermeture prévue.

Au menu, vous trouverez le chili Black Camel (3,25 \$ pour 180 ml ou 5,25 \$ pour 360 ml) et une variété de café torréfié et de thé frais provenant de House of Tea qui est à proximité. Le déjeuner, un sandwich oeufpresso, est servi jusqu'à 11 h. Les oeufs sont cuits d'une manière unique (à la vapeur de la machine espresso) et sont servis avec du jambon pancetta, du steak ou du saumon mariné maison et une sauce à la mayonnaise, au poivron et à l'oignon rouge. Après avoir fait la file environ cinq minutes, nous avons commandé le dîner. D'abord, il y avait la pointe de bœuf cuit au pot-au-feu et servie avec la sauce BBQ signature Black Camel et beaucoup de mayonnaise à la chipotle et du raifort crémeux.

Le gros sandwich épais est garni de poitrine de bœuf tendre extra tendre (cuite lentement pendant 12 heures), dorée avec de la mélasse et des oignons caramélisés, relevé par une sauce barbecue piquante et une pointe de raifort crémeux. C'est incroyablement bon et ça vous remplit l'estomac. Nous avons également essayé le poulet rôti avec des poivrons rouges grillés. C'est moelleux et tout aussi savoureux que l'impressionnant sandwich de poitrine de bœuf, bien que plus léger. Les deux sandwiches sont servis avec de généreuses portions de viande et de sauce, et c'est loin d'être facile de faire tout tenir dans le pain. Le nombre de places est extrêmement limité, et le restaurant peut être plein à l'intérieur, il est donc préférable de s'asseoir dehors ou de trouver une petite place à l'ombre sur l'un des bancs du parc tout proche.

B.10 CSTLP4 Discarded Revision #1

Le Black Camel propose ses merveilleux sandwiches depuis son ouverture en 2004, et son approche est simple et intelligente. On vous propose cinq sandwiches de base au choix, que vous pouvez personnaliser en choisissant parmi une gamme de sauces et garnitures maison de haute qualité. Parmi ces options figurent de la poitrine de bœuf rôtie au pot-au-feu, de l'épaule de porc effilée, du steak grillé, du poulet rôtie effilé, et des tomates italiennes, des poivrons rouges, de la roquette et des aubergines. Le tout à partir de 7,00 \$ (avec une sauce). On peut aussi choisir beaucoup de sauces, condiments et garnitures supplémentaires à moins d'un dollar. Ces sandwiches sont si appréciés que le Black Camel se sert de Twitter pour dire aux gens qu'il ne reste plus de sandwiches, ce qui, la plupart du temps, arrive avant l'heure de fermeture prévue.

Au menu, vous trouverez le chili Black Camel (3,25 \$ pour 180 ml ou 5,25 \$ pour 360 ml) et une variété de café torréfié et de thé frais provenant de House of Tea qui est à proximité. Le déjeuner, un sandwich oeufpresso, est servi jusqu'à 11 h. Les oeufs sont cuits d'une manière unique (à la vapeur de la machine espresso) et sont servis avec du jambon pancetta, du steak ou du saumon mariné maison et une sauce à la mayonnaise, au poivron et à l'oignon rouge. Après avoir fait la file environ cinq minutes, nous avons commandé le dîner. D'abord, il y avait la pointe de bœuf cuit au pot-au-feu et servie avec la sauce BBQ signature Black Camel et beaucoup de mayonnaise à la chipotle et du raifort crémeux.

Le gros sandwich épais est garni de poitrine de bœuf extra tendre (cuite lentement pendant 12 heures), dorée avec de la mélasse et des oignons caramélisés, relevé par une sauce barbecue piquante et une pointe de raifort crémeux. C'est incroyablement bon et ça vous remplit l'estomac. Nous avons également essayé le poulet rôti avec des poivrons rouges grillés. C'est moelleux et tout aussi savoureux que l'impressionnant sandwich de poitrine de bœuf, bien que plus léger. Les deux sandwiches sont servis avec de généreuses portions de viande et de sauce, et c'est loin d'être facile de faire tout tenir dans le pain. Le nombre de places est extrêmement limité, et le restaurant peut être plein à l'intérieur, il est donc préférable de s'asseoir dehors ou de trouver une petite place à l'ombre sur l'un des bancs du parc tout proche.

B.11 CSTLP4 Discarded Revision #2

Black Camel propose ses merveilleux sandwiches depuis son ouverture en 2004, et son approche est aussi simple qu'intelligente. On vous propose cinq sandwiches de base au choix que vous pouvez personnaliser en choisissant parmi une gamme de sauces et garnitures maison de haute qualité. Parmi ces options figurent à la poitrine de bœuf rôtie à feu doux, à l'épaule de porc effilée, au steak grillé, au poulet rôti effilé et aux tomates Roma, poivrons rouges, roquette et aubergines. Les prix débutent à 7,00 \$ (et donnent droit à une sauce). Il y a aussi de nombreux condiments, sauces et garnitures supplémentaires disponibles pour moins d'un dollar. Ces sandwiches sont si appréciés que le Black Camel se sert de Twitter pour informer sa clientèle qu'il ne reste plus de sandwiches, ce qui, la plupart du temps, se produit avant l'heure de fermeture prévue.

Également au menu, vous trouverez le chili Black Camel (3,25 \$ pour 180 ml ou 5,25 \$ pour 360 ml) et une variété de café torréfié et de thé frais provenant du salon de thé House of Tea situé à proximité. Pour le déjeuner, un sandwich express-oeufs est servi jusqu'à 11 h. Les oeufs sont cuits d'une manière unique (à l'aide de la buse à vapeur de la machine espresso) et sont servis avec du jambon pancetta, du steak ou du saumon mariné maison et une sauce à la mayonnaise, au poivron et à l'oignon rouge. Après avoir fait la file environ cinq minutes, nous avons commandé notre dîner. D'abord, il y avait la poitrine de bœuf cuit à feu doux et servie avec la sauce BBQ signée Black Camel et une généreuse portion de mayonnaise au chipotle et de raifort crémeux.

Ce sandwich gargantuesque est garni de poitrine de bœuf extra tendre (cuite 12 heures à feu doux), dorée avec de la mélasse et des oignons caramélisés et relevée par une sauce barbecue et une touche épicée de raifort crémeux. Mets absolument délicieux et satisfaisant. Nous avons également essayé le poulet rôti effilé avec des poivrons rouges grillés. Moelleux et tout aussi savoureux que l'impressionnant sandwich de poitrine de bœuf, il est toutefois plus léger. Les deux sandwiches sont servis avec de généreuses portions de viande et de sauce; par conséquent, il est loin d'être facile de tout faire tenir entre les énormes tranches de pain. Le nombre de places est extrêmement limité, et le restaurant risque d'être plein à l'intérieur, il est donc préférable de se trouver une place à l'extérieur ou encore à l'ombre sur l'un des bancs du parc tout proche.

B.12 CSTLP4 Discarded Revision #3

Le Black Camel propose ses merveilleux sandwiches depuis son ouverture en 2004, et son approche est simple et intelligente. On vous propose cinq sandwiches de base au choix, que vous pouvez personnaliser en choisissant parmi une gamme de sauces et garnitures maison de haute qualité. Parmi ces options figurent de la poitrine de bœuf rôtie lentement, de l'épaule de porc effilée, du steak grillé, du poulet rôtie effilé, et des tomates prune, des poivrons rouges, de la roquette et des aubergines. Les prix débutent à 7,00 \$ (avec une sauce). Il y a aussi de plusieurs sauces, condiments et garnitures supplémentaires disponibles pour moins d'un dollar. Ces sandwichs sont si appréciés que le Black Camel se sert de Twitter pour annoncer à ces clients quand il n'est reste plus de sandwiches, ce qui, la plupart du temps, arrive avant l'heure de fermeture prévue.

Au menu, vous trouverez le chili Black Camel (3,25 \$ pour 180 ml ou 5,25 \$ pour 360 ml) et une variété de café torréfié et de thé frais provenant de House of Tea qui est à proximité. Le déjeuner, un sandwich oeufpresso, est servi jusqu'à 11 h. Les oeufs sont cuits d'une manière unique (à la vapeur de la machine espresso) et sont servis avec du jambon pancetta, du steak ou du saumon mariné maison et une sauce à la mayonnaise, au poivron et à l'oignon rouge. Après avoir fait la file environ cinq minutes, nous avons commandé le dîner. D'abord, il y avait la pointe de bœuf cuit au pot-au-feu et servie avec la sauce BBQ signature Black Camel et beaucoup de mayonnaise à la chipotle et du raifort crémeux.

Le gros sandwich épais est garni de poitrine de bœuf tendre extra tendre (cuite lentement pendant 12 heures), dorée avec de la mélasse et des oignons caramélisés, relevé par une sauce barbecue piquante et une pointe de raifort crémeux. C'est incroyablement bon et ça vous remplit l'estomac. Nous avons également essayé le poulet rôti avec des poivrons rouges grillés. C'est moelleux et tout aussi savoureux que l'impressionnant sandwich de poitrine de bœuf, bien que plus léger. Les deux sandwichs sont servis avec de généreuses portions de viande et de sauce, et c'est loin d'être facile de faire tout tenir dans le pain. Le nombre de places est extrêmement limité, et le restaurant peut être plein à l'intérieur, il est donc préférable de s'asseoir dehors ou de trouver une petite place à l'ombre sur l'un des bancs du parc tout proche.

B.13 CSTNLP2 Rough Draft

Le restaurant "Black Camel" sert des sandwiches exquis depuis son ouverture en 2004. Leur approche est à la fois simple et intelligente. Partant d'un choix de cinq sandwiches de base, vous pouvez personnaliser votre repas avec des sauces et des condiments de haute qualité préparés sur place. Les options comprennent de la poitrine de boeuf à la mijoteuse, viande de porc effilée, steak ou poulet grillé ainsi que des tomates roma, poivrons rouges, roquette et aubergine. Les prix commencent à sept dollars, ce qui comprend une sauce. De plus, plusieurs sauces et condiments sont disponibles pour moins d'un dollar. Leurs sandwiches sont d'une telle popularité que le "Black Camel" utilise Twitter pour faire savoir quand une sorte de sandwich n'est plus disponible. Et ce, fréquemment avant l'heure de fermeture.

Également au menu est un chili Black Camel (3,25 \$ pour 6 onces, \$ 5,25 pour 12 onces) et une variété de torréfaction de café et une sélection de thés à feuilles de la House of Tea à proximité. Pour le petit déjeuner, il y a un sandwich "Eggspresso" servi jusqu'à 11 heures. Les œufs sont cuits un peu différent qu'à l'habitude (ils utilisent une buse à vapeur de machine à espresso) et combinés à un choix de jambon pancetta, de steak ou de saumon fumé maison, avec une mayo aux poivrons et oignons rouges). Après avoir patienté en ligne pendant environ 5 minutes, nous commandons le repas. En premier vient la poitrine de boeuf rôti lentement avec la sauce signature Black Camel BBQ et une couche de mayonnaise au chipotle et au raifort crémeux.

L'immense et ardue sandwich est chargé de poitrine de boeuf super tendre (à la cuisson lente pendant 12 heures) qui est dorée avec de la mélasse et oignons caramélisés et piquante à la saveur de la sauce barbecue et le léger goût du raifort crémeux. C'est stupéfiant et incroyablement satiant. Nous décidons aussi d'essayer le poulet rôti tiré avec poivrons rouges rôtis. C'est tendre et tout aussi savoureux quoique bien plus léger que l'imposant sandwich de poitrine de boeuf. Les deux sandwiches viennent avec de généreuses portions de viande et de sauce et il faut un effort monumental pour garder le tout ensemble dans ce grand pain. Le nombre de places est extrêmement limité et ça peut devenir assez encombré à l'intérieur, il est donc préférable d'essayer d'obtenir des sièges à l'extérieur ou de trouver un endroit ombragé sur l'un des bancs de parc à proximité.

B.14 CSTNLP2 Discarded Revision #1

Le restaurant "Black Camel" sert des sandwichs exquis depuis son ouverture en 2004, et leur approche est à la fois simple et intelligente. Partant d'un choix de cinq sandwichs de base, vous pouvez personnaliser votre repas avec des sauces et des condiments de haute qualité préparés sur place. Les options comprennent de la poitrine de boeuf à la mijoteuse, viande de porc effilée, steak ou poulet grillé ainsi que des tomates roma, poivrons rouges, roquette et aubergine. Les prix commencent à sept dollars (avec une sauce). De plus, plusieurs sauces et condiments sont disponibles pour moins d'un dollar. Leurs sandwichs sont d'une telle popularité que le "Black Camel" utilise Twitter pour faire savoir quand une sorte de sandwich n'est plus disponible. Et ce, fréquemment avant l'heure de fermeture.

Également au menu est un chili Black Camel (3,25 \$ pour 6 onces, \$ 5,25 pour 12 onces) et une variété de torréfaction de café et une sélection de thés à feuilles de la House of Tea à proximité. Pour le petit déjeuner, il y a un sandwich "Eggspresso" servi jusqu'à 11 heures. Les œufs sont cuits un peu différent qu'à l'habitude (ils utilisent une buse à vapeur de machine à espresso) et combinés à un choix de jambon pancetta, de steak ou de saumon fumé maison, avec une mayo aux poivrons et oignons rouges). Après avoir patienté en ligne pendant environ 5 minutes, nous commandons le repas. En premier vient la poitrine de boeuf rôti lentement avec la sauce signature Black Camel BBQ et une couche de mayonnaise au chipotle et au raifort crémeux.

L'immense et ardue sandwich est chargé de poitrine de boeuf super tendre (cuit à la mijoteuse pendant 12 heures) qui est dorée avec de la mélasse et oignons caramélisés et piquante à la saveur de la sauce barbecue et le léger goût du raifort crémeux. C'est stupéfiant et incroyablement satiant. Nous décidons aussi d'essayer le poulet rôti tiré avec poivrons rouges rôtis. C'est tendre et tout aussi savoureux quoique bien plus léger que l'imposant sandwich de poitrine de boeuf. Les deux sandwichs viennent avec de généreuses portions de viande et de sauce et il faut un effort monumental pour garder le tout ensemble dans ce grand pain. Le nombre de places est extrêmement limité et ça peut devenir assez encombré à l'intérieur, il est donc préférable d'essayer d'obtenir des sièges à l'extérieur ou de trouver un endroit ombragé sur l'un des bancs de parc à proximité.

B.15 CSTNLP4 Rough Draft

Au Black Camel, ils vendent des sandwichs sublimes comme des petits pains depuis leur ouverture en 2004 et leur approche est aussi brillante que simple. À partir de cinq choix de sandwich de base, ils vous permettent de personnaliser votre repas en choisissant parmi une gamme de garnitures de qualité et de sauces faites maison. Les options incluent poitrine de bœuf lentement rôtie, épaule de porc, steak grillé, poulet rôti avec des tomates roma, poivron rouge, roquette et aubergine. Les prix commencent à 7,00 \$ (avec une sauce). Il y a aussi plein d'autres sauces, condiments et garnitures disponibles pour moins d'un dollar. Ces sandwichs sont si populaires que Black Camel utilise Twitter pour informer les gens quand ils n'en ont plus, ce qui est souvent le cas avant leur heure de fermeture.

Sont également au menu un chili de Black Camel (3,25 \$ pour 6 oz, 5,25 \$ pour 12 oz), une variété de mélanges de café grillé et des sélections de thé provenant de la maison d'à côté. Pour le petit déjeuner, il y a un sandwich eggspresso servi jusqu'à 11h. Les œufs sont cuits un peu différemment que d'habitude (ils utilisent la vapeur de la machine espresso) et sont accompagnés soit de jambon pancetta, soit de saumon en steak ou mariné maison, soit de poivron et oignon rouge avec mayonnaise. Après une attente d'environ 5 minutes, nous passons commande du diner. On commence par la poitrine de bœuf rôtie et accompagnée de la fameuse sauce barbecue Black Camel, d'une couche épaisse de mayonnaise chipotle et de raifort crémeux.

Cet énorme sandwich est rempli de tendre boeuf rôti (cuit lentement en 12 heures), bruni avec de la mélasse, des oignons caramélisés et parfumé de sauce BBQ et un petit piquant de raifort crémeux. C'est incroyablement copieux. Nous décidions d'essayer aussi le poulet déchiqueté rôti avec poivrons doux rouges. C'est très moelleux et savoureux malgré que plus léger que le gros sandwich de boeuf. Les deux sandwichs sont accompagnés de viande et sauce en quantité généreuse et c'est tout un défi à garder tous les ingrédients à l'intérieur de ce grand pain. Il n'y a pas beaucoup de places pour y manger et l'endroit peut devenir assez congestionné, alors c'est mieux de manger à l'extérieur ou de trouver un petit endroit à l'ombre sur un des nombreux bancs de parc à proximité.

B.16 CSTNLP4 Discarded Revision #1

Au Black Camel, ils vendent des sandwichs sublimes depuis leur ouverture en 2004 et leur approche est aussi brillante que simple. À partir de cinq choix de sandwichs de base, vous pouvez personnaliser votre repas en choisissant parmi une gamme de garnitures et de sauces de qualité, faits maison. Parmi les options on inclut la bavette de bœuf lentement rôti, l'épaule de porc, le steak grillé, le poulet rôti, ainsi que des tomates roma, des poivrons rouges, de la roquette et des aubergines. Les prix commencent à 7,00 \$ pour une sauce. Il y a aussi plusieurs autres sauces, condiments et garnitures disponibles pour un coût de moins d'un dollar. Ces sandwichs sont si populaires que le Black Camel utilise Twitter pour informer les gens quand ils n'en ont plus, ce qui arrive souvent bien avant l'heure de fermeture.

Également au menu, est un chili Black Camel (3,25 \$ pour 6 onces ou 5,25 \$ pour 12 onces), et une variété de mélanges de café et des sélections de thé en feuilles provenant de la maison de thé voisine. Pour le petit déjeuner, servi jusqu'à 11h, il y a le sandwich œufspresso. Les œufs sont cuits un peu différemment qu'à l'habitude (ils utilisent la vapeur d'une machine espresso) et sont accompagnés soit de jambon pancetta, de steak, ou de saumon fumé maison, de poivron doux et de mayonnaise aux oignons rouges. Après une attente en ligne d'environ 5 minutes, nous passons notre commande pour dîner. En premier, c'est la bavette de bœuf rôtie accompagnée de la sauce barbecue Black Camel servie avec une portion épaisse de mayonnaise chipotle et de raifort crémeux.

Cet énorme sandwich est rempli de tendre boeuf rôti (cuit lentement sur 12 heures), brun avec de la mélasse, des oignons caramélisés et parfumé de sauce barbecue ainsi qu'un petit piquant de raifort crémeux. C'est incroyablement copieux. Nous décidons d'essayer aussi le poulet déchiqueté rôti aux poivrons doux rouges. C'est très moelleux et aussi savoureux, malgré que plus léger que le gros sandwich au boeuf. Les deux sandwichs sont accompagnés de viande et sauces en quantité généreuse et c'est tout un défi de retenir tous les ingrédients à l'intérieur de ce grand pain. Il y a peu de places pour manger sur place et l'endroit peut devenir assez congestionné, alors c'est préférable de manger à l'extérieur ou de trouver un endroit à l'ombre sur un des nombreux bancs de parc à proximité.

B.17 CSTNLP4 Discarded Revision #2

Au Black Camel, ils vendent des sandwichs sublimes depuis leur ouverture en 2004 et leur approche est aussi brillante que simple. Avec un choix de cinq sandwichs de base, vous pouvez personnaliser votre repas en choisissant parmi une gamme de garnitures de qualité et de sauces faites maison. Les options incluent: une poitrine de bœuf lentement rôtie, une épaule de porc, steak grillé, un poulet rôti avec des tomates roma, poivrons rouges, roquette et aubergine. Les prix commencent à \$7,00 (y compris une sauce). Il y a aussi plusieurs autres sauces, condiments et garnitures disponibles pour moins d'un dollar. Ces sandwichs sont si populaires que Black Camel utilise Twitter pour informer les gens quand ils n'en ont plus. Ce qui est souvent le cas avant leur heure de fermeture.

Egalement au menu, il y a un chili (\$3,25 pour 6 on, \$5,25 pour 12 on), une variété de mélanges de cafés grillés et des sélections de thés provenant de la maison House of Tea. Pour le petit déjeuner, il y a un sandwich eggspresso servi jusqu'à 11h00. Les œufs sont cuits un peu différemment que d'habitude (ils utilisent la vapeur de la machine espresso) et sont accompagnés soit de jambon pancetta, de saumon en steak ou mariné maison, soit de poivron et oignon rouge avec mayonnaise. Après une attente d'environ 5 minutes, nous passons à la commande du diner. On commence par la poitrine de bœuf rôtie accompagnée de la fameuse sauce BBQ Black Camel, d'une couche épaisse de mayonnaise chipotle et de raifort crémeux.

Cet énorme sandwich est rempli de tendre boeuf rôti (cuit lentement pendant 12 heures), brun avec de la melasse, des oignons caramélisés et parfumé de sauce BBQ avec un petit piquant de raifort crémeux. C'est incroyablement copieux. Nous décidons d'essayer aussi le poulet déchiqueté rôti aux poivrons doux rouges. Il est très moelleux et savoureux et plus léger que le gros sandwich de boeuf. Les deux sandwichs sont accompagnés de viande et sauce en quantité généreuse et c'est tout un défi à garder tous ces ingrédients à l'intérieur de ce grand pain. Il n'y a pas beaucoup de places pour manger et l'endroit peut devenir assez congestionné. Alors, c'est mieux de manger à l'extérieur ou de trouver un petit endroit à l'ombre sur un des nombreux bancs de parc tout proches.

B.18 CSTNLP4 Discarded Revision #3

Au Black Camel, on vent des sandwichs sublimes depuis leur ouverture en 2004, et leur approche est aussi intelligente que simple. À partir de cinq choix de sandwich de base, ils vous permettent de personnaliser votre repas en choisissant parmi une gamme de garnitures et sauces maison de qualité. Les options incluent le bœuf rôti à petit feu, épaule de porc, steak grillé à petit feu, poulet rôti déchiqueté ainsi que des tomates roma, poivrons rouges, roquettes et aubergines. Les prix commencent à 7,00 \$, pour un seul choix de sauce. Une variété d'autres sauces, condiments et garnitures est disponible pour moins d'un dollar. Ces sandwichs sont si populaires que Black Camel utilise Twitter pour informer les gens quand ils n'en ont plus, ce qui est souvent le cas avant leur heure de fermeture.

Également au menu, on trouve le chili Black Camel (3,25 \$ pour 6 oz, 5,25 \$ pour 12 oz), une variété de mélanges de café grillé et un choix de thés provenant de la House of Tea, un établissement voisin. Pour le petit déjeuner, il y a un sandwich eggspresso qui est servi jusqu'à 11h. Les œufs sont cuits de façon originale (ils utilisent la vapeur de la machine expresso) et sont servis avec un choix de jambon pancetta, steak, ou saumon mariné dans une mayonnaise de piment et oignons. . Après environ 5 minutes d'attente, nous passons notre commande. Nous commençons par le bœuf rôti à petit feu et accompagnée de la fameuse sauce barbecue Black Camel, ainsi que d'une couche épaisse de mayonnaise chipotle et de raifort crémeux.

Cet énorme sandwich est rempli de tendre boeuf rôti (cuit à petit feu pendant 12 heures), bruni avec de la mélasse, des oignons caramélisés et parfumé de sauce BBQ avec un petit piquant de raifort crémeux. Il est incroyablement copieux. Nous avons aussi essayé le poulet déchiqueté rôti aux poivrons rouges rôtis. Il est très moelleux et savoureux, quoique plus léger que l'énorme sandwich de boeuf. Les deux sandwichs sont accompagnés de viande et sauce en quantité généreuse et c'est tout un défi de garder tous les ingrédients à l'intérieur du pain, malgré sa grande taille. Il n'y a pas beaucoup de place pour y manger et l'endroit peut devenir assez congestionné, alors il est préférable de manger à l'extérieur ou de trouver un petit endroit à l'ombre sur un des nombreux bancs de parc à proximité.

B.19 Discarded Professional Translation #1

Black Camel confectionne de sublimes sandwiches depuis son ouverture en 2004 et son approche est aussi intelligente que simple. Cinq choix de base sont au départ offerts pour les sandwiches, vous permettant ainsi de personnaliser votre repas en choisissant parmi une gamme de sauces et de garnitures maison de qualité supérieure. Les choix comprennent la pointe de poitrine de boeuf cuite lentement sur le grill, l'épaule de porc effilochée, le bifteck saisi sur le feu, le poulet rôti effiloché servis avec des tomates prunes, des poivrons rouges, de la roquette et de l'aubergine. Les prix à partir de 7 \$ (avec une seule sauce) sont affichés. De nombreux condiments, sauces et garnitures supplémentaires sont disponibles à moins d'un dollar. Ces sandwiches sont tellement prisés que Black Camel utilise son Twitter pour faire savoir à tout le monde que son stock est épuisé, ce qui se produit la plupart du temps avant l'heure de fermeture prévue.

Le chili Black Camel (3,25 \$ pour 6 oz, 5,25 \$ pour 12 oz), une gamme de mélanges de cafés de différents niveaux de torréfaction ainsi qu'un choix de thés en vrac provenant de la boutique House of Tea voisine figurent également au menu. Pour le petit déjeuner, un sandwich oeufpresso est servi jusqu'à 11 h. Les oeufs sont préparés d'une manière un peu différente de celle dont on se sert habituellement (ils utilisent le bec à vapeur d'une machine à café expresso) et jumelés à soit du jambon pancetta, du bifteck ou du saumon salé, séché et fumé maison, du poivron et de la mayonnaise à l'oignon rouge. Après une attente en file d'environ 5 minutes, nous passons la commande pour le repas du midi. C'est d'abord la pointe de poitrine de boeuf grillée lentement, accompagnée de la sauce BBQ exclusive de Black Camel et de mayonnaise au chipotle et de raifort crémeux en abondance.

L'énorme et lourd sandwich est rempli d'une pointe de poitrine de boeuf très tendre (cuite lentement sur le grill pendant 12 heures) brunie dans la mélasse, accompagnée d'oignons caramélisés, de la saveur relevée de la sauce BBQ et d'une touche modérément épicée de raifort crémeux. Le sandwich nous surprend et rassasie de manière incroyable. Nous avons également décidé d'essayer le poulet rôti effiloché accompagné de poivron rouge grillé. Il est moelleux et a également bon goût bien qu'il soit plus léger que l'imposant sandwich de pointe de poitrine de boeuf. Les deux sandwiches sont proposés avec de généreuses portions de viande et de sauce et il faut un effort extraordinaire pour ne rien échapper de ce gros pain à hambourgeois. Le nombre de places assises est extrêmement limité et l'endroit peut devenir plein à craquer; il est donc préférable d'essayer de trouver une place assise à l'extérieur ou à l'ombre sur un des bancs de l'un des parcs avoisinants.

B.20 Discarded Professional Translation #2

Le café Black Camel ne cesse de lancer des sandwiches divins depuis son ouverture, en 2004. Sa démarche est aussi intelligente que simple. À partir de cinq options de sandwich de base, il vous donne l'occasion de personnaliser votre repas en faisant votre choix parmi un éventail de garnitures et de sauces maison de qualité supérieure. Ces options englobent la pointe de poitrine de boeuf grillée lentement, l'épaule de porc effilochée, le steak grillé, le poulet rôti effiloché, ainsi qu'un mélange de tomates italiennes, de poivrons rouges, de roquettes et d'aubergines. Le prix de départ s'établit à 7,00 \$ (sandwich accompagné d'une sauce). Une abondance de sauces, de condiments et de garnitures supplémentaires est également proposée pour moins d'un dollar. Ces sandwiches sont si populaires que le Black Camel utilise sa page Twitter pour informer la clientèle de l'écoulement complet de ses différentes variétés de sandwiches, événement qui se produit, la plupart du temps, avant l'heure de fermeture prévue.

Le chili Black Camel (portions de 6 onces à 3,25 \$ et de 12 onces à 5,25 \$) figure également au menu, en plus d'une sélection de mélanges de café multi-torréfié et de thés en feuilles provenant de la boutique House of Tea, située à proximité immédiate. Pour déjeuner, on sert un sandwich eggspresso jusqu'à 11 h. Les oeufs, cuits selon une méthode qui diffère quelque peu de l'ordinaire (grâce au bec vapeur de la cafetière à expresso), sont jumelés soit avec de la pancetta, soit avec du steak ou du saumon préparé sur place, soit avec des poivrons et de la mayonnaise aux oignons rouges. Après avoir fait la queue pendant près de cinq minutes, nous commandons à dîner. D'abord, nous choisissons la pointe de poitrine de boeuf grillé lentement, agrémentée de la sauce barbecue exclusive à Black Camel, puis une épaisse couche de mayonnaise au chipotle et de raifort crémeux.

Ce sandwich à la fois énorme et pesant regorge de morceaux de pointe de poitrine de boeuf super-tendre (et grillé lentement pendant 12 heures), qu'on a fait brunir dans la mélasse et les oignons caramélisés en le rehaussant de la saveur piquante de la sauce barbecue et d'un doux soupçon vivifiant de raifort. D'un goût renversant, il est incroyablement nourrissant. Nous décidons également d'essayer le poulet rôti effiloché, en l'agrémentant de poivrons rouges grillés. Cette combinaison juteuse est aussi savoureuse que la première, mais plus légère qu'un sandwich imposant à la pointe de poitrine de boeuf. Les deux sandwiches sont accompagnés de généreuses portions de viande et de sauce. Il faut donc déployer un effort monumental pour faire tenir tous les ingrédients dans leur gros pain rond. Par ailleurs, le nombre de places est extrêmement limité à l'intérieur du café, qui peut être très achalandé par périodes. Il vaut mieux essayer de se trouver une place à l'extérieur ou un endroit ombragé en s'asseyant sur l'un des bancs du parc adjacent.

B.21 Discarded Machine Translation #1

Le Camel noir a été élingage sandwichs sublimes puisqu'il a ouvert en 2004 et leur approche est aussi intelligent que c'est simple. Commençant par cinq choix de base sandwich, ils permettent de personnaliser votre repas en choisissant parmi une gamme de sauces préparées en maison de haute qualité et de glaçage. Les options comprennent la poitrine de boeuf rôti lente, extraites d'épaule de porc, steak poêlé, poulet rôti d'extraites et tomate roma, poivrons rouges, roquette et aubergine. Prix de départ \$ 7.00 (avec une sauce). Il existe également beaucoup de sauces et condiments et garniture supplémentaire pour moins d'un euro. Si populaires sont ces sandwichs que Camel noir utilise leur Twitter pour informer les gens quand ils ont épuisé, sur la plupart des jours, c'est-à-dire avant leur heure de fermeture.

Également dans le menu est un chili Camel noir (3.25 \$ pour 6 oz, 5,25 \$ pour 12 onces) et une variété de multi-roast café et thé feuilletés mobiles sélections de la maison de thé à proximité de la fusion. Pour le petit déjeuner, il y a un sandwich eggspresso servi jusqu'à 11 h. Les oeufs sont cuites un peu différent de l'habituel (ils utilisent une baguette de vapeur de machine espresso) et appairée avec soit pancetta jambon, steak ou saumon maison jaune, poivrons doux et mayo oignon rouge). Après la doublure pour environ 5 minutes, nous ordonnons déjeuner. La première est la poitrine de boeuf rôti lente avec la sauce barbecue de Camel noir de signature et une slathering de mayo de piments et crémeux raifort.

Le sandwich énorme et lourd est chargé avec la poitrine de boeuf super offres (lent-cuit pendant 12 heures) Bruni avec MELASSE et oignon caramélisé et âcre avec la saveur de la sauce barbecue et le doux zing de raifort crémeux. Il est incroyable et incroyablement remplissage. Nous également décider d'essayer le rôti extraites de poulet avec rôti poivrons rouges doux. Il est humide et tout aussi savoureux, bien que plus léger que l'imposant sandwich de poitrine de viande bovine. Les deux sandwichs accompagnent helpings généreuses de viandes et sauce et il prend un effort monumental pour conserver tout ensemble dans cette grande bun. Assise est extrêmement limitée et il peut obtient assez entassés à l'intérieur ; il est donc préférable d'essayer d'obtenir des places à l'extérieur ou trouver une place ombragée sur n'importe quel les bancs de parc à proximité.

B.22 Discarded Machine Translation #2

Le chameau noir avait lancé les sandwiches sublimes depuis qu'il s'est ouvert en 2004 et leur approche est aussi futée qu'elle est simple. Commenant par cinq choix de base de sandwich, ils te permettent de personnaliser votre repas par le choix d'une gamme de haute qualité en sauces et écriimages préparés par maison. Les options incluent la poitrine de boeuf rôtie lente, l'épaule de porc tirée, le bifteck desséché, le poulet et la tomate tirée rôtie de Roma, le poivron rouge, l'arugula et l'aubergine. Les prix commencent à \$7.00 (avec de l'une sauce). Il y a également abondance des sauces et des condiments et écriimages supplémentaires disponibles pour moins qu'un mâle. Si populaires sont ces sandwiches que le chameau noir emploie leur Twitter pour faire des personnes connaître quand they' ; le VE s'est vendu dehors qui, la plupart des jours, est avant leur heure exacte programmée.

Également sur le menu est un piment noir de chameau (\$3.25 pour 6 onces., \$5.25 pour 12 onces.) et une série de multi-rôtissez les choix de thé de café et de feuille de mélange à partir de la Chambre voisine du thé. Pour le déjeuner, there' ; s qu'un sandwich à eggspresso a servi jusqu'à ce que 11 AM. Les oeufs sont un peu différents cuit que d'habitude (ils utilisent une baguette magique de vapeur de machine de café express) et appareillé avec du l'un ou l'autre jambon de pancetta, bifteck ou saumons maison-traités, poivron doux et oignon rouge mayo). Après l'alignement pendant environ 5 minutes nous commandons le déjeuner. Est d'abord la poitrine de boeuf rôtie lente avec de la sauce à BBQ de chameau de noir de signature et slathering du chipotle mayo et du raifort crémeux.

Le sandwich énorme et lourd est chargé avec la poitrine de boeuf tendre superbe (lent-cuite pendant 12 heures) qui est brunie avec de la mélasse et l'oignon caramélisé et piquant avec la saveur de la sauce à BBQ et le zing doux du raifort crémeux. It' ; s stupéfiant et remplissant incroyablement. Nous décidons également d'essayer le poulet tiré rôti avec le poivron rouge doux rôti. It' ; s moite et également savoureux bien qu'allumeur que le sandwich imposant à poitrine de boeuf. Les deux sandwiches viennent avec des helpings généreux de viande et la sauce et lui prend un effort monumental de maintenir tout ensemble dans ce grand pain. L'allocation des places est extrêmement limitée et elle peut obtient l'intérieur ainsi assez serré it' ; s meilleur pour essayer et obtenir l'allocation des places dehors ou pour trouver une tache ombreuse sur les bancs de parc voisins l'uns des.

Appendix C – Phase II (Translation Evaluation) Data

This appendix contains PLS path models relating the concepts *Selfish Selection*, *Unselfish Selection*, and *Disinterested Selection* to *Overall Translation Quality*. These hypotheses were not tested as part of the main body of the thesis because these concepts were not identified prior to the beginning of the study. These PLS path models have been included for the sake of completeness and to inform future research. This appendix also includes most of the detailed confidence interval tables that were used to validate **H₂**, **H₄**, and **H₅**. Finally, this Appendix concludes with the instrumentation provided to participants in Phase II (Translation Evaluation).

C.1 Selfish Selection and Overall Translation Quality

Selfish Selection has no effect on the Overall Translation Quality of crowdsourced translation.

This hypothesis is tested with all 128 samples.

Table 50 – Selfish Selection Hypothesis (Structural and Measurement Model)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Selfish Selection	Reflective	SPES	Selfish Paragraph Evaluation Selection (cf. Table 6)
		SRES	Selfish Revision Evaluation Selection (cf. Table 6)
		STES	Selfish Total Evaluation Selection (cf. Table 6)
Overall Translation Quality	Reflective	AQ1	The quality of this text is acceptable for a Canadian Francophone website.
		OTQ11	Most Canadian Francophones would not be able to tell that this text is a translation.
		OTQ21	Most Canadian Francophones would feel confident trusting the content of this translation.
		OTQ31	Most Canadian Francophone readers of this text would feel confident contacting the company in French.
		OTQ41	Most Canadian Francophone readers of this text would feel confident purchasing products of services from this company.

Table 51 – Selfish Selection Hypothesis (Bootstrapping Results)

BOOTSTRAPPING (128 samples with 500 iterations/degrees of freedom)		T-STATISTIC VALUES
Structural Model	Selfish Evaluation Selection → Overall Translation Quality	0.685705 (not statistically significant)
Measurement Model	SPES	1.772176
	SRES	0.559328
	STES	1.932508
	AQ1	4.458593
	OTQ11	4.358256
	OTQ21	4.330068
	OTQ31	4.402717
	OTQ41	4.392277

C.2 Unselfish Selection and Overall Translation Quality

Unselfish Selection has no effect on the Overall Translation Quality of crowdsourced translation.

This hypothesis is tested with all 128 samples.

Table 52 – Unselfish Selection Hypothesis (Structural and Measurement Model)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Unselfish Selection	Reflective	UPES	Unselfish Paragraph Evaluation (cf. Table 6)
		URES	Unselfish Revision Evaluation (cf. Table 6)
		UTES	Unselfish Total Evaluation Selection (cf. Table 6)
Overall Translation Quality	Reflective	AQ1	The quality of this text is acceptable for a Canadian Francophone website.
		OTQ11	Most Canadian Francophones would not be able to tell that this text is a translation.
		OTQ21	Most Canadian Francophones would feel confident trusting the content of this translation.
		OTQ31	Most Canadian Francophone readers of this text would feel confident contacting the company in French.
		OTQ41	Most Canadian Francophone readers of this text would feel confident purchasing products of services from this company.

Table 53 – Unselfish Selection Hypothesis (Bootstrapping Results)

BOOTSTRAPPING (128 samples with 500 iterations/degrees of freedom)		T-STATISTIC VALUES
Structural Model	Unselfish Evaluation Selection → Overall Translation Quality	4.048869 (significant @ 0.0001)
Measurement Model	UPES	18.499549
	URES	1.294507
	UTES	32.259116
	AQ1	62.578185
	OTQ11	54.549948
	OTQ21	79.391200
	OTQ31	63.569421
	OTQ41	49.702780

Figure 45 – Unselfish Selection Hypothesis (PLS Algorithm Results)

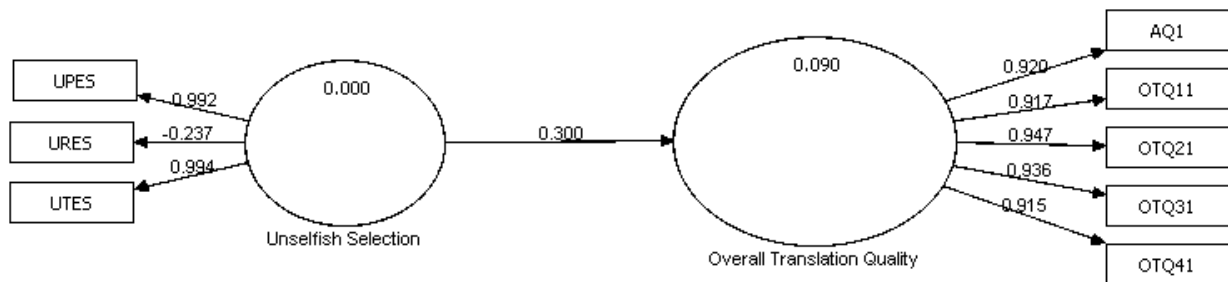


Table 54 – Unselfish Selection Hypothesis (PLS Algorithm Results)

TRANSLATION EVALUATORS – Non-Language Professionals	
PLS ALGORITHM	VALUES
Path Coefficient (Unselfish Evaluation Selection) → Overall Translation Quality)	0.300057
Cronbach's Alpha	0.959184
Composite Reliability	0.968286
Average Variance Extracted (AVE) for Overall Translation Quality	0.859304
R ²	0.090034

Here, the results suggest that however weakly, when during the evaluation tasks of Phase I, crowdsourced participants actually selected other people's work over their own, that this contributed in a small way towards an increase in *Overall Translation Quality*.

C.3 Disinterested Selection and Overall Translation Quality

This hypothesis is tested with all 128 samples.

Table 55 – Disinterested Selection Hypothesis (Structural and Measurement Model)

LATENT VARIABLE	TYPE	INDICATOR	DESCRIPTION
Disinterested Selection	Reflective	DPES	Disinterested Paragraph Evaluation (cf. Table 6)
		DRES	Disinterested Revision Evaluation (cf. Table 6)
		DTES	Disinterested Total Evaluation (cf. Table 6)
Overall Translation Quality	Reflective	AQ1	The quality of this text is acceptable for a Canadian Francophone website.
		OTQ11	Most Canadian Francophones would not be able to tell that this text is a translation.
		OTQ21	Most Canadian Francophones would feel confident trusting the content of this translation.
		OTQ31	Most Canadian Francophone readers of this text would feel confident contacting the company in French.
		OTQ41	Most Canadian Francophone readers of this text would feel confident purchasing products of services from this company.

Table 56 – Disinterested Selection Hypothesis (Bootstrapping Results)

BOOTSTRAPPING (128 samples with 500 iterations/degrees of freedom)		T-STATISTIC VALUES
Structural Model	Disinterested Evaluation Selection → Overall Translation Quality	1.470667 (not statistically significant)
Measurement Model	DPES	2.146470
	DRES	0.996610
	DTES	2.352260
	AQ1	3.551285
	OTQ11	3.561976
	OTQ21	3.547268
	OTQ31	3.563814
	OTQ41	3.537282

C.4 Data Tables for H₂

Table 57 - H₂ (ACQ4)

ACQ4 - The quality of this text is acceptable for a Canadian Francophone website.									
Change	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	CSTLP2 %	CSTNLP2 %	CSTLP4 %	CSTNLP4 %	TOTAL
+ 4	0	0	1	0	0%	0%	3%	0%	1
+ 3	1	1	0	0	3%	3%	0%	0%	2
+ 2	1	1	3	8	3%	3%	9%	25%	13
+ 1	5	8	7	5	16%	25%	22%	16%	25
0	19	16	18	13	59%	50%	56%	41%	66
- 1	5	4	2	6	16%	13%	6%	19%	17
- 2	0	2	0	0	0%	6%	0%	0%	2
- 3	1	0	1	0	3%	0%	3%	0%	2
TOTAL	32	32	32	32	100%	100%	100%	100%	128
Confidence Interval					2.2147	1.9157	2.1116	1.6872	

Table 58 - H₂ (OTQ14)

(OTQ14) - Most Canadian Francophones would not be able to tell that this text is a translation.									
SHIFT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	CSTLP2 %	CSTNLP2 %	CSTLP4 %	CSTNLP4 %	TOTAL
+ 4	0	0	0	0	0%	0%	0%	0%	0
+ 3	1	0	1	0	3%	0%	3%	0%	2
+ 2	2	1	3	4	6%	3%	9%	13%	10
+ 1	5	8	5	5	16%	25%	16%	16%	23
0	17	16	19	16	53%	50%	59%	50%	68
- 1	6	1	3	6	19%	3%	9%	19%	16
- 2	1	5	0	1	3%	16%	0%	3%	7
- 3	0	1	1	0	0%	3%	3%	0%	2
TOTAL	32	32	32	32	100%	100%	100%	100%	128
Confidence Interval					1.9774	1.9424	2.1835	1.8796	

Table 59 - H₂ (OTQ24)

(OTQ24) - Most Canadian Francophones would feel confident trusting the content of this translation.									
SHIFT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	CSTLP2 %	CSTLNLP2 %	CSTLP4 %	CSTNLP4 %	TOTAL
+ 4	0	0	0	0	0%	0%	0%	0%	0
+ 3	0	1	1	0	0%	3%	3%	0%	2
+ 2	0	1	2	3	0%	3%	6%	9%	6
+ 1	3	4	7	7	9%	13%	22%	22%	21
0	23	16	17	17	72%	50%	53%	53%	73
- 1	5	8	4	4	16%	25%	13%	13%	21
- 2	0	2	0	1	0%	6%	0%	3%	3
- 3	1	0	1	0	3%	0%	3%	0%	2
TOTAL	32	32	32	32	100%	100%	100%	100%	128
Confidence Interval					2.7344	1.9157	1.9947	2.0118	

Table 60 - H₂ (OTQ34)

(OTQ34) - Most Canadian Francophone readers of this text would feel confident contacting the company in French.									
SHIFT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	CSTLP2 %	CSTLNLP2 %	CSTLP4 %	CSTNLP4 %	TOTAL
+ 4	0	0	1	0	0%	0%	3%	0%	1
+ 3	0	1	0	0	0%	3%	0%	0%	1
+ 2	0	2	1	4	0%	6%	3%	13%	7
+ 1	5	5	4	8	16%	16%	13%	25%	22
0	20	19	24	16	63%	59%	75%	50%	79
- 1	6	4	0	3	19%	13%	0%	9%	13
- 2	1	1	1	0	3%	3%	3%	0%	3
- 3	0	0	1	1	0%	0%	3%	3%	2
TOTAL	32	32	32	32	100%	100%	100%	100%	128
Confidence Interval					2.3933	2.1913	2.8330	1.9335	

Table 61 - H₂ (OTQ44)

(OTQ44) - Most Canadian Francophone readers of this text would feel confident contacting the company in French.									
SHIFT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	CSTLP2 %	CSTLNL2 %	CSTLP4 %	CSTNLP4 %	TOTAL
+ 4	0	0	0	0	0%	0%	0%	0%	0
+ 3	0	0	1	0	0%	0%	3%	0%	1
+ 2	0	2	1	4	0%	6%	3%	13%	7
+ 1	2	3	5	5	6%	9%	16%	16%	15
0	21	18	21	16	66%	56%	66%	50%	76
- 1	8	8	2	6	25%	25%	6%	19%	24
- 2	1	1	0	1	3%	3%	0%	3%	3
- 3	0	0	2	0	0%	0%	6%	0%	2
TOTAL	32	32	32	32	100%	100%	100%	100%	128
Confidence Interval					2.5595	2.1677	2.4429	1.8796	

Table 62 - H₂ (Shift for Overall Translation Quality)

SHIFT Overall Translation Quality - By Translation/Experiment Type									
SHIFT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	CSTLP2 %	CSTLNL2 %	CSTLP4 %	CSTNLP4 %	TOTAL
+ 4	0	0	2	0	0%	0%	1%	0%	2
+ 3	2	3	3	0	1%	2%	2%	0%	8
+ 2	3	7	10	23	2%	4%	6%	14%	43
+ 1	20	28	28	30	13%	18%	18%	19%	106
0	100	85	99	78	63%	53%	62%	49%	362
- 1	30	25	11	25	19%	16%	7%	16%	91
- 2	3	11	1	3	2%	7%	1%	2%	18
- 3	2	1	6	1	1%	1%	4%	1%	10
TOTAL	160	160	160	160	100%	100%	100%	100%	640
Confidence Interval					5.2793	4.3881	5.1244	4.1196	

C.5 Data Tables for H₄

Table 63 - H_{4A} (ACQ1)

(AQ1) - The quality of this text is acceptable for a Canadian Francophone website.										
SCALE CATEGORIES	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT
Strongly Agree	10	5	3	1	0	31.3%	15.6%	9.4%	3.1%	0.0%
Agree	16	6	12	4	0	50.0%	18.8%	37.5%	12.5%	0.0%
Neither Agree Nor Disagree	3	9	7	5	0	9.4%	28.1%	21.9%	15.6%	0.0%
Disagree	1	6	7	17	5	3.1%	18.8%	21.9%	53.1%	15.6%
Strongly Disagree	2	6	3	5	27	6.3%	18.8%	9.4%	15.6%	84.4%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Interval						4.1326	3.6514	3.7997	4.0917	5.1282

Table 64 - H_{4A} (OTQ11)

(OTQ11) - Most Canadian Francophones would not be able to tell that this text is a translation.										
SCALE CATEGORIES	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT
Strongly Agree	8	3	3	1	0	25.0%	9.4%	9.4%	3.1%	0.0%
Agree	16	11	13	11	0	50.0%	34.4%	40.6%	34.4%	0.0%
Neither Agree Nor Disagree	2	6	4	2	0	6.3%	18.8%	21.9%	6.3%	0.0%
Disagree	4	5	6	10	5	12.5%	15.6%	18.8%	31.3%	15.6%
Strongly Disagree	2	7	3	8	27	6.3%	21.9%	9.4%	25.0%	84.4%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Interval						4.0564	3.7359	3.8374	3.8933	5.182

Table 65 - H_{4A} (OTQ21)

(OTQ21) - Most Canadian Francophones would feel confident trusting the content of this translation.										
SCALE CATEGORIES	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT
Strongly Agree	12	5	4	2	0	37.5%	15.6%	12.5%	6.3%	0.0%
Agree	16	14	14	8	1	50.0%	43.8%	43.8%	25.0%	3.1%
Neither Agree Nor Disagree	3	7	7	10	2	9.4%	21.9%	21.9%	31.3%	6.3%
Disagree	0	3	4	6	8	0.0%	9.4%	12.5%	18.8%	25.0%
Strongly Disagree	1	3	3	6	21	3.1%	9.4%	9.4%	18.8%	65.6%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Interval						4.2472	3.8871	3.8809	3.7359	4.5210

Table 66 - H_{4A} (OTQ31)

(OTQ31) - Most Canadian Francophone readers of this text would feel confident contacting the company in French.										
SCALE CATEGORIES	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT
Strongly Agree	12	4	5	2	0	37.5%	12.5%	15.6%	6.3%	0.0%
Agree	13	13	15	10	1	40.6%	40.6%	46.9%	31.3%	3.1%
Neither Agree Nor Disagree	6	7	7	8	5	18.8%	21.9%	21.9%	25.0%	15.6%
Disagree	0	5	3	5	6	0.0%	15.6%	9.4%	15.6%	18.8%
Strongly Disagree	1	3	2	7	20	3.1%	9.4%	6.3%	21.9%	62.5%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Interval						4.0741	3.8249	3.9606	3.7424	4.3917

Table 67 - H_{4A} (OTQ41)

(OTQ41) – Most Canadian Francophone readers of this text would feel confident purchasing products of services from this company.										
SCALE CATEGORIES	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT
Strongly Agree	11	6	5	2	0	34.4%	18.8%	15.6%	6.3%	0.0%
Agree	16	10	16	11	2	50.0%	31.3%	50.0%	34.4%	6.3%
Neither Agree Nor Disagree	4	11	7	10	1	12.5%	34.4%	21.9%	31.3%	3.1%
Disagree	0	3	2	4	14	0.0%	9.4%	6.3%	12.5%	43.8%
Strongly Disagree	1	2	2	5	15	3.1%	6.3%	6.3%	15.6%	46.9%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Interval						4.2018	3.8311	4.0386	3.8186	4.2922

Table 68 - H_{4A} (Overall Translation Quality – Five Tiers)

Overall Translation Quality - By Translation/Experiment Type										
SCALE CATEGORIES	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT
Strongly Agree	53	23	20	8	0	33.1%	14.4%	12.5%	5.0%	0.0%
Agree	77	54	70	44	4	48.1%	33.8%	43.8%	27.5%	2.5%
Neither Agree Nor Disagree	18	40	35	35	8	11.3%	25.0%	21.9%	21.9%	5.0%
Disagree	5	22	22	42	38	3.1%	13.8%	13.8%	26.3%	23.8%
Strongly Disagree	7	21	13	31	110	4.4%	13.1%	8.1%	19.4%	68.8%
TOTAL	160	160	160	160	160	100%	100%	100%	100%	100%
Confidence Interval						9.2111	8.3452	8.6858	8.3394	10.3148

Table 69 - H₄ (MT vs. CST Rankings in Six Tiers)

Before Knowledge of Methods - Rank – Six Tiers										
RANK ORDER	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT	% CSTLP2	% CSTNLP2	% CSTLP4	% CSTNLP4	% MT
1st Place	16	2	3	1	0	50%	6%	9%	3%	0%
2nd Place	8	5	10	2	0	25%	16%	31%	6%	0%
3rd Place	3	9	9	2	0	9%	28%	28%	6%	0%
4th Place	2	11	6	9	0	6%	34%	19%	28%	0%
5th Place	3	5	4	17	1	9%	16%	13%	53%	3%
6th Place	0	0	0	1	31	0%	0%	0%	3%	97%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Interval						2.0282	1.4314	1.3087	2.2419	4.3588

Table 70 - H₄ (MT vs. CST Rankings in Two Tiers)

Before Knowledge of Methods - Rank – Two Tiers										
RANK ORDER	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	MT	% CSTLP2	% CSTNLP2	% CSTLP4	% CSTNLP4	% MT
1 st – 3 rd	27	16	22	5	0	84%	50%	69%	16%	0%
4 th – 6 th	5	16	10	27	32	16%	50%	31%	84%	100%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%

Table 71 - H₄ (MT vs. CST Aggregated Rank in Six Tiers)

Before Knowledge of Methods – Aggregated Rank - Six Tiers				
RANK ORDER	CST	MT	CST %	MT %
1st Place	22	0	17%	0%
2nd Place	25	0	20%	0%
3rd Place	23	0	18%	0%
4th Place	28	0	22%	0%
5th Place	29	1	23%	3%
6th Place	1	31	1%	97%
TOTAL	128	32	100%	100%
Confidence Interval			1.7892	4.3588

C.6 Data Tables for H₅

Table 72 – H_{5A} (ACQ1)

(AQ1) - The quality of this text is acceptable for a Canadian Francophone website.										
SCALE CATEGORIES	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT
Strongly Agree	10	5	3	1	7	31.3%	15.6%	9.4%	3.1%	21.9%
Agree	16	6	12	4	17	50.0%	18.8%	37.5%	12.5%	53.1%
Neither Agree Nor Disagree	3	9	7	5	4	9.4%	28.1%	21.9%	15.6%	12.5%
Disagree	1	6	7	17	2	3.1%	18.8%	21.9%	53.1%	6.3%
Strongly Disagree	2	6	3	5	2	6.3%	18.8%	9.4%	15.6%	6.3%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Interval						4.1326	3.6514	3.7997	4.0917	4.1093

Table 73 – H_{5A} (OTQ11)

(OTQ11) - Most Canadian Francophones would not be able to tell that this text is a translation.										
SCALE CATEGORIES	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT
Strongly Agree	8	3	3	1	8	25.0%	9.4%	9.4%	3.1%	25.0%
Agree	16	11	13	11	14	50.0%	34.4%	40.6%	34.4%	43.8%
Neither Agree Nor Disagree	2	6	4	2	3	6.3%	18.8%	21.9%	6.3%	9.4%
Disagree	4	5	6	10	6	12.5%	15.6%	18.8%	31.3%	18.8%
Strongly Disagree	2	7	3	8	1	6.3%	21.9%	9.4%	25.0%	3.1%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Interval						4.0564	3.7359	3.8374	3.8933	3.9423

Table 74 – H_{5A} (OTQ21)

(OTQ21) - Most Canadian Francophones would feel confident trusting the content of this translation.										
SCALE CATEGORIES	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT
Strongly Agree	12	5	4	2	6	37.5%	15.6%	12.5%	6.3%	18.8%
Agree	16	14	14	8	19	50.0%	43.8%	43.8%	25.0%	59.4%
Neither Agree Nor Disagree	3	7	7	10	3	9.4%	21.9%	21.9%	31.3%	9.4%
Disagree	0	3	4	6	3	0.0%	9.4%	12.5%	18.8%	9.4%
Strongly Disagree	1	3	3	6	1	3.1%	9.4%	9.4%	18.8%	3.1%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Interval						4.2472	3.8871	3.8809	3.7359	4.2641

Table 75 – H_{5A} (OTQ31)

(OTQ31) - Most Canadian Francophone readers of this text would feel confident contacting the company in French.										
SCALE CATEGORIES	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT
Strongly Agree	12	4	5	2	8	37.5%	12.5%	15.6%	6.3%	25.0%
Agree	13	13	15	10	19	40.6%	40.6%	46.9%	31.3%	59.4%
Neither Agree Nor Disagree	6	7	7	8	2	18.8%	21.9%	21.9%	25.0%	6.3%
Disagree	0	5	3	5	3	0.0%	15.6%	9.4%	15.6%	9.4%
Strongly Disagree	1	3	2	7	0	3.1%	9.4%	6.3%	21.9%	0.0%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Interval						4.0741	3.8249	3.9606	3.7424	4.3256

Table 76 – H_{5A} (OTQ41)

(OTQ41) – Most Canadian Francophone readers of this text would feel confident purchasing products of services from this company.										
SCALE CATEGORIES	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT
Strongly Agree	11	6	5	2	5	34.4%	18.8%	15.6%	6.3%	15.6%
Agree	16	10	16	11	18	50.0%	31.3%	50.0%	34.4%	56.3%
Neither Agree Nor Disagree	4	11	7	10	7	12.5%	34.4%	21.9%	31.3%	21.9%
Disagree	0	3	2	4	1	0.0%	9.4%	6.3%	12.5%	3.1%
Strongly Disagree	1	2	2	5	1	3.1%	6.3%	6.3%	15.6%	3.1%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Interval						4.2018	3.8311	4.0386	3.8186	4.2189

Table 77 – H_{5A} (Aggregated Overall Translation Quality - Five Tiers)

Aggregated Overall Translation Quality – Five Tiers				
SCALE CATEGORIES	PT	CST	PT %	CST %
Strongly Agree	34	104	21.3%	16.3%
Agree	87	245	54.4%	38.3%
Neither Agree Nor Disagree	19	128	11.9%	20.0%
Disagree	15	91	9.4%	14.2%
Strongly Disagree	5	72	3.1%	11.3%
TOTAL	160	640	100%	100%
Confidence Interval			5.0305	5.3071

Table 78 - H_{5A} (PT vs. CST Rankings in Six Tiers)

Before Knowledge of Methods - Rank – Six Tiers										
RANK ORDER	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT	% CSTLP2	% CSTNLP2	% CSTLP4	% CSTNLP4	% PT
1st Place	16	2	3	1	10	50%	6%	9%	3%	31%
2nd Place	8	5	10	2	7	25%	16%	31%	6%	22%
3rd Place	3	9	9	2	9	9%	28%	28%	6%	28%
4th Place	2	11	6	9	4	6%	34%	19%	28%	13%
5th Place	3	5	4	17	2	9%	16%	13%	53%	6%
6th Place	0	0	0	1	0	0%	0%	0%	3%	0%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Interval						2.0282	1.4314	1.3087	2.2419	1.3801

Table 79 - H_{5A} (PT vs. CST Rankings in Two Tiers)

Before Knowledge of Methods - Rank – Two Tiers										
RANK ORDER	CSTLP2	CSTNLP2	CSTLP4	CSTNLP4	PT	% CSTLP2	% CSTNLP2	% CSTLP4	% CSTNLP4	% PT
1 st – 3 rd	27	16	22	5	26	84%	50%	69%	16%	81%
4 th – 6 th	5	16	10	27	6	16%	50%	31%	84%	19%
TOTAL	32	32	32	32	32	100%	100%	100%	100%	100%

Table 80 - H_{5A} (PT vs. CST Aggregated Rank in Six Tiers)

Before Knowledge of Methods – Aggregated Rank - Six Tiers				
RANK ORDER	CST	PT	CST %	PT %
1st Place	22	10	17%	31%
2nd Place	25	7	20%	22%
3rd Place	23	9	18%	28%
4th Place	28	4	22%	13%
5th Place	29	2	23%	6%
6th Place	1	0	1%	0
TOTAL	128	32	100%	100%
Confidence Interval			1.7892	1.3807

C.7 Crowdsourced Translation Acceptance Data

Table 81 - Crowdsourced Translation Acceptance - Five Tiers

SCALE CATEGORIES	CTA1	CTA2	CTA3	CTA4	CTA5	CTA1	CTA2	CTA3	CTA4	CTA5
Strongly Agree	19	1	0	9	2	59%	3%	0%	28%	6%
Agree	10	6	1	15	15	31%	19%	3%	47%	47%
Neither Agree Nor Disagree	2	8	10	6	7	6%	25%	31%	19%	22%
Disagree	0	12	15	1	5	0%	38%	47%	3%	16%
Strongly Disagree	1	5	6	1	3	3%	16%	19%	3%	9%
TOTALS	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Intervals						2.7998	1.3988	2.1720	2.0439	1.7937

Table 82 - Crowdsourced Translation Acceptance - Three Tiers

SCALE CATEGORIES	CTA1	CTA2	CTA3	CTA4	CTA5	CTA1	CTA2	CTA3	CTA4	CTA5
Strongly Agree - Agree	29	7	1	24	17	91%	22%	3%	75%	53%
Neutral	2	8	10	6	7	6%	25%	31%	19%	22%
Strongly Disagree - Disagree	1	17	21	2	8	3%	53%	66%	6%	25%
TOTALS	32	32	32	32	32	100%	100%	100%	100%	100%
Confidence Intervals						5.5038	1.9082	3.4705	4.0603	1.9082

C.8 Survey Instructions – Part I



SURVEY INSTRUCTIONS – PART I

Dear Participant,

You have been selected and consented to participate in my study of the use of Web 2.0 technologies to harness collective intelligence for translation. There are two parts requiring completion, each which you will find in its own large envelope. In Part I, you shall review a single Canadian English source text from which multiple Canadian French translations (a.k.a. target texts) have been produced. A short questionnaire will be completed to first assess the quality of the source text, which is approximately 300 words in length and is of a business/marketing nature. You shall then use additional questionnaires to evaluate the quality of multiple Canadian French translations.

INSTRUCTIONS

1. **Open the envelope marked PART I.** In it, you will find the following items:
 - a. One “Canadian English Source Text”
 - i. One “Source Text Evaluation Survey”
 - b. Six “Canadian French Target Text”
 - i. Six “Target Text Evaluation Survey”.
2. **Read all of the enclosed material.**
3. **Complete the Source Text Evaluation Survey.**
 - a. Staple the completed Source Text Evaluation Survey to the “Canadian English Source Text”
4. **For each Target Text...**
 - a. Complete a Target Text Evaluation Survey.
 - b. Staple the completed Target Text Evaluation Survey to its Target Text
5. **After completing all Target Text Evaluation Surveys...**
 - a. Using your judgement, place the stapled Target Texts **in order** from best quality to worst quality.
 - i. The best quality Target Text should be on the top
 - ii. The worst quality Target Text on the bottom.
 - b. On the upper right corner of each Target Text Evaluation Survey, write the number that represents the rank of the target text.
 - i. The best quality target text, which should be on top, should be marked as #1.
 - ii. The worst quality target text, which should be on the bottom, should be marked as #6.
6. **Return all of the Evaluation Survey materials back into the envelope marked PART I.**
7. **Seal the envelope marked PART I.**
8. **Proceed to open the envelope marked PART II.**
 - a. Inside you will find instructions and material on how to complete the next step of your participation.

You have two weeks to complete and return all survey materials. Should you have any questions, you may contact me via email.

Sincerely,

Steven O'Brien
Graduate Student – Master of Science – Electronic Business Technologies - University of Ottawa

C.9 Source Text Evaluation Survey



SOURCE TEXT EVALUATION SURVEY

For each statement, circle your choice

1. **This text uses correct terminology.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
2. **This text uses correct grammar.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
3. **This text has an understandable meaning.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
4. **This text is written in a style the calls the reader to action.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
5. **This text is written for an educated middle class audience.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
6. **This text is written in the style of a professional blogger.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
7. **This text is written in a style that is intended to be read instead of a style that is intended to be spoken.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
8. **This text is written in a style that is:**
1 – Frozen, 2 – Formal, 3 – Consultative, 4 – Casual, 5 – Intimate
9. **This text appears to be written by a fluent Canadian Anglophone.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
10. **This text appears to be written by someone with a university education.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
11. **This text is written using language that was appropriate in 2010.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree

C.10 Target Text Evaluation Survey



TARGET TEXT EVALUATION SURVEY For each statement, circle your choice

RANK:

1. **This text uses correct terminology.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
2. **This text uses correct grammar.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
3. **This text has an understandable meaning.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
4. **This text is written in a style the calls the reader to action.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
5. **This text is written for an educated middle class audience.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
6. **This text is written in the style of a professional blogger.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
7. **This text is written in a style that is intended to be read instead of a style that is intended to be spoken.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
8. **This text is written in a style that is:**
1 – Frozen, 2 – Formal, 3 – Consultative, 4 – Casual, 5 – Intimate
9. **This text appears to be written by a fluent Canadian Francophone.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
10. **This text appears to be written by someone with a university education.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
11. **This text is written using language that was appropriate in 2010.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
12. **The quality of this text is acceptable for a Canadian Francophone website.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
13. **Most Canadian Francophones would not be able to tell that this text is a translation.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
14. **Most Canadian Francophones would feel confident trusting the content of this translation.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
15. **Most Canadian Francophone readers of this text would feel confident contacting the company in French.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree
16. **Most Canadian Francophone readers of this text would feel confident purchasing products or services from this company.**
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree

C.11 Survey Instructions – Part II



SURVEY INSTRUCTIONS – PART II

Dear Participant,

If you are reading this, it should be because you have completed all of the instructions for PART I. If this is not the case, please do so before you complete PART II.

INSTRUCTIONS

1. In the envelope marked **PART II** are the following items:
 - a. Six "Canadian French Target Text"
 - i. Six "Target Text Evaluation Survey".
 - b. One "Supplemental Evaluation Survey".
2. Read all of the enclosed material.
 - a. At the top of each Target Text, you will find information describing how the translation was created.
3. For each Target Text...
 - a. Complete a Target Text Evaluation Survey.
 - b. Staple the completed Target Text Evaluation Survey to its Target Text
4. After Completing All Target Text Evaluation Surveys...
 - a. Using your judgement, place the stapled Target Texts in order from best quality to worst quality.
 - i. The best quality Target Text should be on the top.
 - ii. The worst quality Target Text on the bottom.
 - b. On the upper right corner of each Target Text Evaluation Survey, write the number that represents the rank of the target text.
 - i. The best quality target text, which should be on top, should be marked as #1.
 - ii. The worst quality target text, which should be on the bottom, should be marked as #6.
5. Complete the Supplemental Evaluation Survey.
6. Return all of the Evaluation Survey materials back into the envelope marked **PART II**.
7. Place the envelopes (**PART I** and **PART II**) inside of the envelope marked **TRANSLATION RESEARCH**.
8. Seal the envelope marked **TRANSLATION RESEARCH**.
9. Send an email to [REDACTED] indicating that you have completed your participation in the study.
 - a. Include in the email instructions on how you would like me to pick up the completed surveys.

You have two weeks to complete and return all survey materials. Should you have any questions, you may contact me via email.

Sincerely,

Steven O'Brien
Graduate Student – Master of Science – Electronic Business Technologies - University of Ottawa

C.12 Supplemental Evaluation Survey



SUPPLEMENTAL EVALUATION SURVEY

For each statement, circle your choice

1. I prefer crowdsourced translation more than machine translation.
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree

2. I prefer crowdsourced translation more than professional translation.
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree

3. Crowdsourced translation should be performed by non-language professionals.
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree

4. Crowdsourced translation should be performed by language professionals.
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree

5. Crowdsourced translation should be used to translate website content in Canada.
1 – Strongly Disagree, 2 – Disagree, 3 – Neither Disagree nor Agree, 4 – Agree, 5 – Strongly Agree

C.13 Screening Letter



Greetings,

I am looking for research participants who are bilingual in both Canadian English and Canadian French. I would like to invite you to participate in my study of the use of Web 2.0 technologies to harness collective intelligence for translation. The purpose of the study is to gather data to understand better if crowdsourced translation may be an effective means of bridging the Canadian linguistic digital divide.

For this portion of the study, participants shall review a single Canadian English source text from which multiple Canadian French translations have been produced. Participants shall complete a short questionnaire to first assess the quality of the source text, which is approximately 300 words in length and is of a business/marketing nature. Participants shall then use additional questionnaires to evaluate the quality of multiple Canadian French translations.

I anticipate that the time required to participate in this part of the study should take between one and two hours depending upon the speed at which each participant works. Each participant shall have a maximum of two weeks to complete and return the work for this part of the study. Each participant shall be paid \$40 in cash for their participation in this part of the study.

Please note that participation is voluntary and participants may refuse and withdraw at any time. If you are interested in participating in this research, I ask that you answer the following qualification screening questions:

1. Please indicate the highest level of education you have achieved:
1 - High School, 2 - Some College/University, 3 - Bachelor's Degree, 4 - Master's degree, 5 - Ph.D.

2. Please indicate the discipline of each degree (if any).

3. Please write the number of years you have worked as a professional translator (if any):

4. Please write the names of any professional translation certifications that you have acquired (if any):

5. Yes or No – Do you consider yourself to be fluent in Canadian English? _____

6. Yes or No – Do you consider yourself to be fluent in Canadian French? _____

Upon receiving confirmation from you indicating your interest in the study and your answers to these questions, I will respond to you and let you know if you have qualified to participate in this part of the study. In the event that you do not qualify, participants may be offered to join a research participant group that is more appropriate for their background.

In the event that you do qualify, you will be asked to complete and sign a research participant consent form. Upon your successful completion and return of the research participant consent form, you will receive additional information from me enabling you to participate. The research participant consent form may be returned to me via email or in person (if applicable).

Sincerely,

Steven O'Brien
Graduate Student – Master of Science – Electronic Business Technologies - University of Ottawa

C.14 Statement of Informed Consent



Title of the study:

HARNESSING COLLECTIVE INTELLIGENCE FOR TRANSLATION: AN ASSESSMENT OF CROWDSOURCING AS A MEANS OF BRIDGING THE CANADIAN LINGUISTIC DIGITAL DIVIDE

Name of researcher: Steven O'Brien, Electronic Business Technologies Department, Faculty of Graduate and Postdoctoral Studies, University of Ottawa, Email: [REDACTED]

Thesis Supervisor: Dr. Ajax Persaud, Telfer School of Management, Faculty of Graduate and Postdoctoral Studies, [REDACTED]

Invitation to Participate: I am invited to participate in the above mentioned research study conducted by Steven O'Brien and Dr. Ajax Persaud.

Purpose of the Study: The purpose of the study is to assess the effectiveness of applying crowdsourced translation techniques to website content as a means of bridging the linguistic digital divide within Canada.

Participation: My participation will consist essentially of evaluating the quality of several Canadian French translations and the associated Canadian English source text. I will be asked to rank Canadian French translations in order of quality from best to worst. I will perform these tasks by filling out a survey and placing the translations in order of quality, from best to worst. I understand that the researcher estimates that this should take no more than two hours to complete. I understand that I may complete these tasks at any time and that I will have a period of two weeks from the time the researcher has acknowledged the receipt of this completed consent form.

Risks: I do not perceive any risk in participating in this study.

Benefits: My participation in this study will allow the researcher to collect data for the purposes of validating if the technique of crowdsourcing is appropriate and potentially effective as a means of bridging the linguistic digital divide within Canada. The research conducted by this study may be useful for others who wish to further examine solutions to this or similar problems. My participation in this study may help to facilitate these goals. Specifically, my participation will contribute to a better understanding of the types of translation errors that diverse groups of people are good at finding and correcting in addition to which aspects of crowdsourced translation were found to be the most acceptable for bridging the linguistic digital divide within Canada.

Confidentiality and anonymity: By signing this consent form, I understand and consent to the following: that the information that I will share with the researcher shall remain strictly confidential and will be used only for the purpose of this study; that my confidentiality will be protected by the researcher; that any personally identifying information collected by the researcher shall be removed and replaced with unique codes by the researcher after the data collection phase of the research; and that at no time shall my identity be published.

Conservation of data: The data collected will be conserved for a period of five years following the successful completion of the thesis. During the period of conservation, all data shall be securely stored in the home of the researcher, under lock and key. Following the period of conservation, all electronic data shall be securely deleted and all physical records such as completed surveys and translations shall be shredded.

Compensation: I understand that I will be paid \$40 in cash for my participation in this study.

Voluntary Participation: I am under no obligation to participate and if I choose not to participate, I can withdraw from the study at any time and/or refuse to answer any questions, without suffering any negative consequences. If I choose to withdraw, all data gathered until the time of withdrawal may be used for data analysis purposes by the researcher.

Acceptance: I, _____ agree to participate in the above research study conducted by Steven O'Brien of the Electronic Business Technologies Department, Faculty of Graduate and Postdoctoral Studies, University of Ottawa, whose research is under the supervision of Dr. Ajax Persaud.

If I have any questions about the study, I may contact the researcher or his supervisor, Dr. Ajax Persaud.

If I have any questions regarding the ethical conduct of this study, I may contact the Protocol Officer for Ethics in Research, University of Ottawa, [REDACTED]

Should I consent, I shall return this document to the researcher, indicating my name and the date in which I gave my consent.

Participant's signature:

Date:

Researcher's signature:

Date: