

Can machine translation facilitate outreach to newcomers? A pilot study investigating the needs of Spanish-speaking users of the Ottawa Public Library

Lynne Bowker

School of Information Studies

University of Ottawa

lbowker@uottawa.ca

Final Report

2013 OCLC/ALISE Library and Information Science Research Grant Project

08 March 2014

© 2014 Lynne Bowker

Published by permission.

<http://www.oclc.org/research/grants/>

Reproduction of substantial portions of this publication must contain the copyright notice.

Suggested citation:

Bowker, Lynne. 2014. "Can machine translation facilitate outreach to newcomers? A pilot study investigating the needs of Spanish-speaking users of the Ottawa Public Library." 2013 OCLC/ALISE research grant report published electronically by OCLC Research. <http://www.oclc.org/research/grants/reports/2013/bowker2013.pdf>

EXECUTIVE SUMMARY

This report describes a research project conducted between January and December 2013 under the direction of Dr. Lynne Bowker (University of Ottawa) with funding from the OCLC/ALISE Library & Information Science Research Grant Program. The project sought to investigate the potential of machine translation for aiding Spanish-speaking newcomers to Canada's National Capital Region to make better use of the Ottawa Public Library's (OPL) website.

This investigation adopted a community-based participatory approach to research in which academic researchers worked alongside the community of Spanish-speaking immigrants in Ottawa, and the Ottawa Public Library, which seeks to offer improved services to this community. Community participation helps to ensure that the goals of the study are relevant to the population, which can in turn help to identify service and investment priorities and to better allocate resources.

Three professional translators were enlisted to test three different machine translation systems – Google Translate, Reverso and Systran – and to evaluate which seemed most promising as a tool for helping to translate the OPL website into Spanish. They also produced translated and post-edited texts.

One-hundred and fourteen community participants responded to a survey in which they provided some basic demographic information and also indicated the reasons why they would like to have access to the OPL website in Spanish.

These participants then participated in a recipient evaluation in which they were asked to read four different translated versions of a text, which had been produced using four different translation methods: human translation, maximally post-edited machine translation, rapidly post-edited machine translation, and raw machine translation. In the first instance, participants simply read four unlabelled versions of the translated text and indicated their preference. In a subsequent step, participants were informed how much time and money were required to produce each version, and, taking this information into account, they were asked to once again choose the text that best meets their needs.

This study revealed that many – though not all – of the translation-related needs of this community can be satisfied with texts that are semantically accurate, but which need not be stylistically elegant. For example, in many cases, community members use the translated texts in order to be able to process information more quickly, or to confirm their understanding of the original text. This experiment demonstrated that techniques such as maximal and rapid post-editing can be combined with automatic machine translation to provide translated texts that sufficiently address these needs for a high percentage of the population. Moreover, post-edited texts can be produced in a quicker turnaround time and at a fraction of the cost of professionally translated texts. In the case of rapid post-editing, texts produced as part of this

Bowker, Lynne: Can machine translation facilitate outreach to newcomers? A pilot...

experiment offered a cost savings of 67 percent over human translation, and could be produced in one-third of the time.

This study also confirmed that some types of texts, such as texts written in an instructional format, appear to have a particularly high acceptance rate among the community when produced using raw machine translation or rapid post-editing.

Overall, the results of this small-scale pilot study appear to be promising and would seem to merit additional investigation.

CONTENTS

EXECUTIVE SUMMARY.....2

ACKNOWLEDGMENTS5

1. INNOVATIVE ASPECTS OF THE RESEARCH6

2. RESEARCH QUESTION7

3. OVERVIEW OF PROJECT SCOPE AND COMPLETED PHASES OF RESEARCH 11

4. METHODOLOGY AND DISCUSSION OF FINDINGS..... 11

 4.1 Phase I: Evaluating machine translation systems..... 12

 4.1.1 Phase I Findings and Discussion 13

 4.2 Phase II: Surveying the community about their translation needs 14

 4.2.1 Phase II Findings and Discussion 16

 4.3 Phase III: Conducting a recipient evaluation of translated texts 18

 4.3.1 Preparatory work..... 20

 4.3.2 Recipient evaluation..... 22

 4.3.3 Phase III Findings and Discussion 22

5. OUTCOMES AND OUTPUTS 24

6. FUTURE RESEARCH 25

7. BUDGET REPORT..... **Error! Bookmark not defined.**

REFERENCES 27

APPENDICES..... **Error! Bookmark not defined.**

ACKNOWLEDGMENTS

We are grateful to the following people and organizations who contributed to and supported the research reported here:

- Jairo Buitrago Ciro, project research assistant and student on the Master of Information Studies program at the University of Ottawa
- Ottawa Public Library, and in particular Alexandra Yarrow (Acting Manager of Diversity and Accessibility Services); Céline St-Martin (Newcomer Services Librarian); and Craig Ginther (Manager of Digital Services)
- Translators, Raúl Colon Rodriguez, Daniel Josephy Hernandez and Julian Zapata Rojas
- Settlement workers at the Ottawa Community Immigrant Services Organization (OCISO) and the Ottawa Catholic Immigration Centre
- Chaplains at the Iglesia Sagrada Familia and the Cathédrale Saint-Joseph
- Embassies of Colombia, El Salvador and Guatemala
- Online newspaper *Mundo Español* (Ottawa edition) and the Ottawa-based distribution list "Spanish-language meet-up"
- Members of the Spanish-speaking newcomer community in the Ottawa region who participated in the survey and recipient evaluation
- ALISE/OCLC Library and Information Science Research Grant Program (2013)
- University of Ottawa

1. INNOVATIVE ASPECTS OF THE RESEARCH

In August 2013, the International Federation of Library Associations and Institutions (IFLA) released an [Insights Document](#) which pulls together and summarizes all of the information contained on IFLA's Trend Report website. The [IFLA Trend Report](#) identifies five key high level trends in the global information environment, setting out existing and likely future trends which characterize the new digital paradigm in libraries. One of these trends focuses on the fact that new technologies are transforming the global information economy, and a particular technology that is identified as being important is machine translation. IFLA raises some interesting questions about the use of such technologies in a library setting, such as "Machine translation will change the way we communicate, but will it increase our understanding?" and "What is the cultural impact of using machine translations without the benefit of cultural context?" IFLA has only very recently identified these questions as being important to the library and information science community, and to the best of our knowledge, no research has yet been carried out on the potential use of machine translation as a tool for helping newcomers to make better use of the public library. Our work is therefore an original contribution.

To carry out this investigation, we adopted an approach that has not yet been widely applied in either the library science or the translation fields: community-based participatory research. CBPR is a collaborative approach that includes not only researchers, but also community members and community service providers. It involves meaningful participation by all groups, with the recognition that each brings different strengths. CBPR seeks to increase knowledge and understanding of a given phenomenon and to integrate the knowledge gained to improve the quality of life of community members. The principles of CBPR were first elaborated (e.g. Israel *et al.* 1998) – and are still most often applied – within the context of public health research (e.g. Minkler and Wallerstein 2008; Horowitz *et al.* 2009; Willison 2013). However, they can be adapted and extended to apply to research conducted in other areas.

In CBPR, the intent is to transform research from a relationship where researchers *carry out investigations on a* community to one where researchers work *with* a community, in order to decide on the questions and methods and to disseminate and apply the findings (Fitzpatrick 2013: 14). In other words, community members become part of the research team and researchers become engaged in the activities of the community. Unlike many forms of conventional research, CBPR challenges researchers to listen to, learn from, solicit and respect the contributions of the groups that they are trying learn about and help.

Within the general CBPR approach, we carried out a form of machine translation evaluation known as a recipient evaluation. Although the literature on machine translation evaluation is quite extensive, comparatively little attention is paid to recipient evaluations. While it is likely that many organizations that employ machine translation do conduct some type of recipient evaluation, the details of these evaluations are not widely published. A literature survey turned up a number of indirect references to recipient evaluations, but for most of these, little specific information about the evaluation procedure was provided, and the sparse information that was available was largely anecdotal in nature. There is therefore a need to develop a more systematic framework for carrying out recipient evaluations, and the research reported here is a step in that direction.

Finally, many machine translation evaluation projects focus exclusively on the quality of the output. As such, they often take place in an academic or research context, where real-world considerations, such as deadlines or labour costs, are not taken into consideration. We have situated this project firmly within a real-world context, and we have examined the delivery of translation services for the newcomer community from a business perspective, where the quality of the translation must be weighed against the constraints of time and cost.

2. RESEARCH QUESTION

As reported by MacDonald (2012: 2), Canada's contemporary response to the challenges associated with immigrant integration has been to turn to cities and municipalities to administer immigrant settlement policies and programs. In particular, there has been a growing interest in the role that public libraries, as municipal institutions, play in immigrant settlement and integration in Canada (e.g. Frisken and Wallace 2000). Indeed, public libraries have invested much time and effort into providing newcomers to Canada with information and resources in various forms, including efforts that specifically aim to address the language-related needs of newcomers. For example, the Newcomer Services at OPL attest to an active Spanish speaking newcomer community, noting that in 2011, a total of 12,559 Spanish-language items circulated, which represents a 37 percent increase since 2008 (not including the Spanish titles offered as part of the digital collection). This is in line with information reported in the Census of Canada, which notes that Spanish is the second-most common immigrant language spoke in the Ottawa area (after Arabic). The Census reports that in 2011 there were 10,930 Spanish-speakers in Ottawa, representing 1.3 percent of the city's total population of 856,845 residents, which is up from 9,860 Spanish-speakers (1.2 percent) in 2006.

However, as reported in MacDonald (2012: 2), historically, public libraries have struggled with how to implement multiculturalism beyond collection development. While multilingual collections go some way towards meeting the needs of newcomers, there has been increasing recognition (e.g. George 2002; Jang *et al.* 2004; Caidi and Allard 2005) that newcomers have other types of information-related needs, including the need to access information about library and other services in their own language (Holt 2009: 123).

As observed by MacDonald (2012: 2), the problem of immigrant integration had once been defined as one of systemic inequality in access to services that resulted from a lack of coordination and integration between service providers; however, more recently, the responsibility for integration has shifted to the newcomers themselves, where the problem is being framed as a lack of awareness about settlement services. Citizenship and Immigration Canada (CIC) has recognized that many newcomers turn to public libraries when they first arrive in this country and, in response, is attempting to increase visibility of services in these venues. In 2008, CIC formally launched the Library Settlement Partnerships (LSP) program in which partnerships between 49 public libraries and 23 immigrant-serving organizations were established in 11 Ontario communities that have high newcomer populations requiring services in various languages. This program helps newcomers settle into the community by linking individuals and families to community and settlement services through partner-led group information sessions. OPL is one of the public library partners participating in the LSP program.

The partner-led sessions may address issues such as job-seeking, housing support, legal aid, English language training, support for students, cross-cultural education, and community activities and events, among others. However, it remains the responsibility of the library to provide information about library-specific services. This is challenging given the immense linguistic diversity of the immigrant population in the region. Quite understandably, the two Newcomer Services Librarians employed by OPL are not able to address this broad linguistic spectrum directly. Therefore, other potential solutions may be of interest.

It has been suggested that a useful strategy for making a library more accessible to newcomers is to translate the library's website into the most common languages spoken among local immigrants (e.g. Holt 2009: 123). As part of their effort to facilitate access to information for newcomers who speak languages other than English or French, OPL has undertaken to translate some of the material describing their services into nine languages: Arabic, simplified and traditional Chinese, Hindi, Russian, Somali, Spanish and Urdu. These translations, which have been produced by professional translators, are available on a section of the OPL website titled "Newcomers Services".¹ These translated sections include information organized into the following categories:

- Get a library card
- Access material in your language
- Improve your language skills (English or French)

While this laudable, it is nonetheless important to put this effort into context. The total length of the original English text that has been translated into these nine languages is less than 1000 words in length, while the amount of material on the OPL website that is available in Canada's official languages of English and French totals hundreds of thousands of words. Moreover, as observed by the OPL librarians, material on the website is updated regularly, which means that corresponding updates need to be made in each of the translated versions, which is labour-intensive and costly.

Unsurprisingly, cost and time are the principal reasons why a greater volume of information is not translated by more public institutions, including public libraries. For example, a Canadian professional translators' association – the *Ordre des traducteurs, terminologues et interprètes agréés du Québec* – recently conducted a survey of rates and income (Gauthier 2012). The resulting report indicates that based on 336 respondents, the average rate charged by Canadian professional translators for translating from English or French into another language is 22 cents per word (Gauthier 2012:3). Therefore, the cost of translating a 1000-word text into one language would be approximately \$220 (+13% tax = \$228.60), while the cost of translating that text into ten languages would be about \$2,200 (+13% tax = \$2,286). It is not hard to see that the cost of providing information in a range of immigrant languages quickly adds up. Moreover, as previously mentioned, not only is translation a costly service, it does not represent a one-off cost but rather requires an ongoing commitment as web content is updated regularly.

¹ The Newcomers Services section of the OPL website may be accessed at the following location: <http://bibliooottawalibrary.ca/en/interest/newcomers> [Retrieved February 22, 2014]

Like all public institutions, libraries are under pressure to reduce expenses, and multilingual services may be seen as an easy target. For example, the Canadian Broadcasting Company reported a story where, faced with having to make cuts, the city of Toronto's budget chief suggested that libraries might be able to preserve current operating hours if fewer resources were spent on foreign-language materials (CBC 2011). The chair of the Toronto Public Library board defended the multilingual collection as being critical for helping newcomers to integrate to life in the city, but it is clear that the offer of services in immigrant languages is vulnerable.

Therefore, while translation plays a key role in helping newcomers integrate into their communities, it can be costly and time-consuming, thus limiting the range of foreign-language information services that libraries are able to provide. For immigration-related policies such as settlement and multiculturalism to succeed, their associated costs must be kept down; however, if a cost-effective means of translation were available, public libraries might feasibly offer additional services to newcomers in a range of languages.

Indeed the OPL librarians were in favour of exploring possibilities for making a greater portion of the OPL website available in a range of languages. In fact, OPL had already begun discussions about whether to integrate a Google Translate widget directly into the OPL website as a translation solution, in keeping with the strategy that has already been adopted by the City of Ottawa for their own website.²

The City of Ottawa currently makes the information on its website available in the two official languages of Canada – English and French. However, the City's website has also integrated a Google Translate widget so that anyone can simply go to the dropdown list, select one of the languages listed, and click on Translate. The website content will then be automatically translated into that language by the Google Translate machine translation system. The City's website also includes the following disclaimer:

The material on this website is provided as general information only and is not intended to provide specific legal advice for any individual and should not be relied upon in that regard. [...]

The City of Ottawa assumes no responsibility for the accuracy of translations that you perform using the Google Translate tool. Google Translate is a third-party automated translator and, as such, may make mistakes that a skilled human translator would not otherwise make.³

A perusal of the machine translated version of the City of Ottawa site reveals that this disclaimer is warranted as the automatically translated version of the site contains a number of meaning errors as well as awkward expressions. Although interested in being compatible with and offering parallel service to other city institutions, OPL librarians nonetheless expressed concern about adopting the same model of providing unverified raw machine translation output to their clients.

² The City of Ottawa website can be found at the following location: <http://ottawa.ca> [Retrieved February 22, 2014]

³ The full disclaimer notice on the City of Ottawa website can be viewed at: <http://ottawa.ca/en/city-hall/accountability-and-transparency/disclaimer> [Retrieved February 22, 2014]

A possibility proposed by the researchers was to explore the degree to which post-edited machine translation output could offer a solution that was more cost-effective and efficient than professional translation, but of a higher quality than raw machine translation. Post-editing refers to a situation where a text is first translated automatically by a machine translation system, and then the raw output is corrected and improved by a professional translator (Allen 2003). The level of editing that is applied can vary depending on the needs of the users.

This question of user needs is a very important one. In the context of this project, the intended users are members of the Spanish-speaking immigrant community in the Ottawa region. Therefore, in order to best determine which type of translation – professional translation, raw machine translation, or post-edited machine translation – would best meet the needs of these target users, it is necessary to ask them and to include them in the decision-making process.

Therefore, using a community-based participatory research approach, this project seeks to address the following general research question:

Given a context where there is a limited budget and a time constraint for producing translations for the Ottawa Public Library website, which of the following types of translation – human translation, maximally post-edited machine translation, rapidly post-edited machine translation, or raw machine translation – most reasonably meets the needs of Ottawa's Spanish-speaking immigrant community?

3. OVERVIEW OF PROJECT SCOPE AND COMPLETED PHASES OF RESEARCH

Task	Period of Activity	Completion Date
Literature survey	February-March 2013	March 2013
Community building	February-November 2013	November 2013
Acquisition of approval from University Research Ethics Board	March-September 2013	9 April 2013 (Level 1A approval received for pilot phase) 23 September 2013 (Level 1B approval received for full study)
Discussions with OPL librarians to refine research question and to identify the portions of the website to be translated for the experiment	April 2013	April 2013
Testing and evaluation of three machine translation systems	May 2013	May 2013
Development of draft survey	June 2013	June 2013
Production of translated and post-edited texts	July 2013	July 2013
Design of recipient evaluation	August 2013	August 2013
Pilot testing of survey and recipient evaluation	August 2013	August 2013
Refinement of survey and recipient evaluation based on pilot feedback	September 2013	September 2013
Distribution of survey and recipient evaluation	25 October-27 November 2013	27 November 2013
Analysis of results	December 2013-January 2014	January 2014
Dissemination of results	January 2014 - present	Ongoing

4. METHODOLOGY AND DISCUSSION OF FINDINGS

The project was divided into three main phases. The goal of the first phase was to establish whether the Google Translate machine translation system was a reasonable choice to use as the machine translation engine, or whether another product might produce better results. The second phase aimed to develop a survey that could be distributed to the target community in order to gather basic demographic information, as well as to uncover any reasons why members of this community would like to have content on the OPL website made available to them in Spanish. Finally, the third phase consisted of developing a recipient evaluation in which the target users of the translated material could indicate which type of translation best meets their needs. Each of these phases is described in more detail below.

4.1 Phase I: Evaluating machine translation systems

Evaluation of machine translation can take different forms depending both on the goal of the evaluation and on which stakeholders are involved (White 2003). Many sophisticated metrics have been developed to attempt to automatically evaluate the quality of a machine translation systems output (e.g. Papineni *et al.* 2002; Coughlin 2003). However, we elected to adopt a more modest and goal-oriented approach within the specific context of this project. We wanted to establish whether Google Translate – the machine translation engine adopted by the City of Ottawa – was a reasonable system to use for the upcoming phases of this pilot study at the OPL, or whether a competing product on the market would be a better choice.

Currently, machine translation systems adopt one of three main architectures, which means that the underlying approach used by a system may be a statistical approach, a rule-based approach, or a hybrid of the two. Google Translate is recognized as a leading statistical machine translation system. We also elected to test well-known systems from the other two categories. Reverso was selected as a rule-based machine translation system, and Systran as a hybrid system. All three machine translation tools are freely available online, which is an important consideration for OPL.

The evaluation approach that we adopted was relatively crude, but it was in line with the general goal of the evaluation. In consultation with the OPL librarians, we selected three different extracts of approximately 300 words from the Newcomers Services section of the website. The texts were of three different types: 1) an informative text written in a prose style, 2) an informative text written in a question-and-answer style, and 3) an instructional text. Each of these text types was similar in style to the different sections of the website that the OPL librarians wanted to make available in Spanish. Since it is very well known (e.g. Polvsen *et al.* 1998) that not all types of texts are equally suited for machine translation, and since different machine translation architectures have different strengths, we felt it was important to conduct the evaluation using the different text types that were featured on the website with each of the machine translation systems.

Therefore, each of these three English-language extracts was translated into Spanish by the three different machine translation systems, which were anonymized as systems A, B and C. Note that for each individual text, a different system was identified as “system A”, etc., as illustrated in Table 1. This step was taken in order to better anonymize and randomize the material that would be evaluated.

Text 1: informative prose	Text 2: informative Q&A	Text 3: instructional
System A=Google Translate	System A=Systran	System A=Reverso
System B=Reverso	System B=Google Translate	System B=Systran
System C=Systran	System C=Reverso	System C=Google Translate

Table 1. Table identifying which machine translation systems were used to produce each version.

The resulting nine translations, along with the original English-language texts, were then presented to three different professional translators. The translators were asked to look at the

three versions (A, B, C) of each text, and to comparatively rank them, where 1 was the best translation and 3 was the worst. The translators were advised that all the translations had been produced automatically using machine translation, and that it was possible that none of the translations would be “good”. However, the goal was simply to rank them in comparison to one another. The translators were instructed that the general criteria that were to be used as a basis for the comparison were the criteria of *fidelity* and *intelligibility*. Fidelity is a measure of accuracy that aims to determine how well the contents of the translation reflect the contents of the original text. In other words, has the information been translated correctly with regard to its meaning. Intelligibility is more of a stylistic measure that seeks to determine how understandable or easy-to-read each text the text is, in comparison to the others. In making their assessment, translators were also instructed to consider how much editing would be required to “repair” each text.

4.1.1 Phase I Findings and Discussion

Working independently, and taking into account fidelity, intelligibility and degree of editing required, the translators were asked to providing a comparative ranking for each text. We have un-anonymized the results and presented the rankings for the three machine translation systems in Table 2.

	Google Translate	Reverso	Systran
TEXT 1 (informative prose)			
Translator 1	1	2	3
Translator 2	2	1	3
Translator 3	3	1	2
TEXT 2 (informative Q&A)			
Translator 1	1	3	2
Translator 2	1	3	2
Translator 3	1	3	2
TEXT 3 (instructional)			
Translator 1	1	3	2
Translator 2	1	3	2
Translator 3	1	2	3
MODE	1 (7/9)	3 (5/9)	2 (6/9)

Table 2. Results of the ranking of machine translation output by three professional translators.

Overall, the results show that seven times out of nine, the Google Translate system was deemed to have produced the best results. Meanwhile, the output from Systran was ranked in second place six out of nine times, while Reverso’s output was ranked third in five out of nine cases.

While this crude method of evaluation does not permit a fine-grained analysis of the strengths and weaknesses of the different systems, we would nonetheless tentatively suggest that the systems that have a purely statistical architecture performed very well on texts that are in the form of a set of instructions or a question-answer format. Meanwhile, the prose format

appeared to pose a greater challenge for the statistical systems, and was better handled by the rule-based system.

Additionally, there was a stronger inter-rater reliability for the informative text in the Q&A style (Text 2) and the one in the instructional style (Text 3), while the informative text written in prose format (Text 1) produced a lower inter-rater reliability. This might suggest that the evaluation of texts written in a prose format is more open to subjectivity on the part of evaluators.

At the end of this phase, it was agreed that the Google Translate system would be retained as the machine translation system to be used for the recipient evaluation phase of the project (Phase III).

4.2 Phase II: Surveying the community about their translation needs

The next task was to develop a survey to be distributed among the community of Spanish-speaking immigrants in the Ottawa region. The goal of the survey was to gather basic demographic information, but also identify the reasons *why* members of this community might find it valuable to have the information on the OPL website made available to them in Spanish. While it would be easy to assume that Spanish-speaking immigrants would want access to these texts in Spanish because they do not understand English, this could in fact be an erroneous assumption. As stressed by Edwards (1992: 48), probing the underlying reason, rather than simply asking yes/no type questions, provides valuable insight into a linguistic minority community. In this case, the motivation for wanting a text translated might have a bearing on the type of translation that is most appropriate for meeting these needs.

The initial draft of the survey was prepared by the university researchers, one of whom is also a Spanish-speaking immigrant to Canada. The survey was designed to be both voluntary and anonymous, and the final instrument was approved by Research Ethics Board of the University of Ottawa (see Appendices). Questions that were aimed at establishing the demographic profile of the respondents included questions relating to age; sex; country of origin; length of time in Canada, and also more specifically in the Ottawa region; reason for moving to the region; projected length of time for remaining in the region; level of education; dominant language; ability to understand English and French; whether they currently use the public library and with what frequency; and whether they would be more likely to use the public library, or to use it more often, if the website were available in Spanish. As noted above, there was also a question about why the respondents wanted to have the website available in Spanish.

In recognition of the fact that the targeted respondents were likely to come from a wide variety of educational backgrounds, with varying levels of literacy, every effort was made to phrase the questions in a clear and straightforward way. Moreover, the survey was prepared using a neutral “international” Spanish, rather than in the dialect of any particular region.

Once the survey had been drafted by the researchers, it was pilot tested by four volunteers from the target community, who were identified by the project research assistant. As pointed out by Horowitz *et al.* (2009: 2635), “Who would know better whether the research methods and tools are sensible and engaging and how to structure recruitment so that participants want to take

part than those very targets?" This community input proved valuable for helping to make the survey instrument more user-friendly, applicable and culturally appropriate. Using feedback from the pilot test participants, we made some minor changes to the survey, which included rephrasing and clarifying some of the questions, as well as changing the order in which some of the questions were asked.

In addition, these community members also gave important advice with regard to the recruitment of participants. While there are numerous benefits associated with CBPR, several researchers have identified a number of associated challenges as well, and many of these have to do with recruiting participants or getting the "buy in" of the community. For instance, a number of researchers (e.g. Israel *et al.* 2001: 185; Horowitz *et al.* 2009:2639; Fitzpatrick 2013:14; Willison 2013:11) have observed that developing trust and relationships with community members requires time and patience.

The project research assistant took the lead role in the community-building efforts. As a native speaker of Spanish and an immigrant, he already had a certain standing within the community, but he was still required to invest time in order to build relationships and gain "buy in" for the project. He began cultivating these relationships in the early stages of the project, making appointments to meet personally with key figures and to explain the goals of the project and to provide essential documentation, such as a summary of the project description and a certificate of approval from the University of Ottawa's Research Ethics Board. These key figures often played a "gatekeeper" role, so in order to gain a more direct access to the community, it was necessary to first gain the trust of these individuals. The gatekeeper community members with whom the project research assistant met personally included the ambassadors of the embassies of Colombia, El Salvador and Guatemala; the priests at two local churches that offer Spanish-language mass; and the Spanish-speaking settlement workers in immigrant-serving organizations such as the Catholic Immigration Centre (CIC) and the Ottawa Community Immigrant Services Organization (OCISO).

Once a trust relationship had been built with the gatekeeper figures, the project assistant continued to build relationships with the community members more directly. For example, he attended the services at the Spanish-speaking churches – not just once, but for several weeks – and he was given permission to address the congregation directly in order to explain the project and its potential benefits for the community, to answer any questions they might have, and to gather early stage input. In this way, by the time we were ready to distribute the survey, the community already had a good understanding of the project and its intentions, and some members (i.e., the pilot test participants) had already given input into the development of the survey tool. In addition, there was also a sort of positive snowball effect in that some of these community members went on to recommend additional distribution channels that we had not initially thought of, such as a local online Spanish-language newspaper *Mundo en español* (Ottawa edition), and the Ottawa "Spanish-language meet-up" distribution list.

Although some of the community members were clearly comfortable with information technology and with the idea of responding to an online survey, this was not the case for

all of the potential participants. We therefore faced an additional challenge – one that had also been identified by Willison (2013:11):

...doing CBPR is an important means to help identify the needs of marginalized (vulnerable) communities. [...] However, such marginalized communities face varied systemic barriers to participating in research and/or policy discussions. On-line surveys, town hall meetings and use of traditional focus group techniques – as used to conduct CBPR - simply do not practically work for many people and communities. They may, for example, not have access to technology, language, and other resources to enable them to participate.”

Indeed during the course of community-building, it became evident that some potential respondents would not be able to participate easily in an online survey. Therefore, the decision was made to distribute the survey in both an online version (using Fluid Surveys⁴) and a paper-based format. The survey was conducted over a five-week period during October and November 2013. This dual format delivery was a good decision as 38 of the 114 completed surveys that were received came in hard copy, which represents 33% of the total responses.

4.2.1 Phase II Findings and Discussion

Overall, 114 completed responses were received. Fifty-three percent of the respondents were women and 47 percent were men. Sixty percent of the respondents were in the age range between 31 and 50 years old, while 20 percent were under the age of 30 and the remaining twenty percent were over the age of 50. The respondents were from 17 different countries – including Spain and 16 countries in Central and South America – with larger groups of respondents coming from Colombia⁵ (30 percent), Mexico (18 percent) and Chile (9 percent).

Fifty-seven percent of the respondents had come to Canada to work and 29 percent to study, while the remaining 20 percent did not specify a reason. Thirty-four percent had been residing in Canada for less than two years, 28 percent for two to five years, and 38 percent for more than five years. Seventy-five percent intend to stay in Canada for ten years or more, while 23 percent plan to remain for a period between two and ten years, and two percent for less than two years.

With regard to the highest level of education completed, 29 percent of respondents had a graduate degree, 22 percent had an undergraduate degree, and an additional 25 percent had completed some type of post-secondary vocational training. Meanwhile, 19 percent had completed secondary school, and only 5 percent had ended their formal education before completing secondary school. Therefore, we can say that the Spanish-speaking immigrants who responded to this survey were relatively well-educated on the whole, with 76 percent of them having completed some type of post-secondary education or training.

⁴ For more information on Fluid Surveys, please consult: <http://www.fluidsurveys.com>

⁵ The high response rate among the Colombian community may be partially explained by the fact that the project research assistant is of Colombian origin and had particularly good connections within this community.

In response to questions about degree of comfort in understanding Canada's two official languages, 60 percent of the respondents indicated that their comprehension of English was either good or excellent, 21 percent considered their knowledge to be fair, and 19 percent felt it was poor. With regard to French, 36 percent of the respondents reported a good or excellent level of comprehension, 11 percent fair, and 53 percent poor. Overall, a greater number of the Spanish-speaking immigrants have stronger second language skills in English than in French.

If we consider the question of whether respondents are able to function in at least one of the two official languages, we see that 31 percent claimed a poor ability to understand in either English or French, while the remaining 69 percent are able to function reasonably well in at least one of the two official languages. This is perhaps not too surprising given that 66 percent of the respondents had already been living in the region for more than two years. Of these longer-term residents, 58 percent claimed a good or excellent knowledge of English and 32 percent claimed a good or excellent knowledge of French. Meanwhile, among the more recently arrived newcomers, the number of people claiming a good or excellent knowledge of English fell to 18 percent and those having a good or excellent knowledge of French fell to 24 percent.

In response to questions regarding library use, 68 percent of respondents claimed to be patrons of the public library, while 32 percent are not currently library users. Of those who use the library, 48 percent do so at least once per month, while the remaining 52 percent are less frequent users. Meanwhile, in response to a question that asked whether respondents would be more likely to begin using or to make an increased use of the library if a greater volume of informative material, such as the website, were available in Spanish, 73 percent of respondents said definitely or probably, whereas 20 percent responded possibly, and only seven percent indicated that it was unlikely.

This time, when we compare the responses of longer-term residents with more recently arrived newcomers, we see some surprising results. Of the respondents who had been living in the region for less than two years, 69 percent indicated that they would definitely or probably use the library more if there were Spanish material available. Interestingly, this number was even higher – 76 percent – among the longer term residents. This would seem to indicate that it is not only the recently arrived newcomers who are interested in having material available in Spanish.

Finally, we asked the respondents to indicate the reason(s) *why* they would like to have the website made available in Spanish.⁶ Only 11 percent of respondents indicated that they wished to have the material translated because they were unable to understand the original text.⁷ This is not too surprising if we recognize that, as noted above, only 19 percent of respondents felt that their ability to understand English was poor. Meanwhile, 25 percent of respondents indicated that they were reasonably sure that they understood the English text, but they would appreciate having access to a Spanish translation in order to be able to verify their understanding. In other words, the translation would be used for confirmation purposes, as a sort of confidence booster. Forty-four percent of respondents had a pragmatic reason, reporting

⁶ Because respondents were permitted to give more than one reason, the percentages do not add up to 100.

⁷ Note that this experiment was conducted exclusively between English and Spanish. It is possible that some of the respondents who indicated that they could not understand the English source text may have been able to understand a French source text.

that while they were fully able to comprehend the source text, they would be able to read more quickly and easily in Spanish. In this case, it is a question of efficiency, rather than one of ability.

A relatively small percentage of respondents – 11 percent – replied that they enjoy having the possibility of working with Spanish texts because it is one means of preserving or reinforcing their cultural heritage. Meanwhile, another 20 percent of respondents indicated that they would use the availability of Spanish-language material as an opportunity to teach this language to others, such as their children. Twenty-five percent of respondents would make use of the translated version as an aid for their own English-language learning efforts. In other words, they would consult the English and Spanish versions in a contrastive way to help them learn more about English terminology and usage.

In addition to choosing from among the list of possibilities suggested, respondents were also invited to suggest other reasons that they may have for wanting a Spanish translation of the text. One additional reason was put forward by a respondent, who noted that having access to a translated version could ease some of the pressure of integration. The respondent commented that, while largely valuable, the immersion process can be very intense and overwhelming and that even occasional opportunities for lightening the cognitive load were welcome.

Two other respondents added a comment to indicate that their needs had changed over time. They commented that while they having been living in the region for some time and are now quite comfortable functioning in English, they would certainly have appreciated having access to this type of service at the time when they had recently arrived.

Meanwhile, one additional respondent warned of potential dangers associated with the provision of material in immigrant languages, suggesting that it could contribute to a sort of linguistic ghettoization. This respondent recommended providing support to learn Canada's official languages, rather than providing translation in Spanish. However, as described above, 25 percent of the respondents indicated that a motivation for wanting to have text available in Spanish was to use this as a tool for helping to improve English-language skills, by allowing them to learn new vocabulary, for example. Similarly, another 25 percent of respondents wanted to use the Spanish version as a tool for confirming their comprehension of English. When viewed in this light, rather than promoting linguistic ghettoization, the service of providing translated texts can help immigrants to build knowledge and confidence in English so that they can integrate more fully into their new country.

4.3 Phase III: Conducting a recipient evaluation of translated texts

The final phase of the project was to solicit the opinion of the community members regarding the usefulness of different forms of translated texts. While machine translation is unquestionably faster and cheaper than professional human translation, in the vast majority of cases, so-called "raw" or unedited machine translation output is of a lesser quality than human translation. A key question that remains to be answered is whether members of the Spanish-speaking newcomer community will find some form of machine translation output to be useful for helping to meet their translation needs. In order for the use of machine translation to be considered a viable option for OPL, the intended recipients of the translated texts must be willing to accept the output produced by machine translation systems.

As noted above, there are numerous possible approaches for evaluating the quality of machine translation systems. However, in this CBPR project, we elected to adopt the view put forward by Chesterman and Wagner (2002: 80-84), who note that one approach is to view translation as a service, intangible but wholly dependent on customer satisfaction; therefore, to measure translation quality, one needs to measure customer satisfaction. This approach to evaluating machine translation is referred to as a *recipient evaluation*, which is described by Trujillo (1999: 255) as an evaluation performed by the recipients of the translation in order to evaluate quality, cost, and speed. As pointed out by Loffler-Laurian (1996: 69), for example, it is the intended recipients of the translated texts who are in the best position to decide whether or not their needs can be suitably met by those texts. Therefore, to find out if members of the Spanish-speaking newcomer community would be accepting of machine translation, we conducted a recipient evaluation in order to determine the extent to which some form of machine translation can help to meet some of their translation needs.

Experience has clearly shown that machine translation is not a viable option for all types of texts or situations (e.g. Church and Hovy 1993; L'Homme 2008: 273). It is also generally accepted that machine translation is better viewed as a translation aid, rather than as an outright replacement for human translators. With this in mind, we sought to experiment with different text types, and with both raw and edited machine translation output.

With regard to text types, as noted above, the OPL librarians had identified a section of the OPL website under the "Newcomer Services" tab for which they thought it would be useful to provide information in Spanish. From this section of the website, we selected three different text extracts of between 300 and 400 words in length. These texts were not the same extracts that were used in the machine translation evaluation phase described above; however, they did belong to the same general text types or genres. The following three texts were used for the recipient evaluation:

- Text 1: "Your Library Card" (380 words) – informative prose format
- Text 2: "FAQs on borrowing materials" (368 words) – informative Q & A format
- Text 3: "How to place a Hold on a Book, DVD or Music CD" (301 words) – instruction format

As for editing, Allen (2003: 302) describes two levels of post-editing that can be applied to raw machine translation output. The first is known as Rapid Post-Editing (RPE), which he describes as follows:

...strictly minimal editing on texts in order to remove blatant and significant errors and therefore stylistic issues should not be considered. The objective is to provide the minimum amount of necessary correction work that can be made on a text in order for the text to be fully understandable as a comprehensible element of information.

In other words, RPE is concerned only with ensuring the accuracy of the content, and it does not attempt to address readability or stylistic issues. The text may sound like it was generated by a computer, the syntax might be somewhat unusual, the grammar may not be perfect, but the message is accurate.

In contrast, Maximal Post-Editing (MPE) addresses both content and style. The goal of MPE is to edit the machine translation output to such a level that the resulting text closely resembles a professional human translation. In other words, once a text has been corrected using MPE techniques, it should be both accurate and well-written so that it is virtually impossible to distinguish it from a professional translation. The Translation Automation User Society (TAUS) has developed two sets of guidelines to help editors apply RPE and MPE to machine translation output, and these guidelines were provided to the translators who participated in this project.

4.3.1 Preparatory work

For each of the three source texts, four different translated versions were produced:

- 1) A professional human translation (HT);
- 2) A maximally post-edited (MPE) machine translation (i.e., a translation where both content and style were corrected to produce a text that resembles as closely as possible a human translation);
- 3) A rapidly post-edited (RPE) machine translation (i.e., a translation where content errors were corrected, but stylistic changes were not made);
- 4) A raw or unedited machine translation (MT).

The time and cost for producing each of the four versions was also calculated because whenever a text needs to be translated in a context where there are time and budget constraints – as is the case at OPL – the parameter of quality must be weighed against the parameters of time and cost. The methods used for determining the time and cost are described in the following paragraphs, while the actual time and cost required to produce each version are summarized further down (see Tables 4 and 5).

The raw machine translations were produced by running the texts through the Google Translate machine translation system. The translation is produced almost instantaneously, and there is no cost involved. Therefore, the time and cost for producing raw machine translation (MT) are considered to be zero in the context of this experiment.

Three professional translators were hired to help produce the remaining target texts, which included the RPE text, the MPE text and the human translation (HT). Clearly it is not possible for three different translators to have precisely the same ability; however, every effort was made to find three translators who had a comparable background and level of skill as described above. Rather than assigning each translator a single method (i.e., one of RPE, MPE or HT) to be carried out on three texts, we elected to have all three translators apply all three methods. However, as summarized in Table 3, each translator used a different method on each of the texts. Moreover,

in an effort to control for potential order effect, they also carried out the tasks in a different order.

	Translator 1	Translator 2	Translator 3
Text 1	HT	RPE	MPE
Text 2	RPE	MPE	HT
Text 3	MPE	HT	RPE

Table 3. Task assignments and order for the translators.

By having each translator apply each of the three methods – RPE, MPE and HT – we were able to calculate an average time per 100 words for each of the methods. In our opinion, calculating an average time gives more representative results than would be achieved by having each translator apply only a single method. The average time required to apply each of the methods used in this experiment is summarized in Table 4.

Method	Average time required to produce 100 words	Savings as compared to the time required to produce HT
HT	13.3 minutes	--
MPE	9.3 minutes	4 minutes (30%)
RPE	4.5 minutes	8.8 minutes (66%)
MT	0 minutes	13.3 minutes (100%)

Table 4. Comparison of the average time to produce 100 words of text using the methods of HT, MPE, RPE and MT, as well as the savings associated with each method.

To determine the cost associated with producing each of the three versions, we referred to the “2012 Survey on Rates and Income” produced by the *Ordre des traducteurs, terminologues et interprètes agréés du Québec*, which is a large and well-respected professional translators’ association in Canada. The survey is conducted every two years, and in 2012, there were 336 respondents.

According to this survey, the average cost of translation is \$56.73 per hour, while the average hourly rate for editing is slightly cheaper at \$54.85 (Gauthier 2012: 3). Using these figures, we were able to calculate the cost of producing the different versions, which is summarized in Table 5.

Method	Average cost for producing 3 texts (1049 words) for OPL website	Savings as compared to the cost of HT
HT	\$131.42	--
MPE	\$89.59	\$41.83 (32%)
RPE	\$43.88	\$87.54 (67%)
MT	\$0	\$131.42 (100%)

Table 5. Comparison of the average cost of producing texts using the methods of HT, MPE, RPE and MT, as well as the savings associated with each method.

Not surprisingly, the data show that raw MT was always the fastest and cheapest method of producing a text, followed by RPE, then MPE, and finally HT. For the texts used in this

experiment, it is interesting to note that those produced using MPE – which aims to produce texts that are comparable in quality to HT – were on average 32 percent cheaper and required 30 percent less time to produce than professional human translations. In a context where an organization, such as OPL, seeks to provide translations in multiple languages, this could potentially add up to significant total savings of time and money. Of course, any such savings would only be meaningful if the texts that were produced were deemed to be useful for and usable by the target community for which they are intended. Therefore, the next important step was to present these different translated versions to the Spanish-speaking newcomer community to determine which, if any, of these texts can best meet their needs.

4.3.2 Recipient evaluation

The recipient evaluation had several steps. First, participants were presented with a choice of the following three texts and asked to select one:

- Text 1: “Your Library Card”
- Text 2: “FAQs on borrowing materials”
- Text 3: “How to place a Hold on a Book, DVD or Music CD”

Note that in the online version, the randomize feature was employed so that the order in which the three source texts were presented varied from participant to participant. For the paper-based version, three different versions were prepared – one for each source text – and these were “shuffled” and distributed in a random fashion.

Once they had selected a text, the participants were shown four different Spanish-language translations of that text, produced using the four methods described above (HT, MPE, RPE and MT). These four versions were also presented in a random order, and there was no indication as to which method had been used to produce each version. Participants were simply asked to read each of the four versions and to indicate which version they found to be the one that best met their needs.

In a subsequent step, participants were once again shown the same four translated versions of the text, but this time, they were also given information about the method used to produce each version, as well as the time and cost required to produce them. Balancing the parameters of time, cost and quality, the participants were asked to select the text that best met their needs.

4.3.3 Phase III Findings and Discussion

The results of the recipient evaluation are summarized in Table 6 below. The results are first reported individually for each of the three texts, and then averaged. The method used to produce each version is identified in the left-hand column. The second column shows the percentage of participants who selected each version when the method and the production time and cost were unknown. The third column shows the percentage of participants who selected each version during the subsequent step, when the method and the production time

and cost were revealed and taken into account in the decision-making process. Finally, the right-hand column shows the difference in percentage of people selecting each method once the information about the method, time and cost had been provided.

	Method/Time/Cost Unknown	Method/Time/Cost Revealed	Difference
Text 1: "Your Library Card" (informative prose) (26 respondents)			
HT	27%	0%	-27%
MPE	19%	19%	No change
RPE	50%	62%	+12%
MT	4%	19%	+15%
Text 2: "FAQs on borrowing materials" (Q&A style) (49 respondents)			
HT	57%	12%	-45%
MPE	27%	18%	-9%
RPE	14%	57%	+43%
MT	2%	12%	+10%
Text 3: "How to place a Hold on a Book, DVD or Music CD" (instructional) (39 respondents)			
HT	33%	3%	-30%
MPE	23%	13%	-10%
RPE	23%	46%	+23%
MT	21%	38%	+17%
Average for all texts (114 respondents)			
HT	42%	6%	-36%
MPE	24%	17%	-7%
RPE	25%	54%	+29%
MT	9%	23%	+14%

Table 6. Respondents' preferences for translation methods before and after the information about method and production time and cost had been revealed.

In all cases, the findings reveal that the raw machine translation was selected by the fewest number of participants when the method, time and cost were unknown. This number rose substantially – by an average of 14 percent – once this information was revealed.

In contrast, the number of participants who selected either human translation or maximally post-edited translation (which strives to imitate human quality) in the first instance was very high; on average, these two categories combined were preferred by 66 percent of the respondents. However, once the details regarding the method, time and cost were made clear, the number fell to just 23 percent.

In the final analysis, once the details of method, time and cost were revealed, the method that was preferred by the greatest number of participants – in the case of each individual text and the average of all three – was rapidly post-edited machine translation. This form of text, in which the content is accurate but the style may not be elegant, was deemed by more than half the respondents to provide the best value for money given the translation needs at hand.

If we recall the discussion above, it was observed that 44 percent of the respondents wished to have the translated text available because they could process it more quickly, while an additional 25 percent were looking to the translated version merely to confirm that they had understood the English text correctly. Another 11 percent of respondents wished to have a translation because they had very limited proficiency in English and could not understand the source text. For this range of needs, an elegant text is not required; a translation that provides the correct gist of the text is sufficient.

Of course, some respondents did indicate other types of reasons for wishing to have the text translated. For instance, 11 percent indicated that they would use a Spanish translation as a means of preserving their culture, and another 20 percent of respondents indicated that they would use the availability of Spanish-language material as an opportunity to teach this language to others, such as their children. For this type of need, a higher quality of text is certainly desirable. This may explain why a relatively small number of participants – an average of 23 percent – still indicated a preference for having either a human translation or a maximally post-edited translation, even once the time and cost had been revealed. It is interesting to note, however, that of this 23 percent, 17 percent were willing to accept the more economical option of MPE, while a mere 6 percent insisted on HT.

When considering text type, it is worth noting that the text that is written in an instructional style is the one for which the greatest number of respondents indicated a willingness to accept a text that had been rapidly post-edited (46 percent) or even produced as raw machine translation (38 percent). This is not entirely surprising and is in keeping with the widely held observations that have been made about texts that are generally considered suitable for machine translation (e.g. Polvsen *et al.* 1998). The very purpose of an instructional text is to make it clear to the reader what steps should be followed in order to achieve a goal. As such, these texts tend to contain short sentences with simple structures and precise terminology. The relative lack of syntactic and semantic ambiguity in this type of text makes it a promising candidate for machine translation. This general observation would seem to be supported by the findings of this study also.

5. OUTCOMES AND OUTPUTS

To date, the results of the research have been presented at the following two conferences:

- “Can Machine Translation Facilitate Outreach to Newcomers? A Pilot Study Investigating the Needs of Spanish-Speaking Users of the Ottawa Public Library” presented at session 7.1 of the [2014 Association for Library and Information Science Education \(ALISE\) Annual Conference](#), Philadelphia, PA, January 21-24, 2014.
- “Machine Translation and Minority Language Communities: Compatible or in Conflict?” presented at the [Symposium on Translation and Technology: Challenges in a Multilingual Society](#), Antwerp, Belgium, February 18, 2014

In addition, we have prepared a short summary (circa 700 words) that we have submitted for

possible publication in *Access OLA*, the professional magazine of the Ontario Library Association:

- “Can machine translation help newcomers to access the public library?” article submitted to [Access OLA](#) (submission sent on February 20, 2014).

We also plan to give a public presentation of the project at an upcoming “Research Conversation”. This is a monthly “brown bag lunch” series organized at the University of Ottawa’s School of Information Studies. These talks are open to the public and the announcements and invitations are widely distributed. The roster for the 2013/2014 academic year is already full, but we will present during the 2014/2015 series.

Additionally, we are preparing an article for submission to a peer-reviewed scientific journal. Although we have not yet finalized the article nor decided where to send it, possible target journals include the [Journal of the Association for Information Science and Technology](#) or the [Canadian Journal of Information and Library Science](#). A copy of any paper accepted for publication will be deposited with OCLC as requested.

A popularized, one-page summary of the findings written in Spanish is currently being prepared for distribution to the community members who participated in the survey and recipient evaluation. It will be distributed using the same channels that were used to distribute the initial invitation to participate in the project.

A copy of the present report, as well as a copy of any paper eventually accepted for publication, will also be provided to the Ottawa Public Library.

6. FUTURE RESEARCH

The research conducted as part of this project was a pilot study with one library branch (Ottawa Public Library) and one newcomer community (Spanish-speaking immigrants). This project has the potential to be expanded in a number of directions.

One way to expand the project would be to extend the languages involved. The 2011 Canadian census reports over 120 immigrant languages that are used as the “language spoken most often in the home”. There is therefore considerable scope to test this approach with other immigrant communities. Machine translation may have greater or lesser appeal to other groups, and indeed, the work of post-editing machine translation output may be more arduous in other languages. In the longer term, it would be important to do testing with other language families before deciding to adopt this approach on a large scale. In the short term, however, we plan to do a related project using French because a small grant has already been awarded by the *Centre canadien de recherche sur les francophonies en milieu minoritaire*, who are interested in supporting the integration of French-speaking immigrants.

Another important area for development could be a deeper analysis of the types of texts that are most suited for translation with the help of machine translation systems. The pilot study revealed that some types of texts resulted in better “raw” output than others, and therefore required less post-editing. Identifying those types of texts that can most easily be translated

with machine assistance will help libraries know where to focus their limited resources for maximum effect.

Finally, while this study focused on post-editing the output of a machine translation system, another interesting avenue of exploration could be to investigate “pre-editing”. In other words, to what extent is it possible to write/revise the original English-language text in such a way that it is easier for the machine translation system to process? If the original text could be improved by eliminating ambiguities or difficult syntactic constructions, then the resulting machine translation output should be better and require less post-editing. This becomes increasingly important if the original text is to be translated into multiple target languages. If a difficult passage is encountered in the original text, it will likely be mistranslated in all the target languages, and therefore it will need to be corrected in all of those languages in a post-editing phase. This could mean that 10, 15 or even more than 20 post-editors need to make those corrections in all those different languages. In contrast, if those problems could be eliminated just a single time in the original text by making it more “translation-friendly”, then the total time and cost required for translation and post-editing would be reduced significantly.

REFERENCES

- Allen, Jeffrey. (2003). "Post-editing," in Harold Somers Computers and Translation: A Translator's Guide. Amsterdam/Philadelphia: John Benjamins, 297-317.
- Caidi, N. and Allard, D. (2005). "Social inclusion of newcomers to Canada: An information problem?" *Library & Information Science Research*, 27(3), 302-324.
- Canadian Broadcasting Corporation (CBC) (2011, Nov. 30) "Toronto Library Chair Defends Multilingual Collection." <http://www.cbc.ca/news/canada/toronto/story/2011/11/30/toronto-libraries.html> [Retrieved February 22, 2014]
- Chesterman, Andrew & Emma Wagner. (2002). *Can Theory Help Translators? A Dialogue Between the Ivory Tower and the Wordface*, Manchester: St. Jerome Publishing.
- Church, Kenneth W. & Eduard H. Hovy. (1993). "Good Applications for Crummy Machine Translation". *Machine Translation* 8, 239-258.
- Coughlin, D. (2003). "Correlating Automated and Human Assessments of Machine Translation Quality" in *MT Summit IX*, New Orleans, USA pp. 23–27
- Edwards, J. (1992). "Sociopolitical aspects of language maintenance and loss: Towards a typology of minority language situations," in W. Fase, K. Jaspaert and S. Kroon (eds) *Maintenance and Loss of Minority Languages*, Amsterdam/Philadelphia: John Benjamins, 37-54.
- Fitzpatrick, Joey. (2013). "Nothing about us without us." *University Affairs*. 6/13: 12-17. <http://www.universityaffairs.ca/nothing-about-us-without-us.aspx>
<http://www.universityaffairs-digital.com/universityaffairs/201306#pg16> [Retrieved February 22, 2014]
- Gauthier, François (2012) "2012 Survey on Rates and Income," Quebec, Canada: Ordre des traducteurs, terminologues et interprètes agréés du Québec. [Translated by Karin Montin]
- George, U. (2002) "A needs-based model for settlement service delivery for newcomers to Canada," *International Social Work* 45 (4): 465-480.
- Holt, G. (ed) (2009) "Library Services for Immigrants, an Abridged Version," *Public Library Quarterly* 28(2): 120-126.
- Horowitz, Carol R., Robinson, Mimsie and Seifer, Sarena. (2009). "Community-Based Participatory Research From the Margin to the Mainstream: Are Researchers Prepared?" *Circulation* 119: 2633-2642. <http://circ.ahajournals.org/content/119/19/2633.full> [Retrieved February 21, 2014]

Bowker, Lynne: Can machine translation facilitate outreach to newcomers? A pilot...

Israel, Barbara, Schulz, Amy, Parker, Edith, and Becker, Adam. (1998). "Review of Community-Based Research: Assessing Partnership Approaches to Improve Public Health." *Annual Review of Public Health* 19: 173-202.

International Federation of Library Associations and Institutions (IFLA). 2013. Riding the Waves or Caught in the Tide? Navigating the Evolving Information Environment. Insights from the IFLA Trend Report. The Hague: IFLA. http://trends.ifla.org/files/trends/assets/insights-from-the-ifla-trend-report_v3.pdf [Retrieved March 2, 2014]

Israel, Barbara, Schulz, Amy, Parker, Edith, and Becker, Adam (2001). "Community-based Participatory Research: Policy Recommendations for Promoting a Partnership Approach in Health Research" *Education for Health* 14 (2): 182-197. https://depts.washington.edu/ccph/pdf_files/EducforHealthIsrael.pdf [Retrieved February 20, 2014]

Jang, W., Basi, R., Vanderkooy, M. and Sullivan, P. (2004) "Cross-Canada Checkup: Multicultural Services in Public Libraries," *Feliciter* 50 (1): 26.

L'Homme, Marie-Claude (2008). *Initiation à la traductique*, 2e édition. Brossard, Quebec: Linguattech.

Library Settlement Partnerships (LSP) program: <http://www.lsp-peb.ca/> [Retrieved February 22, 2014]

Loffler-Laurian, Anne-Marie (1996). *La traduction automatique*. Villeneuve d'Ascq: Presses Universitaires du Septentrion.

MacDonald, S. (2012) "Conceptualizing Public Library Service to Newcomers in Canada," A. Quan-Haase, V.L. Rubin & D. Chaves (eds). *Proceedings of the 40th Annual Conference of the Canadian Association for Information Science*. http://www.cais-acsi.ca/proceedings/2012/caisacsi2012_60_macdonald.pdf [Retrieved March 2, 2014]

Minkler, Meredith and Wallerstein, Nina (eds). (2008). *Community-Based Participatory Research for Health: From Process to Outcomes* (2nd ed.) San Francisco: Jossey-Bass.

Papineni, K., Roukos, S., Ward, T., and Zhu, W. J. (2002). "BLEU: a method for automatic evaluation of machine translation" in *ACL-2002: 40th Annual meeting of the Association for Computational Linguistics* pp. 311–318.

Polvsen, Claus, Nancy Underwood, Bradley Music, Anne Neville (1998). "Evaluating Text-type Suitability for Machine Translation a Case Study on an English-Danish MT System" In Rubio, A., N. Gallardo, R. Castro & A. Tejada (eds.) *Proceedings of the First International Conference on Language Resources and Evaluation*, Granada, Spain 28-30 May 1998, Vol. 1, 27-31.

Bowker, Lynne: Can machine translation facilitate outreach to newcomers? A pilot...

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.131.2836&rep=rep1&type=pdf>

[Retrieved March 2, 2014]

Statistics Canada. (2011) 2011 Census of Canada. <http://www12.statcan.gc.ca/census-recensement/index-eng.cfm> [Retrieved February 20, 2014]

Translation Automation User Society (TAUS). Post-editing guidelines:

<https://evaluation.taus.net/resources/guidelines/post-editing/machine-translation-post-editing-guidelines> [Retrieved March 2, 2014]

Trujillo, Arturo. (1999). Translation Engines: Techniques for Machine Translation. London: Springer.

White, John S. (2003). "How to Evaluate Machine Translation". Harold Somers (ed). Computers and Translation: A Translator's Guide. Amsterdam/Philadelphia: John Benjamins, 211-244.

Willison, Kevin D. (2013). "Maximizing Chronic Disease Prevention and Management through Community Based Participatory Research and Inter-collaborative Practices." Open Journal of Social Science Research 1(1): 7-14.

http://manuscript.sciknow.org/uploads/ojsr/pub/ojsr_1362045231_2013_1_2_web.pdf

[Retrieved February 20, 2014]