Translation and Interpreting Studies 10:2 (2015), 165–186. doi 10.1075/tis.10.2.01bow
issn 1932–2798 / e-issn 1876–2700 © John Benjamins Publishing Company

Title: Investigating the usefulness of machine translation for newcomers at the public library

Running head: Machine translation for newcomers at the public library

Authors:

Lynne Bowker School of Translation and Interpretation University of Ottawa 70 Laurier Ave East (ART-401) Ottawa, ON, K1N 6N5 Canada Email: lbowker@uottawa.ca	Jairo Buitrago Ciro School of Information Studies University of Ottawa 55 Laurier Ave East (DMS-11101) Ottawa, ON, K1N 6N5 Canada Email: jbuit008@uottawa.ca

Bio-sketches:

Lynne Bowker holds a BA and an MA in Translation (University of Ottawa, Canada) and a PhD in Language Engineering (UMIST, UK). She is a Full Professor at the School of Translation and Interpretation at the University of Ottawa, with a cross-appointment to the School of Information Studies. Her main areas of teaching and research include translation technologies, language for special purposes, knowledge organization and library and information science.

Jairo Buitrago Ciro holds a BA in Civil Theology (Saint Paul University, Canada). He is currently completing a Master of Information Studies at the University of Ottawa (Canada), which includes an 8-month work placement at the library of the Université du Québec en Outaouais.

Abstract:

This study investigates the potential of machine translation for offering an efficient and cost-effective means of translating sections of the Ottawa Public Library website into Spanish to better meet the linguistic needs of the Spanish-speaking newcomer community. One-hundred and fourteen community members participated in a recipient evaluation, where they evaluated four different versions of a translated portion of the library's website—a professional human translation, a maximally post-edited machine

translation, a rapidly post-edited machine translation, and a raw machine translation. Participants also considered metadata such as the time and cost required to produce each version. Findings show that while machine translation cannot address all needs, there are some needs for which the faster and cheaper post-edited versions are considered to be useful and acceptable to the community.

Keywords: machine translation, post-editing, recipient evaluation, immigrant languages, public library

Investigating the Usefulness of Machine Translation for Newcomers at the Public Library

1. Introduction

According to the 2011 Census, Canada's population has increased by 38% since 1981 (Statistics Canada 2011). During this period, Canada welcomed about 235,000 new immigrants each year. Canada has two official languages—English and French—but the mother tongue of over 80% of newcomers is neither English nor French. In 2011, 19.8% of Canada's total population reported having an immigrant language as mother tongue—with over 150 different immigrant languages being reported—and this group is growing at a faster rate than the country's English- and French-speaking populations.

The picture is one of a country with a diverse and evolving linguistic profile. This diversity is also reflected at the municipal level since 86% of immigrants live in one of Canada's six largest cities, including the national capital region of Ottawa. To take Spanish-language immigrants as an example, in 2011 there were 10,930 Spanish-speakers in Ottawa, representing 1.3% of the city's total population, up from 9,860 (1.2%) in 2006.

An obvious question prompted by such a profile is "What type of linguistic support does the newcomer population need?" However, given the diversity of this group of newcomers, the answers will likely differ depending on the geographic location, the language community and the social context in question. This investigation therefore focuses on the potential for machine translation to help Spanish-speaking newcomers in Ottawa, and the social context is limited to the Ottawa Public Library. The research question that will be addressed is as follows:

Given a context where there is a limited budget and a time constraint for producing

translations for the Ottawa Public Library website, which type of translation—human translation, maximally post-edited machine translation (MT), rapidly post-edited MT, or raw MT—meets the needs of Ottawa's Spanish-speaking immigrants?

2. The public library and the newcomer community

The public library was chosen as the social context because many newcomers turn to libraries upon arrival in Canada. At a public library, community outreach aims to develop library programs for non-users, the under-served, and people with special needs. Communities are constantly evolving, so libraries must monitor changes and adapt their services to meet diverse community needs, including linguistic diversity.

As MacDonald (2012, 2) reports, Canada's recent response to challenges associated with immigrant integration has been to turn to municipalities to administer settlement policies and programs. There is a growing interest in the public library's role in immigrant settlement and integration (Frisken and Wallace 2002), and libraries have invested much time and effort into providing newcomers with various forms of information and resources, including efforts to address their language-related needs. For example, the Ottawa Public Library (OPL) is a municipal library system with 33 branches. OPL offers various resources and programs for newcomers, and its website offers information about how to obtain and use a library card, participate in free library programs, access material in a variety of languages or find resources and programs to help improve English or French language skills. In addition, OPL is a partner in the Library Settlement Partnerships (LSP) program, which helps newcomers settle into the community by linking them to services delivered in partnership with a dozen local settlement agencies and city partners. OPL's Newcomer Services attest to an active

Spanish-speaking community. In 2012, a total of 17,454 Spanish-language items circulated. In 2013, the number rose to 19,282, a 10.5% increase.

However, public libraries have struggled to implement multiculturalism beyond collection development (MacDonald 2012, 2). While multilingual collections meet some needs of newcomers, there is growing recognition (e.g., Caidi and Allard 2005; George 2002) that newcomers have other types of information needs, including the need to access information about library and other services in their own language (Holt 2009, 123).

As MacDonald (2012, 2) observes, 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, responsibility for integration is shifting to the newcomers themselves, where the problem is framed as a lack of awareness about settlement services. Citizenship and Immigration Canada (CIC) has recognized that many newcomers turn to public libraries and, in response, is attempting to increase visibility of services in these venues. In 2008, CIC 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. As noted above, OPL is a participant in the LSP program, which helps newcomers by linking them to community and settlement services through partner-led information sessions. These sessions may address issues such as job-seeking, housing support, legal aid, local language training, support for students, cross-cultural education, and community activities and events; however, the it remains the responsibility of the library to provide information about library-specific services. This is challenging given

the linguistic diversity of the immigrant population. Understandably, the two OPL Newcomer Services librarians cannot address this broad linguistic spectrum directly.

One potential strategy for making a library more accessible to newcomers is to translate its website into languages spoken by local immigrants (e.g., Holt 2009, 123). As part of an effort to facilitate access to information for newcomers, OPL has translated some material describing its services into nine languages: Arabic, simplified and traditional Chinese, Hindi, Persian, Russian, Somali, Spanish and Urdu. These translations have been produced by professional translators and are available on a page of the OPL website entitled "Newcomers." Though laudable, it is important to put this effort into context. The volume of text available in these nine languages is less than 1000 words, while the amount available in Canada's two official languages totals hundreds of thousands of words. Moreover, the OPL librarians note that web content is updated regularly, so frequent updates must also be made in each translated version.

Unsurprisingly, cost and time are principal reasons why a greater volume of information is not translated by public institutions. For example, a Canadian professional translators' association recently conducted a survey of rates and salaries. The resulting report indicates that based on 336 respondents, the average rate charged is 22 cents per word (Gauthier 2012, 3). Therefore, translating 1000 words into one language would cost $220, while translating that text into ten languages would cost about $2,200. The cost of providing information in immigrant languages quickly adds up. Moreover, translation is not a one-time cost but is an ongoing commitment because websites are updated regularly.

Like all public institutions, libraries are under pressure to reduce expenses, and

multilingual services are an easy target. Therefore, while translation plays a key role in newcomers' integration, it is costly and time-consuming, thus limiting the range of foreign-language information services that libraries can provide. For immigration-related policies to succeed, their associated costs must be kept down. However, if cost-effective means of translation were available, public libraries might feasibly offer additional newcomer services in various languages.

Indeed OPL wishes to explore possibilities for making a greater portion of its website available in a range of languages. In fact, OPL had already begun to consider whether to integrate a Google Translate widget into the OPL website as a translation solution, in keeping with the strategy already adopted by the City of Ottawa for automatically translating its own website. The City's website also includes the following disclaimer: "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." (City of Ottawa)

A look at the machine translated version of the City's site in Spanish shows that the disclaimer is warranted as this version contains errors. Although interested in offering parallel service to other city institutions, OPL librarians were concerned about providing unverified machine translation (MT) output.

The researchers suggested exploring whether post-editing the MT output could be more cost-effective and efficient than professional translation, and could produce higher quality than raw MT. In this scenario, a text is first translated automatically by an MT system, and then the output is corrected by a professional translator. The degree of post-

editing varies depending on user needs.

This question of user needs is important. For this project, the intended users are Spanish-speaking immigrants in the Ottawa region. To best determine which type of translation would best meet their needs, it is necessary to include them in the decision-making process. Therefore, the framework adopted for this project is community-based participatory research (CBPR), a collaborative approach that seeks to transform research from a relationship where researchers carry out investigations *on* a community to one where researchers work *with* a community to increase understanding of a given phenomenon and to integrate the knowledge gained to improve the quality of life of community members. CBPR is typically applied within the context of public health research (e.g., Horowitz et al., 2009; Minkler and Wallerstein 2008; Willison 2013), but it can be adapted to other fields.

3. Methodology and discussion

The project was divided into three main phases: 1) MT system evaluation, 2) target community survey, and 3) recipient evaluation.

3.1 Phase I: Evaluating machine translation systems

Evaluation of MT can take different forms depending on the goal and on the stakeholders involved (White 2003). Sophisticated metrics have been developed to attempt to automatically evaluate the quality of MT output (e.g., Papineni et al. 2002), but this phase of the project had a goal of establishing which of the three test systems best met the requirements of this pilot study and should be retained for the next phase.

Currently, MT systems adopt one of three main architectures: statistical, rule-based

or a hybrid of the two. Google Translate is a leading statistical MT system. We also tested well-known systems from the other two categories: Reverso (rule-based) and Systran (hybrid). All three tools are freely available online, which is an important consideration for OPL.

Our evaluation approach was relatively rudimentary, but it was in line with the general goal of the evaluation. In consultation with OPL librarians, we selected three different 300-word extracts 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 (Q&A), and 3) an instructional text. Each of these 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 well known (e.g., Polvsen et al. 1998) that not all text types are equally suited to MT, and that different MT architectures have different strengths (e.g. Thurmair 2004; Costa-Jussà et al. 2012) , we felt it was important to conduct the evaluation using the different text types that were featured on the website with each of the MT systems. Therefore, each of the English-language extracts was translated into Spanish by the three different MT systems, which were anonymized as systems A, B and C. To better anonymize and randomize the material to be evaluated, a different system was identified as "system A" for each text, as illustrated in Table 1.

[Insert Table 1 near here]

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 presented to three different professional translators in order to determine whether the Google MT system could be retained for the next phase of the study. The decision to consult professional translators, rather than randomly selected members of the participant pool, was based on the fact that it is necessary to understand both the source and the target language in order to determine the accuracy of a translation. All three professional translators are native Spanish speakers, originating from Cuba, Costa Rica and Colombia respectively. All received graduate-level training in translation at the University of Ottawa and all have at least three years of translation experience. The translators were advised that the three versions (A, B, C) of each text had been produced automatically using MT, and that it was possible that none of the translations would be "good." The translators were instructed that the goal was simply to rank the three versions in comparison to one another using 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 a stylistic measure that seeks to determine how readable each 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 target text.

3.1.1 Phase I findings and discussion

Working independently, and taking into account fidelity, intelligibility and degree of editing required, the translators provided a comparative ranking for each text. We have

un-anonymized the results (i.e., system A, B and C) and presented the rankings for the three MT systems in Table 2.

[Insert Table 2 near here]

	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, Google Translate was deemed to have produced the best results. Meanwhile, Systran's output 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 rudimentary evaluation method does not permit a fine-grained analysis of the different MT systems' strengths and weaknesses, we nonetheless tentatively suggest that the systems that have a statistical architecture consistently outperformed rule-based and hybrid systems when translating texts that are in an instructional or a Q&A format. Meanwhile, the prose format seemed to pose a greater challenge for the statistical system,

and was better handled by the rule-based system. Additionally, there was a stronger inter-rater reliability for the text in the Q&A style (Text 2) and the one in the instructional style (Text 3), while the informative text 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 Google Translate would be retained as the MT system to be used for the recipient evaluation in Phase III.

3.2 Phase II: Surveying the community about their translation needs

The next task was to develop a survey to distribute to Spanish-speaking immigrants in Ottawa. Horowitz et al. (2009, 2634) recommend that in CBPR, "community" should be interpreted broadly as all who will be affected by the research. We cast our net widely to potentially include all adults who are native Spanish speakers and who have settled in the Ottawa region. We did not screen participants based on the amount of time that they had spent in the region. This community constituted the target audience for the survey as well as for the subsequent recipient evaluation of the translated texts.

The goal of the survey was not only to gather basic demographic data, but also to identify the reasons *why* members of this community might find it valuable to have the information on the OPL website made available in Spanish. As stressed by Edwards (1992, 48), probing the underlying reason, rather than simply asking yes/no type questions, provides valuable insight into a linguistic community. In this case, the reason for wanting a text translated might have a bearing on the translation approach that is most appropriate for meeting these needs.

The university researchers prepared the voluntary and anonymous survey in

"neutral" Spanish. In other words, the survey was not localized to any particular regional variety of Spanish but was intended to be comprehensible by all Hispanophones. This survey was approved by the University's Research Ethics Board. Questions aimed at establishing the respondents' demographic profile included those relating to age; sex; country of origin; length of time in Canada and in Ottawa; reason for moving to the region; projected length of stay in the region; level of education; dominant language; ability to understand English and French; frequency of use of the public library; 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. In addition, there was a question about why the respondents wanted to have the website available in Spanish.

A draft was pilot tested by four volunteers from the target community. 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.

These community members also gave useful advice regarding participant recruitment. While there are numerous benefits associated with CBPR, there are also challenges, many of which have to do with gaining community "buy in." For instance, several researchers (e.g., Horowitz et al. 2009, 2639; Willison 2013, 11) have observed that it takes time and patience to develop trust and relationships with the community. The project research assistant took the lead in community-building efforts. He began cultivating relationships in the early stages, making appointments to meet personally with

key figures to explain the goals of the project and to provide essential documentation, such as a project description summary and a certificate of approval from the university's Research Ethics Board. These key figures often act as gatekeepers, so to gain more direct access to the community, it was necessary to first gain their trust. Gatekeepers with whom the research assistant met personally included the ambassadors of the embassies of Colombia, El Salvador and Guatemala; chaplains of two local Spanish-language congregations; and Spanish-speaking settlement workers in local immigrant-serving organizations.

Once a trust relationship had been built with the gatekeepers, the research assistant continued to build relationships with the community more directly (e.g., attending church services and community events) to explain the project and its potential benefits to the community, to answer questions, and to gather early stage input. By the time the survey was distributed, the community had an understanding of the project and its intentions, and some members had already given input into the development of the survey tool or recommended additional distribution channels (e.g., a local online Spanish newspaper and distribution list).

While some community members were comfortable with information technology, this was not the case for all potential participants. We therefore faced an additional challenge—one that had also been identified by Willison (2013, 11), who noted that while CBPR can be a valuable means for identifying the needs of marginalized communities, these same communities may face barriers that hinder their participation, such as lack of access to or comfort with technology. 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, we decided to distribute the survey in both an online format (using Fluid Surveys[1]) and a paper-based format. The survey was conducted over a five-week period in October and November 2013. This dual format delivery proved to be a prudent course of action as one third of the completed surveys were ultimately returned in hard copy.

3.2.1 Phase II findings and discussion

Overall, 114 completed responses were received. Fifty-three percent of the respondents were women and 47% were men. Sixty percent of the respondents were aged between 31 and 50 years, while 20% were under the age of 30 and 20% were over the age of 50. The respondents were from 17 different countries—Spain and 16 countries in Latin America—with larger groups of respondents from Colombia (30%), Mexico (18%) and Chile (9%).

Fifty-seven percent of respondents had come to Canada to work and 29% to study, while 20% did not specify a reason. Thirty-four percent had been residing in Canada for less than 2 years, 28% for 2 to 5 years and 38% for over 5 years. Seventy-five percent intend to stay in Canada for 10 years or more, while 23% plan to remain for a period between 2 and 10 years, and 2% for less than 2 years.

With regard to the highest level of education completed, 29% of respondents had a graduate degree, 22% had an undergraduate degree, and 25% had completed post-secondary vocational training. Meanwhile, 19% had completed secondary school, and only 5% had not completed secondary school. Therefore, the Spanish-speaking immigrants who responded to this survey were relatively well educated, with 76% having

[1] Fluid Surveys: http://www.fluidsurveys.com [Retrieved February 22, 2014]

completed some type of post-secondary education or training.

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

In response to the question asking whether they can function in at least one of the official languages, 31% of respondents claimed poor comprehension in both English and French, while the remaining 69% reported that they are able to function reasonably well in at least one of these two languages. This is not too surprising given that 66% of respondents had been living in the region for more than two years. Of these longer-term residents, 58% claimed good or excellent knowledge of English and 32% claimed good or excellent knowledge of French. Meanwhile, among the more recently arrived newcomers, the number claiming good or excellent knowledge of English fell to 18% and those having good or excellent knowledge of French fell to 24%.

In response to questions regarding library use, 68% of respondents stated that they use use the public library, while 32% reported that they do not. Of those who use the library, 48% do so at least once per month, while the remaining 52% are less frequent users. Meanwhile, in answer to a question about whether they would be more likely to increase their library use if more informative material, such as the website, were available in Spanish, 73% of respondents said definitely or probably, whereas 20% responded possibly, and only 7% indicated that it was unlikely.

Comparing the responses of longer-term residents with more recently arrived newcomers reveals some surprising results. Of those who had been in the region for less than two years, 69% indicated that they would definitely or probably use the library more if Spanish material were available. However, this number rose to 76% among longer-term residents, which seems to indicate that it is not only recently arrived newcomers who appreciate having Spanish material.

Finally, respondents were asked to give the reason(s) *why* they would like to have the website available in Spanish.[2] Only 11% indicated that they were unable to understand the English text.[3] This is not too surprising if we recognize that, as noted above, only 19% of respondents felt that their ability to understand English was poor. Meanwhile, 25% of respondents indicated that they were reasonably sure that they understood the English text, but they would like to have access to a Spanish translation 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, noting that while they were 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 of ability.

A relatively small percentage of respondents,11%, replied that they enjoy having Spanish texts because it is a means of preserving or reinforcing their cultural heritage. Meanwhile, another 20% indicated that they would use the availability of Spanish as an

[2] Respondents were permitted to give more than one reason, so percentages do not add up to 100.

[3] For resource-related reasons, this experiment was conducted exclusively between English and Spanish. Some respondents who indicated that they could not understand English may have been able to understand French.

opportunity to teach this language to others, such as their children. One quarter of respondents wanted to use the Spanish translation as an aid for learning English by consulting the two texts comparatively.

One other reason was put forward by a respondent who noted that having access to a translation could ease the pressure of integration. The respondent commented that, while largely valuable, the immersion process can be intense and that occasional opportunities for lightening the cognitive load were welcome.

Two others observed that their needs had evolved over time. They noted that while they have been living in the region for some time and are now comfortable in English, they would have appreciated this service at the time of their arrival.

Meanwhile, one additional respondent argued against providing material in immigrant languages, claiming that it contributes to linguistic ghettoization. This respondent suggested providing more support to learn Canada's official languages, rather than providing Spanish texts. However, as described above, 25% of the respondents indicated that a motivation for wanting the text in Spanish was to use this as a tool to help improve their second-language skills. Similarly, another 25% of respondents wanted to use the Spanish version to confirm their comprehension of English. When viewed in this light, rather than promoting linguistic ghettoization, translations can help immigrants to build knowledge and confidence in English and thus to ease their integration.

3.3 Phase III: Conducting a recipient evaluation of translated texts

Phase III polled community members about the usefulness of different forms of translated texts. While unquestionably faster and cheaper than human translation (HT), raw MT output is typically of a lower quality than HT. If MT is to be considered a viable

option for OPL, the intended recipients must be willing to accept it. A key question is whether Spanish-speaking immigrants will find some form of MT output to meet their needs.

There are multiple approaches for evaluating the quality of MT systems. This CBPR project views translation as a service that is dependent on customer satisfaction (Chesterman and Wagner 2002, 80–84). To measure translation quality, we therefore need to measure customer satisfaction, which can be done using a *recipient evaluation* in order to evaluate quality, cost and speed (Loffler-Laurian 1996, 69; Trujillo 1999, 255–256). All three parameters are pertinent in the context of the present investigation. The ability of newcomers to effectively use the target texts will be largely determined by the translation quality. In addition, the level of quality required will have an impact on both the cost and speed of producing translations. If a lower quality translation can be produced in a quicker turnaround time, it will make it easier for the library to update the website in a timely fashion. Finally, given that the library has a fixed budget for translation, the cost of translation will determine the volume of text that can be translated. For instance, if customers are willing to accept lower quality texts that can be produced more cheaply, then it will be possible to translate a higher volume of text for the same price.

Experience has long shown that MT is not a good option for all types of texts or situations (e.g., Church and Hovy 1993). It is also generally accepted that MT makes a better aid than a replacement for human translators (e.g. Koehn and Haddow, 2009; Simard and Isabelle, 2009). We therefore sought to experiment with different text types, and with both raw and edited MT output.

With regard to texts, the OPL librarians had identified a section of the site for Newcomers that they thought may be useful to translate into Spanish.[4] From this section, we selected three extracts to use in the recipient evaluation. These were not the same texts used in the MT system evaluation phase, but they were of the same general types.

Text 1: "Your Library Card" (380 words)—informative prose format

Text 2: "FAQs on borrowing" (368 words)—informative Q&A format

Text 3: "How to place a hold" (301) words)—instruction format

Raw MT can be more or less heavily edited based on user requirements. Allen (2003 distinguishes between two levels of intervention: Maximal Post-Editing (MPE) and Rapid Post-Editing (RPE). MPE addresses both content and style with the goal of editing the MT output to a level that closely resembles a professional human translation, whereas RPE only addresses the accuracy of the content, but not readability or stylistic issues. The Translation Automation User Society has developed guidelines for applying MPE and RPE to MT output (TAUS 2014), and these were provided to the translators.

3.3.1 Preparatory work

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

1) A professional human translation (HT);
2) A maximally post-edited (MPE) machine translation, whose content and style were corrected to produce a text resembling a human translation;
3) A rapidly post-edited (RPE) machine translation (i.e., content errors were corrected, but no stylistic changes were made);

[4] The Newcomers section of the OPL website can be viewed at http://biblioottawalibrary.ca/en/newcomers

4) A raw (unedited) machine translation (MT).

The time and cost of producing each version was also calculated because the parameter of quality must be weighed against those of time and cost. The methods used to determine time and cost are described below, while the actual time and cost required to produce each version are summarized in Tables 4 and 5.

Raw MT output was produced using Google Translate. The translation is produced almost instantaneously, and there is no cost involved. Therefore, the time and cost for producing raw MT are considered to be zero for this experiment.

Three professional translators were hired to produce the remaining texts (RPE text, MPE text and 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 with a comparable background and experience as described above. As summarized in Table 3, each translator undertook a different task for each text and to control for potential order effect, they carried out the tasks in a different order. By having each translator perform each of the three tasks, we were able to calculate an average time per 100 words for each task, as shown in Table 4.

[Insert Tables 3 and 4 near here.]

	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 in which tasks were performed by each translator.

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 HT, MPE, RPE and MT, as well as the associated savings.

Table 5 shows the cost of producing the different versions, which we calculated using data from the "2012 Survey on Rates and Income" produced by a large and well-respected Canadian professional translators' association (Gauthier 2012, 3). According to the 336 respondents, the average cost of translation is CAD$56.73 per hour, while the average hourly rate for editing is CAD$54.85.

[Insert Table 5 near here.]

Method	**Estimated average cost of producing 3 texts (1049 words) for OPL website**	**Savings compared to cost of HT**
HT	CAD$131.42	--
MPE	CAD$89.59	CAD$41.83 (32%)
RPE	CAD$43.88	CAD$87.54 (67%)
MT	CAD$0[5]	CAD$131.42 (100%)

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

[5] In fact, there are some costs associated with MT. For example, among other tasks, it would be necessary for someone to submit the source texts to the MT system, and then to insert the MT output into the target-language versions of pages and to publish those pages on the OPL website.

Not surprisingly, the data show that raw MT is always the fastest and cheapest form of translation, followed by RPE, then MPE, and finally HT. In this experiment, texts produced using MPE required 30% less time and were on average 32% cheaper than HT. If OPL wants to provide translations in multiple languages, this could potentially add up to significant savings. Of course, such savings are only meaningful if the texts are accepted by the target community. Therefore, the next important step was to present these different translations to the Spanish-speaking immigrants in order to gauge the relative acceptability of the translations.

3.3.2 Recipient evaluation

The recipient evaluation had several steps. First, participants were presented with the following three texts and asked to select one[6]: Text 1—"Your Library Card"; Text 2—"FAQs on borrowing"; or Text 3— "How to place a hold."

After selecting a text, participants were shown four different Spanish translations of that text, produced using the four methods described above (HT, MPE, RPE and MT). These four versions were presented in random order with no indication of which method had been used to produce each. Participants were simply asked to read each version and to indicate which one best met their needs.

Next, participants were shown the same four translations, but now, they also saw the metadata about the production method, time and cost. Balancing these parameters,

[6] Note that in the online survey, the randomize feature of the Fluid Surveys tool was used to present the participants with the choice of the three possible texts in a random order. For the paper-based survey, three different versions of the survey were prepared (one for each text) and these were randomly distributed.

participants were asked to select the text that best met their needs[7].

3.3.3 Phase III findings and discussion

The results of the recipient evaluation are summarized in Table 6. They are first reported for each of the three texts, and then averaged. The left column shows the production method for each version. The second column shows the percentage of people who stated that a particular version (HT, MPE, RPE or MT) best met their needs when the production method, time and cost were unknown. The third column shows the percentage of people who stated that a particular version (HT, MPE, RPE or MT) best met their needs when the production method, time and cost were known. The final column shows the change in the number of people selecting a particular version once the metadata about method, time and cost were provided. This change is recorded as an increase (+) or decrease (-) over the original, expressed as a percentage.

[Insert Table 6 near here.]

	Method/Time/Cost Unknown	**Method/Time/Cost Revealed**	**Difference (change in number of people selecting this option)**
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" (Q&A style) (49 respondents)			
HT	57%	12%	-45%

[7] As noted previously, each of these parameters is relevant to the end-user or recipient. Quality affects the usability of the text, time is pertinent because it determines how quickly information on the website can be updated, and cost is relevant because it determines the total volume of text that can be translated.

MPE	27%	18%	-9%
RPE	14%	57%	+43%
MT	2%	12%	+10%
Text 3: "How to place a hold" (instructional) (39 respondents)			
HT	33%	3%	-30%
MPE	23%	13%	-10%
RPE	23%	46%	+23%
MT	21%	38%	+17%
Average for all three 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 metadata about production method, time and cost had been revealed.

The findings show that the raw MT was always selected by the fewest number of people when the method, time and cost were unknown. This number rose substantially—by an average of 14 percentage points—once this metadata was revealed.

In contrast, the number selecting either HT or MPE in the first instance was high; on average, these two categories combined were preferred by 66% of the respondents. However, once method, time and cost were given, this preference fell to 23%.

In the final analysis, after method, time and cost metadata were revealed, the method that was preferred by the greatest number of participants—for each individual text and the average of all three—was RPE. This form of text, in which the content is accurate but the style may not be elegant, was deemed by over half the respondents to provide the best value given their translation needs.

In the discussion in section 3.2.1 concerning the reasons that newcomers wanted to have these texts translated, 44% of respondents expressed a desire to have a Spanish

translation because they could process it more quickly, while another 25% wanted the Spanish translation to confirm that they had understood the English. Another 11% wished to have a Spanish translation because they had very limited proficiency in English and could not understand the original. For such needs, an elegant text is not required; one that provides the correct gist is sufficient.

However, some respondents gave other reasons for wanting a translation. For instance, 11% saw a Spanish version as a means of preserving their culture, while 20% of respondents would use the Spanish-language material to teach this language to others (e.g., children). To address these needs and achieve customer satisfaction, a translated text must be more than merely accurate with regard to meaning. This may explain why an average of 23% of respondents still preferred either HT or MPE, even after the time and cost had been revealed. It is interesting to note, however, that of this 23%, 17% were willing to accept the more economical option of MPE, while only 6% insisted on HT.

When considering text type, it is worth noting that the instructional text is the one for which the greatest number of respondents indicated a willingness to accept RPE (46%) or even MT (38%). This is in keeping with the widely held observations that have been made about text types that are generally considered suitable for MT (e.g., Polvsen et al. 1998). The purpose of an instructional text is to make it clear to readers what steps must be followed to achieve a goal. As such, these texts tend to contain short, simple structures and precise terminology. The relative lack of syntactic and semantic ambiguity in this text type makes it a good candidate for MT. The findings of our study support this general observation.

4. Concluding remarks

This investigation adopted a CBPR approach in which academic researchers worked alongside the Spanish-speaking immigrant community in Ottawa, and alongside the OPL, which seeks to offer improved services to this community. As observed by Horowitz et al. (2009, 2636) and Willison (2013, 7), community participation can help to ensure that the goals of the study are relevant to the population, which can in turn help to identify service priorities and to better allocate resources.

This study revealed that many—though not all—of the translation 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 translations to be able to process information more quickly, or to confirm their understanding of the original. This experiment demonstrated that raw MT output that is lightly or maximally edited can produce translations that address these needs for approximately two-thirds of the population. Moreover, post-edited texts can be produced more quickly and cheaply than HT. In the case of RPE, texts produced during this experiment offered a cost savings of 67% over professional HT, and required only one-third the time to produce.

This study also confirmed that translations of some types of texts, such as instructional texts, seem to have a high user acceptance rate when produced using raw MT or RPE. The general question surrounding translatability and text types leads to an important point about the value of planning and writing for translation. On this subject, Sichel (2009, 3-4) advises “Take the time to think your project through from the perspective of the recipient, and do some research if you don’t know the recipient’s perspective.” She goes on to note that where cost is a factor, it may not be possible (or

necessary) to translate everything, but that understanding the recipients' needs can help an organization, such as OPL, to identify which texts are most needed by the community in question, and to prioritize those in the translation queue. Meanwhile, Brown (2003, 4) emphasizes that if an organization knows in advance which texts will be translated, then these texts can be constructed in a way that will facilitate their eventual translation. This process is sometimes referred to as controlled authoring (Ó Broin 2009, 12) or content internationalization (Brown 2003, 4), and the general goal is produce a text that is clear and unambiguous, so that it can be translated more easily. Numerous sets of guidelines are available for helping content developers to write their texts in a way that will be translation-friendly, whether this translation is to be carried out by a professional translator or by an MT system (e.g. Kohl 2008; O'Brien 2003). This becomes particularly important if the source text is to be translated into multiple target languages because eliminating a potential problem (e.g. an ambiguous construction) in the source text is much faster and cheaper than having to address the resulting problem in each of the different target texts. In the case of OPL, which would like to translate sections of their Newcomer website into nine languages, an approach that specifically includes planning and writing for translation could result in significant savings.

Overall, the results of this pilot project appear promising and seem to merit further study with both additional immigrant languages and additional text types and samples. These results are also in line with other recent findings. For example, Guerberof Arenas (2009) suggests that translators produce higher quality and are more productive when post-editing MT output than when processing fuzzy matches from translation memories. Meanwhile, Garcia (2011) found that post-editing MT can lead to increased quality

among trainee translators, and that with adequate training, it can potentially lead to increased productivity also.

The present research project was launched just prior to the release of the International Federation of Library Associations and Institutions' (IFLA) *Trend Report* and *Insights Document* (IFLA 2013). However, it is encouraging to note that MT figures among the five key high-level trends in the global information environment that IFLA identifies as characterizing the new digital paradigm in libraries. IFLA raises interesting questions about the use of MT 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?" We hope that the work reported here has taken a step towards addressing these important questions.

Acknowledgments

This research was supported by a 2013 OCLC/ALISE Library and Information Science Research Grant. Thanks to all project participants: A. Yarrow (Acting Manager, Diversity and Accessibility Services, OPL); C. St-Martin (Newcomer Services Librarian, OPL); C. Ginther (Manager of Digital Services, OPL); translators, R. Colon Rodriguez, D. Josephy Hernandez and J. Zapata Rojas; embassies of Colombia, El Salvador and Guatemala in Ottawa; congregations at the Iglesia Sagrada Familia and the Cathédrale Saint-Joseph; settlement workers at the Ottawa Community Immigrant Services Organization and the Ottawa Catholic Immigration Centre; *Mundo Español* (Ottawa) and the Ottawa "Spanish-language meet-up"; and respondents from Ottawa's Spanish-speaking newcomer

community, including the survey testers. Finally, we are grateful to the anonymous reviewers who provided thorough, constructive and thought-provoking comments.

References

Allen, Jeffrey. 2003. "Post-editing." In *Computers and Translation: A Translator's Guide*, ed. by Harold Somers. 297–317. Amsterdam: John Benjamins.

Brown, M. Katherine. 2003. "Trends in Writing for Translation." *MultiLingual Computing & Technology: Supplement on Writing for Translation* 59 (October/November): 4–8.

Caidi, Nadia, and Danielle Allard. 2005. "Social Inclusion of Newcomers to Canada: An Information Problem?" *Library and Information Science Research* 27(3): 302–324.

Chesterman, Andrew, and Emma Wagner. 2002. *Can Theory Help Translators? A Dialogue Between the Ivory Tower and the Wordface*. Manchester: St. Jerome.

Church, Kenneth W., and Eduard H. Hovy. 1993. "Good Applications for Crummy Machine Translation." *Machine Translation* 8: 239–258.

City of Ottawa: http://ottawa.ca [Retrieved February 22, 2014]

Costa-Jussà, Marta R., Mireia Farrús, José B. Mariño, and José A. R. Fonollosa. 2012. "Study and Comparison of Rule-Based and Statistical Catalan-Spanish Machine Translation Systems." *Computing and Informatics* 31: 245–270.

Edwards, John. 1992. "Sociopolitical Aspects of Language Maintenance and Loss: Towards a Typology of Minority Language Situations." In *Maintenance and Loss of Minority Languages*, ed. by Willem Fase, Koen Jaspaert, and Sjaak Kroon, 37–54. Amsterdam: John Benjamins.

Frisken, Frances, and Marcia Wallace. 2002. *The Response of the Municipal Public Service Sector to the Challenge of Immigrant Settlement (Revised Edition).* Citizenship and Immigration Canada. http://www.settlement.org/downloads/municipal_sector.pdf [Retrieved March 24, 2014].

Garcia, Ignacio. 2011. "Translating by Post-editing: Is it the Way Forward?" *Machine Translation* 25(3): 217–237.

Gauthier, François. 2012. "2012 Survey on Rates and Income." Translated by Karin Montin. Quebec, Canada: Ordre des traducteurs, terminologues et interprètes agréés du Québec.

George, Usha. 2002. "A Needs-based Model for Settlement Service Delivery for Newcomers to Canada." *International Social Work* 45 (4): 465–480.

Guerberof Arenas, Ana. 2009. "Productivity and Quality in the Post-editing of Outputs from Translation Memories and Machine Translation." *Localisation Focus* 7(1): 11–21.

Holt, Glen E. (ed). 2009. "Library Services for Immigrants, an Abridged Version." *Public Library Quarterly* 28(2): 120–126.

Horowitz , Carol R., Mimsie Robinson, and Sarena Seifer. 2009. "Community-Based Participatory Research From the Margin to the Mainstream: Are Researchers Prepared?" *Circulation* 119: 2633–2642.

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.

Koehn, Philipp, and Barry Haddow. 2009. "Interactive Assistance to Human Translators using Statistical Machine Translation Methods." In *Proceedings of MT Summit XII*, Ottawa, Canada (August 26-30, 2009). http://www.mt-archive.info/MTS-2009-TOC.htm [Retrieved November 25, 2014].

Kohl, John. R. 2008. *The Global English Style Guide: Writing Clear, Translatable Documentation for a Global Market*. Cary, NC: SAS Institute, Inc.

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, Susan. 2012. "Conceptualizing Public Library Service to Newcomers in Canada." In *Proceedings of the 40th Annual Conference of the Canadian Association for Information Science*, ed. by Anabel Quan-Haase, Victoria L. Rubin, and Debbie Chaves. http://www.cais-acsi.ca/proceedings/2012/caisacsi2012_60_macdonald.pdf [Retrieved March 2, 2014]

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

Ó Broin, Ultan. 2009. "Controlled Authoring to Improve Localization." *MultiLingual: Getting Started Guide on Writing for Translation* (October/November), 12-14.

O'Brien, Sharon. 2003. "Controlling Controlled English: An Analysis of Several Controlled Language Rule Sets." *Proceedings of the Joint Conference combining*

the 8th International Workshop of the European Association for Machine Translation and the 4th Controlled Language Applications Workshop (Dublin, May 15-17, 2003) http://www.mt-archive.info/CLT-2003-Obrien.pdf [Retrieved November 28, 2014]

Ottawa Public Library (Newcomers): http://biblioottawalibrary.ca/en/newcomers [Retrieved November 25, 2014]

Papineni, Kishore, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. "BLEU: A Method for Automatic Evaluation of Machine Translation." In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, 311–318. Stroudsberg, PA.

Polvsen, Claus, Nancy Underwood, Bradley Music, and Anne Neville. 1998. "Evaluating Text-type Suitability for Machine Translation a Case Study on an English-Danish MT System." In *Proceedings of the First International Conference on Language Resources and Evaluation, Vol. 1*, ed. by Antonio Rubio, Natividad Gallardo, Rosa Castro, and Antonio Tejada, 27–31. Paris: European Language Resources Association.

Sichel, Barb. 2009. "Planning and Writing for Translation." *MultiLingual: Getting Started Guide on Writing for Translation* (October/November), 3-4.

Simard, Michel, and Pierre Isabelle. 2009. "Phrase-based Machine Translation in a Computer-assisted Translation Environment." In *Proceedings of MT Summit XII*, Ottawa, Canada (August 26-30, 2009). http://www.mt-archive.info/MTS-2009-TOC.htm [Retrieved November 25, 2014].

Statistics Canada. 2011. *2011 Census of Canada*. http://www12.statcan.gc.ca/census-

recensement/index-eng.cfm [Retrieved February 20, 2014]

Thurmair, Gregor. 2004. "Comparing Rule-based and Statistical MT Output." In *Proceedings of the Language Resources and Evaluation Conference (LREC) 2014 Workshop on the Amazing Utility of Parallel and Comparable Corpora*: http://www.lrec-conf.org/proceedings/lrec2004/ [Retrieved November 25, 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." In *Computers and Translation: A Translator's Guide*, ed. by Harold Somers, 211–244. Amsterdam/Philadelphia: John Benjamins.

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.