Gist MT Users: A Snapshot of the Use and Users of One Online MT Tool

Mary Nurminen
University of Tampere
Kalevantie 4
FI-33100 Tampere, Finland
mary.nurminen@uta.fi

Niko Papula
Multilizer
Kimmeltie 3
FI-02110 Espoo, Finland
niko.papula@multilizer.com

Abstract

This study analyzes usage statistics and the results of an end-user survey to compile a snapshot of the current use and users of one online machine translation (MT) tool, Multilizer's PDF Translator[1]. The results reveal that the tool is used predominantly for assimilation purposes and that respondents use MT often. People use the tool to translate texts from different areas of life, including work, study and leisure. Of these, the study area is currently the most prevalent. The results also reveal a tendency for users to machine translate documents that are in languages they have some understanding of, rather than texts they do not understand at all. The findings imply that gist MT is becoming a part of people's everyday lives and that perhaps people use gist MT in a different way than they use publishing-level translations.

1. Introduction

Online machine translation (MT) tools have been in use for almost 25 years and people are finding numerous ways to integrate MT into the processes of their everyday lives. However, although research on professional translators' use of MT has grown rapidly, the literature on all other users of MT remains limited. This paper aims to contribute to that limited body of research with a study on the users of one online MT tool, Multilizer's PDF Translator.

1.1. Purpose of the Study

Our study focuses on users of MT for assimilation, or scenarios in which people use raw, unedited machine translated text for some other purpose than editing it for publication. Because users most often want just a basic understanding of the information (or gist) of the text, we term them *gist MT users.* We also use it because it is shorter than the term *users of MT for assimilation*; however, we use the two terms interchangeably.

The overall purpose of the study is to present a snapshot of the use and users of one online MT tool. Our questions concern who is using MT, where these users are, how they are using it, when they are using it, and in what areas of life they are using it.

We had several motivations in doing the study. First, because online MT is in such wide use today, we can assume that the number of gist users is much larger than the number of professional translator users. Yet the latter group has been studied far more than gist MT users. We believed it was time to put some focus on other user groups and we hoped to contribute to that with this study. Second, our literature review revealed only one gist MT user survey conducted in the past 10 years. We felt it was time to conduct another one. Finally, this analysis will serve as a basis for a second study we are planning, a qualitative study that will probe more deeply into the specific ways people are using gist MT.

1.2. Related Work

The pioneer study of MT users, by Henisz-Dostert in 1979, was also the first study on gist MT users. In the 40 years since it was published, a relatively small number of articles have been written about gist MT users. These studies can be grouped into two categories: experimental studies on *potential* users of gist MT and survey studies on *actual* gist MT users.

In the experimental studies, groups of potential gist MT users were asked to evaluate specific aspects of MT or the use of MT. Fuji et al. (2001) tested user success with machine translat-

[1] pdf.multilizer.com

Pérez-Ortiz, Sánchez-Martínez, Esplà-Gomis, Popović, Rico, Martins, Van den Bogaert, Forcada (eds.)
Proceedings of the 21st Annual Conference of the European Association for Machine Translation, p. 199–208
Alacant, Spain, May 2018.

ed texts, measured through reading comprehension, against users' *impressions* of comprehensibility and awkwardness. Gaspari (2006) had users evaluate their *confidence* in understanding raw MT. Bowker and Ehgoetz (2007), Bowker (2009), Bowker and Buitrago (2015) and Castilhjo and O'Brien (2017) had users evaluate the *acceptability* of raw MT. They often had users compare preference or acceptability of raw MT, post-edited MT, and human translation. Gaspari (2004), Stewart et al. (2010), and Doherty and O'Brien (2012) had users evaluate raw MT against traditional usability criteria. Finally, Doherty and O'Brien (2014) used eye tracking to measure MT output usability.

The studies on actual gist MT users include research on market or usage reports, end-user surveys, or a combination of the two. A small number of these were agnostic to MT systems, focusing on groups who were using any number of the systems available at the time. A larger group of research focuses on users of one specific system. A limitation of this second group of studies is that they describe only a specific type of user and therefore the results cannot be considered representative of all MT users. However, they do contribute information on those users and, seen collectively, help to paint an overall picture of gist MT usage.

The first studies on users of various MT systems were sponsored by the International Association for Machine Translation (IAMT) in 1993 and 1995. These studies used participants they recruited through the manufacturers of MT systems or through the AMTA website. Although they focused mainly on professional translators, who used MT for dissemination, they did include a small amount of data on gist MT users in the form of eight testimonials (Lawson and Vasconcellos, 1993). The Asia-Pacific Association for Machine Translation (AAMT) recruited participants for a series of studies in 2003-2005 through their website, so the user group represented was again not specific to any one tool. These surveys focused much more on gist MT users, indicating that "the main use of machine translation" was assimilation (Yamada et al. 2005, p. 58). The final study that was not dependent on any one MT tool was that carried out by Gaspari in 2007. The survey, conducted at several UK university campuses, used students as informants and covered user demographics, experience with computers and MT, languages translated, use of MT for assimilation vs. dissemination, genres translated, and user evaluations of MT.

The first study that focused on users of a specific system was the study on the users of the Georgetown MT system cited earlier in this article (Henisz-Dostert, 1979). It used a survey, although that survey was administered almost entirely through face-to-face interviews. It provided a rich and multifaceted description of the users, how they used the system, and their experience regarding usefulness, speed, and quality. The study also included a few interesting questions on how users experience cognitive processes, which subsequent surveys have not touched on. These included questions such as "If the style of the MT is awkward, can you correct it mentally?" and "Do you get 'used to' reading MT?" (Henisz-Dostert 1979, p. 193) The only other study we are aware of that address cognitive processes was Doherty and O'Brien's (2014) previously mentioned eye tracking study.

The next study of the users of one system was conducted in Japan by Hoshino (1995), focusing on users of the Korya Eiwa ("It's Nice! English–Japanese") consumer desktop system. The survey was comprehensive, covering user demographics, genres and subject matters translated, users' fluency in English, experience with MT, purpose, motivations, and expectations for MT. Flanagan's (1996) paper described the usage of CompuServe's online MT service as well as users' reactions to it. Another online service, AltaVista Translation with Systran, was the focus of a study by Yang and Lange (2003). The study included both an analysis of usage and feedback data in the form of 5,005 e-mails received in 1998.

A few studies have been conducted on company-internal MT systems and their users. Smith (2003) analyzed PriceWaterhouseCooper's intranet-based MT system and its users. This was perhaps the first study on a system that supports a large number of language pairs, 37 in total. It described how people used the system, their reactions to it, and factors that affected users' satisfaction with the system. Another company-internal study was conducted by Nuutila (2005), who reported on a survey conducted with users of Nokia's Roughlate MT service.

The latest user study we are aware of was a study by Burgett (2015) on the users of Intel's machine-translated support content. This study asked users to perform usability tests while working with Intel's machine translated content.

2. Multilizer's PDF Translator

The tool in our study, PDF Translator, is an online MT tool that translates full documents that are in either PDF or Word format. A user submits a document, then the tool extracts the texts, puts them through machine translation, rebuilds the document with the original pictures in place, and returns it to the user in the requested language. PDF Translator utilizes the MT engines of Microsoft, Google and PROMT to perform the translations. Due to the proprietariness of the engines and the dynamic nature of MT development, we do not have information on the exact type of MT (rule-based, statistical or neural) used for each language pair during the time of the study.

PDF Translator is meant for any type of document that people want to have translated, so it is not trained for specific genres or subject matters. Two versions are available, a desktop and an online version. The desktop version, which was developed first, is downloaded onto the user's computer and used from there. Its user interface is available in 14 languages. Users can translate up to 3 pages at a time for a total of 15 pages for free. PDF Translator offers three levels of paid licenses: Standard, Pro and Business, and after initial purchase of a license, additional pages can be purchased in batches. The desktop version supports 47 source languages and 39 target languages. The newer online version has been in use since 2016 and it is currently available through an English, Spanish or Chinese user interface. Users can translate a small amount of text (one page) free of charge and thereafter they can purchase packages of translation (10, 50, 100, etc. pages). The online version supports translation between 42 languages.

2.1. MT for PDF and DOC Documents

One important aspect of PDF Translator is that it translates entire documents instead of pieces of text typed or copy/pasted into a text field. This holds several implications for our study and the types of users it addresses. First, it excludes incidences when people enter only one or two words, essentially using MT as a bilingual dictionary. Previous studies have found this to constitute a large portion of MT use. For example, Yang and Lange reported that "more than 50% of translations are of one- or two-word phrases" (Yang and Lange, 2003, p. 199) and Gaspari was led to devote a whole section of his PhD to "(Mis-)Using Free Web-based MT Services as Online Dictionaries" (Gaspari, 2007, p. 108). Another implication of translating whole documents is that the materials people submit for translation tend to be well-formed and written, published documents instead of more informal texts such as chat messages or personal correspondence. This can influence the areas of life where people use MT – for work and study or in their free time. A final implication is that, due to the very nature of PDF as a publication instead of an editing format, users are far more likely to be gist MT users than to be people who want to edit the material for publication. All of these factors contribute to profiling a specific type of user and need to be kept in mind when reading this study.

3. Materials and Methods

Our goal was to capture a snapshot of the use and users of PDF Translator in a short, specific point of time. We chose a four-month period, November 1, 2017 through February 28, 2018, and collected two types of data from the period for analysis. We collected log files from both the desktop and the online systems, and we conducted an online end-user survey with users of the desktop system.

Our first batch of data consisted of the logs from the desktop and online versions of PDF Translator. We used the logs to examine the times that submissions for translation were made, the places they were made from, and the source and target languages involved.

The end-user survey was short, consisting of eight questions in three categories:

Category	Questions
Basic demographics	1. What is your gender? 2. What is your age? 3. What language are you most proficient in? 4. What is the highest degree or level of school you have completed?
Frequency of use of MT tools	5. How often do you use tools that automatically translate texts, similarly to PDF Translator or Google Translate?
Questions on the specific document submitted for translation	6. Why did you want to translate the document? 7. Did you need the document for work, study, or leisure purposes? 8. How well did you understand the language of the original written document?

Table 1. Survey questions.

The reason for the brevity of the survey was that, in keeping with the idea of a snapshot, our focus was on quantity more than quality. The survey needed to be short enough so that a large number of people would be willing to answer it.

Besides keeping the survey short, we used other strategies to encourage users to respond. We offered all respondents the chance to participate in a drawing for five small prizes: 100 pages of free translation through PDF Translator. We also named it *3-minute Survey for Users of PDF Translator* under the assumption that precise information on how long it would take to answer the survey would encourage people decide to devote time to it. The average response time was, in fact, three minutes.

Due to limited resourcing, we had to make decisions on what languages to offer the survey in. We decided to offer the survey to users of the most popular 6 of the 14 languages the desktop version of PDF Translator is available in: English, Spanish, Portuguese, French, Russian and Indonesian.

An invitation to answer the survey was offered to users after they had submitted a document into PDF Translator and received the translation back. It was offered to everyone who submitted a document during that period, meaning that both heavy users of the tool and first-timers could answer.

4. Discussion

Besides the log files, our data included 1,579 responses to the three-minute survey. The response distribution by language survey is displayed in the following table.

Language survey	Number of responses
Spanish	652
Portuguese	283
French	211
Russian	188
English	147
Indonesian	98
Total	**1579**

Table 2. Survey response distribution.

PDF Translator has a large customer base in Spanish-speaking countries and this is reflected in the high number of responses to the Spanish survey. The placement of the other language surveys correlate roughly with our statistics on the countries and target languages with the most traffic during the study period. While compiling responses, we noticed that a large number of responses to the English survey (49 responses, comprising 25% of all responses), were from people who marked Indonesian as their most proficient language. We did not observe a similar phenomenon in any other language survey. We decided to move these 49 responses from the English survey to the Indonesian one. The previous table reflects the numbers *after* that change.

4.1. Locations and Languages

PDF Translator is used widely across the world. Our logs indicated that requests for translation during the study period came from 181 countries and territories. The tool's large customer base in Spanish-speaking countries is reflected in the list of the countries with the most traffic, with 10 of the top 20 spots being occupied by those countries. Other countries in the top 20 include Brazil, Indonesia, Poland, Germany, Italy, Russia, Turkey, France, Ukraine, and Portugal.

English was the most popular source language, with 85% of all documents translated during the study period being originally in English. The next languages on the list of source languages included German, Spanish, French, Portuguese, Italian, Russian, Polish, Dutch and Indonesian. Spanish led the list of the most popular target languages, followed by Portuguese, English, French, Russian, Indonesian, German, Polish, Italian and Turkish.

The top language pair of English–Spanish comprised 47% of all requests. This was expected, considering PDF Translator's customer base. Also, this language pair has appeared at the top of lists in survey and market studies for a long time, including those by Yang and Lange (2003), Smith (2003), Gaspari and Hutchins (2007) and Turovsky (2016).

Indonesian's position near the top of the language lists was interesting. The past ten years have seen a major expansion in the language palette of online MT tools (e.g. Turovsky, 2016). It appears that this expansion is beginning to produce results and new language pairs are emerging at the top of the lists of the most-translated languages. For example, Google's recent reports on the most-translated languages include the ones that have appeared at the top of these lists for years—Spanish, Russian and Portuguese—but also relative newcomers to online MT, such as Arabic and Indonesian (Turovsky, 2016). Indonesian proved to be an interesting and different market in other areas of our study as well.

4.2. Survey Participant Demographics

The overall gender demographic of survey participants showed males comprising 68% of responses, females 32%, and the group of *other*, 3%. Small differences surfaced when comparing the results of different language areas. In the Portuguese, Spanish and English surveys, males made up 61–68% of responses while in the French and Russian surveys, 82–83% of respondents were male. Indonesia was the only country in which female respondents outnumbered male (54% and 46%, respectively). The high proportion of men in most of the language surveys seems to be typical in studies of technological systems.

The age distribution shown in survey answers was also typical of that shown in technology studies, with the 19–29 age group providing the largest number of responses, 46% altogether. Similarly to the results of the gender demographic, the age demographic also contained differences in the results from different language surveys, as is shown in the following figure.

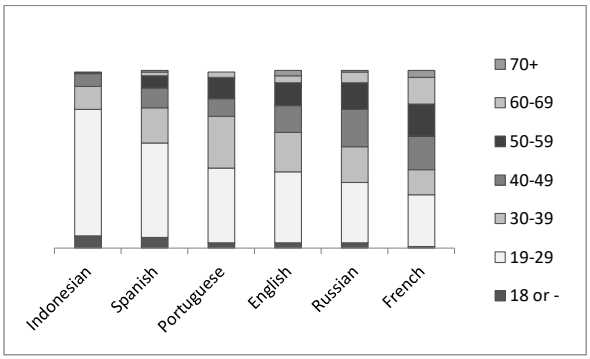

Figure 1. Age distribution of respondents in different language surveys.

Indonesian again displayed a different profile from the other surveys. In that survey, the 19-29 age group made up 71% of the total, 18 percentage points higher than the next (Spanish) survey. The French and Russian surveys were again at the opposite end of the scale, with a much more even distribution of ages. Another interesting point was that the French-speaking older respondents seem to be the most active. Whereas in most of the language surveys, the two highest age groups comprised 3–7% of respondents, in the French survey this group comprised 19% of all respondents. Although the total overall number of answers in the highest age groups, 60–69 and 70 or older, was small (68 and 19 responses respectively), it was good to note that people in these age groups are also using MT actively.

The following figure shows how much of each respondent age group was comprised of female, male and other genders.

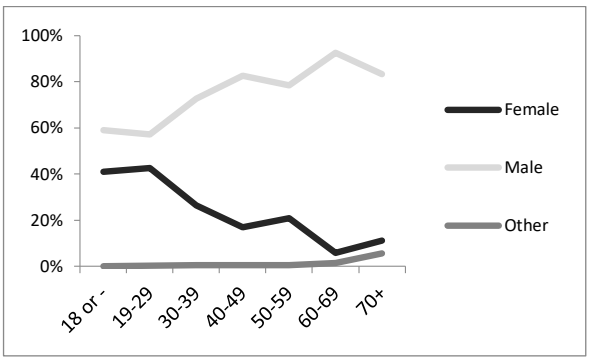

Figure 2. Percentage distribution of survey respondents by gender and age group.

The chart shows that in the younger age groups, a smaller gap exists between the male and female composition of the respondents. This gap grows and peaks in the 60–69 age group before becoming smaller again in the 70 or older group. A somewhat even number of people identify as some other gender throughout all age groups, although the relatively small overall number of respondents in the 70 or older group resulted in the *other* group comprising a higher percentage of the whole.

The highest degree or level of school reported by respondents is shown in the following table.

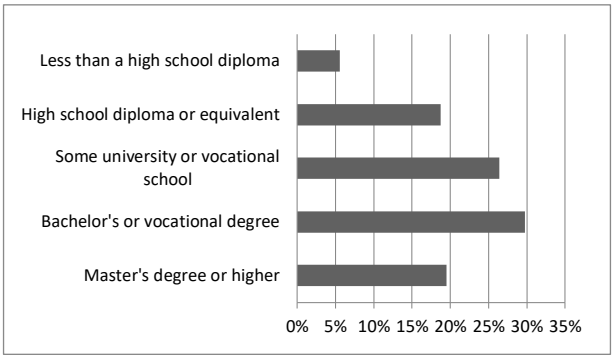

Figure 3. Highest level of education of respondents.

Respondents appear to be fairly highly educated, with the largest group being comprised of people who already have a vocational or bachelor's degree. In comparing the different language surveys, the French and Russian surveys once again stood out in that they had high percentages of respondents who held a master's degree or higher. In fact, the educational level with the most responses in both surveys was a master's degree.

4.3. Frequency of MT Use

As has been noted by Gaspari (2007) and others, a self-administered survey such as this one can result in responses being given by people who are relatively more active in the technology area than the general user population. This factor needs to be considered when examining the responses to our survey question on how often respondents use MT, which are displayed in the following chart.

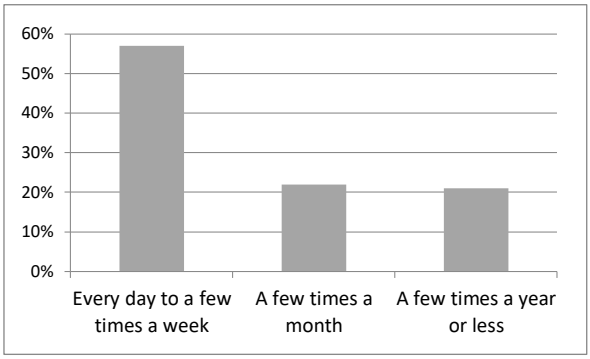

Figure 4. How often respondents report using machine translation.

These results indicate that a majority of the overall respondents of this survey tend to use MT on a very regular basis. In comparing to previous studies that have asked this question, Yamada et al. (2005) reported that only 13–18% of users used MT as frequently in 2003–2005. However, Nuutila's (2005) study showed that 63% of Nokia's in-house Roughlate system users reported using the system several times a day or at least every week.

The next chart shows a breakdown of reported frequency of use by age group.

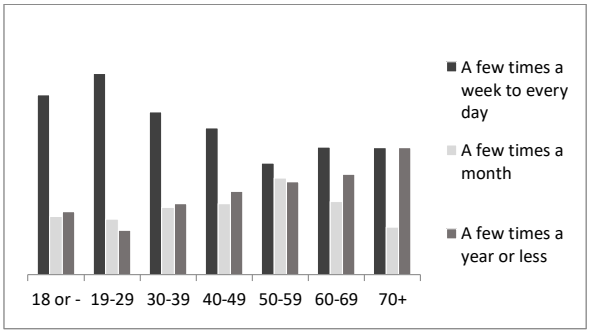

Figure 5. Frequency of MT use by different age groups.

As is shown here, the younger age groups, 18 or under and 19–29, showed a stronger tendency to use MT very frequently than respondents in older age groups. In fact, the level of very frequent use for the 19–29 group was remarkably high, 67%.

4.4. Purpose: Assimilation, Dissemination, or Something Else

To explore users' purposes for using MT, the area of life they were using MT in, and their proficiency in the languages involved, the survey included three questions that asked specifically about the document the respondent had submitted for translation right before being invited to take the survey. The first of these questions concerned whether users were using the submitted document for assimilation, dissemination, or some other purpose. Although we could assume that people translating whole documents (many of them PDFs) are mainly using MT for assimilation purposes, we wanted to verify this. We started with the questions and answer choices used by Gaspari in his survey of students (Gaspari, 2007, p. 102–103) and edited them a bit. The following table shows the overall responses.

Why did you translate the document?	**% of responses**
I wanted to understand it myself. (assimilation)	58%
I wanted to verify that I understood it myself. (assimilation)	18%
I wanted to translate it into my own language so that someone else can understand it. (assimilation for other person)	14%
I wanted to translate it from my language into another language so that someone else can understand it. (dissemination)	6%
Some other reason (please specify).	4%

Table 3. Purpose of translating the document submitted for translation.

Combining the first and second answers gives an overall view to assimilation and shows that a majority of respondents, 76%, are indeed using the machine-translated documents for their own assimilation. However, the second answer taken alone is also interesting in that it shows that people are using MT for understanding documents, but also for verifying their understanding. Another interesting point arises when comparing the responses of different language surveys. In Indonesia, 25% of respondents reported that they translated the document into their own language so that someone else could understand it. In other language surveys, the rate was only 10–16%. Combining this with the relatively young de-

mographics of that market, could this reflect an effort by younger people to help their technologically more reticent elders?

4.5. Area of Life Where MT was Used

The second of the questions we asked about the document the respondent had translated regarded the area of life that the document concerned: work, study, or leisure. We allowed respondents to select more than one choice in case the document was used in various areas. However, only 11% chose more than one area. The following figure displays the overall compiled results of responses to the question.

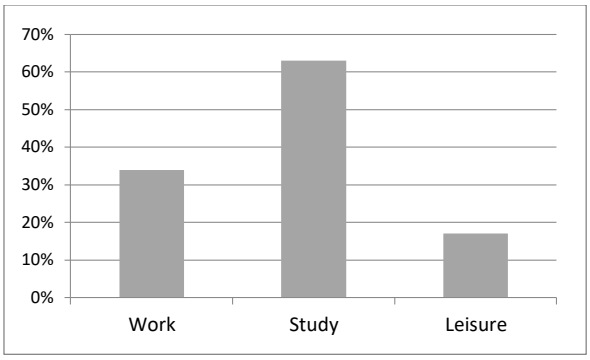

Figure 6. Percentage of respondents who listed work, study, and/or leisure as the purpose of the document they translated.

Overall, 63% of the respondents reported that at least one of the areas of life in which they needed the translated document was study. This would indicate that, at least for the type of user who is translating whole documents (and willing to answer surveys), MT is being used widely for learning purposes.

This figure shows the responses by age group.

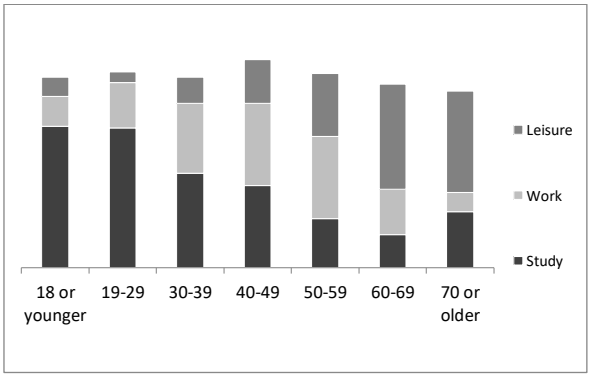

Figure 7. Reported area of life where machine translated document was used, by age group.

This distribution seems logical and perhaps expected, with users in the younger age groups showing a relatively strong emphasis on study. It is interesting that the study category increased again in the 70 or older age group, though it should be kept in mind that the number of responses in that group was small (19), and that respondents who are active users of MT, and are willing to answer surveys, might well also have a keen interest in self-study.

Two factors seem to have contributed to making *study* the top area reported. First, a relatively high number of responses to the survey came from the 19–29 age group. Second, responses from the Spanish and Portuguese surveys were also relatively high, and as can be seen in the following table, both of those languages showed very high scores for *study*.

Survey	Work	Study	Leisure
Indonesian	19%	88%	4%
Portuguese	30%	73%	15%
Spanish	31%	75%	9%
English	46%	49%	19%
French	43%	34%	39%
Russian	44%	36%	31%

Table 4. Percentage of respondents who listed work, study, and/or leisure as the purpose of the document they translated in different surveys.

In this table, the English, French, and Russian answers reflect more of an emphasis on work. In fact, in the French and Russian results, work surpasses study as the area of life the translated document concerned. As discussed earlier in this article, the demographics of the French and Russian respondents were somewhat different than those of the other language surveys. These differences seem to indicate that the way MT is used can be different in different groups or geographical areas.

In addition to analyzing the responses to our survey, we also used the log files to analyze the day of the week and time of day when people requested translations. We converted all log time stamps to local times. The results of that analysis are presented in the following figure, which shows usage levels for the seven days of the week and hour-by-hour. Each of the seven lines in the graph represents one day of the week. Black lines were used for Monday–Thursday and gray for Friday–Sunday.

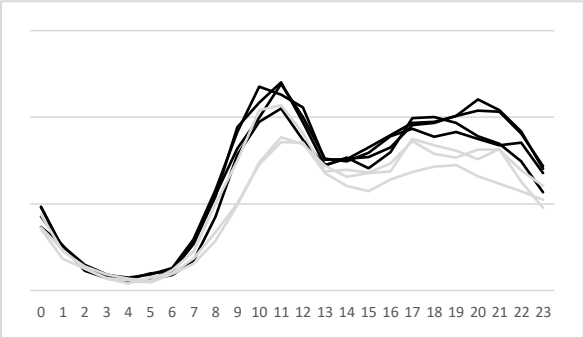

Figure 8. Usage by day of the week and time of day. The black lines are Monday–Thursday and gray lines Friday–Sunday.

Although all lines demonstrate activity during the evening hours, a clearly higher activity level emerges on Monday–Thursday than on Friday–Sunday. This analysis seemed to support the result that study and work are areas of life where users of the tool request translations, more than leisure.

It should be noted that these results reflect the situation for one tool at a specific point in time. As the technology and users mature, the overall emphasis could shift from study to other areas of life. Another point of consideration is that our results do not provide details on the level of education users are at when they use MT for study. It could be anything from grade school through Ph.D. research. The results also do not tell us exactly how users are using the machine-translated information: to help them in language production, for self-study, or to read scientific articles in a language they do not know. These questions should be addressed in future studies.

4.6. Understanding of Source Language

The third question in the survey related to the document that each respondent had submitted for translation was the following: *How well do you understand the language of the original written document (before it was translated)?* The possible answers were *Very well, Well, A little* and *Not at all.*

In the overall results, 51% of people reported that they understood *a little* of the source text and 33% said they understood the source text *well* or *very well*. By contrast, only 17% labeled their understanding as *not at all*. A few differences emerged when comparing the results of different language surveys. The Portuguese and English surveys had the highest percentage of people answering that their understanding of the source language was *not at all* (23% in English, 36% in Portuguese). In all other languages, 15% or fewer reported having no understanding.

As participants reported using PDF Translator for a variety of purposes, including dissemination, we conducted a separate analysis of people who specifically used it for assimilation, or gist users. For that analysis, we used only the answers of respondents who said their reason for translating the text was either that they wanted to understand it themselves or that they wanted to verify that they understood it themselves. As is shown in the following chart, a large majority of this specific group displayed at least a basic understanding of the source texts they translated. This result was similar to the overall results.

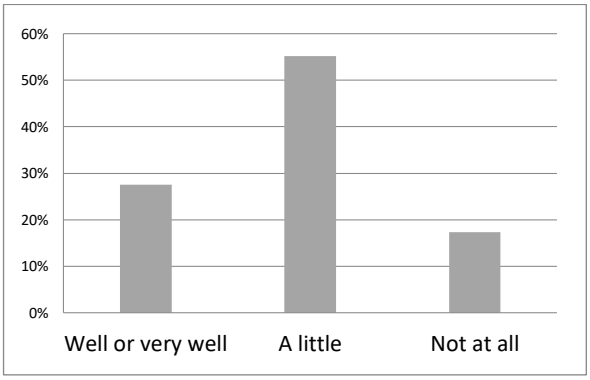

Figure 9. Reported understanding of the source text of the document submitted for translation by gist users.

The responses showed that in general, users of this tool often seem to translate texts that are in languages in which they already have some proficiency. Some previous survey studies have asked about users' competence in the source language, including Henisz-Dostert (1979), Hoshino (1995) and Yamada et al. (2005). A few other studies have uncovered indications of a link between knowledge of the source language and use of MT (Nurminen, 2016; Ogura et al., 2004).

Of course, people who are translating documents they already have some understanding of might also be simply testing PDF Translator or MT. Although we offered such people the choice to answer that their reason for translating the document was *some other reason*, some people may have instead indicated that their purpose was their own understanding and are therefore included in the assimilation group. In spite of this, there did appear to be a tendency to translate documents that respondents already had some understanding of, and this tendency has some interesting implications. First, this might be one

reason why, even with the onslaught of new language pairs available in online MT tools, the same European-based languages still tend to dominate the lists of the most translated languages. Because they are being taught widely in schools, these are languages in which people may have low-to-medium (although existing) competence.

Second, this could reflect a tendency to use MT with caution. Users want to be able to compare the machine translated text to the original so that they can evaluate the general level of MT output. This tendency might decrease in the future, as MT improves and users' trust in its quality increases.

Third, this raises a question that was asked in Henisz-Dostert's survey (1979): how do people find the texts they have machine translated? Do they need to have a basic understanding of the text (or even the title) to be able to make the decision to machine translate it? This would restrict the texts and the languages involved in gist MT use.

Finally, the phenomenon raises a question about how people use MT. Is MT in these cases being used as some type of language tool, which users can combine with other resources, such as their limited competence in the source language or their familiarity with the topic of the text, to gain understanding of a text in another language? If so, does this mean that the way people use gist MT (in raw or possibly also lightly post-edited form) is inherently different than the way they use publishing-level translations? Perhaps we need to begin seeing gist MT as a different translatorial activity than human translation, and to stop comparing them to each other.

5. Conclusions

This study provided a snapshot of the use and users of a specific type of gist MT tool. It presented a picture of who is using PDF Translator, where these users are, how they are using it, when they are using it, and in what areas of life they are using it.

The study confirmed some findings of previous studies. English continues to be the most-translated language and English-Spanish the most commonly translated language pair. However, it also showed that new languages such as Indonesian are beginning to appear at the top of lists of languages involved in MT. The demographics of the survey respondents indicate that, even though overall statistics reflect a bias toward young and male users, which is commonly found in technology studies, differences do emerge in the demographics of different language areas.

A few new tendencies that deserve further study surfaced also. First, gist MT users who translate whole documents seem to use MT often, multiple times a week. Second, the importance of MT in the area of study, at least for the current users of PDF Translator, was a noteworthy result. Finally, users' tendency to machine translate texts in a language that they have some level of proficiency in was a new finding.

Our study shares a limitation with a number of similar surveys in that it studied the users of only one tool and therefore cannot be considered representative of any larger or more general population of users. A second limitation was the use of a self-administered survey, which can lead to a disproportionately enthusiastic picture of MT users. A more random sampling of respondents could produce different results.

The study nevertheless contributes to the small body of literature on gist MT users. The main contribution is that that users' competence in the source language seems to play some role in their use of MT. Users' reports on having some level of proficiency in the source language of the document they translated, plus the tendency some users have to use MT not only for assimilation but also for verifying their understanding of documents, lead to questions of exactly how people are using gist MT. Is it comparable to their use of human translation, or do they use MT in very different ways?

Further studies on how people are using MT in their studies would be called for. We would also like to see new studies that focus on general populations of gist MT users, instead of the users of one tool. However, the most urgent need we envision right now is for deep, qualitative data on exactly how people use gist MT. After the first study in 1979, very little insight has been gained as to how people have integrated MT into their daily lives, what types of processes they use, and the cognitive processes they rely on to extract meaning from imperfect language. As the quality of MT improves and more uses are found for MT in its raw form, the already-pressing importance of this type of data will increase.

References

Bowker, Lynne, and Jairo Buitrago Ciro. 2015. Investigating the usefulness of machine translation for newcomers at the public library. *Translation and Interpreting Studies* 10 (2): 165-186.

Bowker, Lynne, and Melissa Ehgoetz. 2007. Exploring user acceptance of machine translation output: A recipient evaluation. In *Across boundaries: International perspectives on translation studies.* Eds. Dorothy Kenny & Kyongjoo Ryou, 209-224. Newcastle, UK: Cambridge Scholars Publishing.

Bowker, Lynne. 2009. Can machine translation meet the needs of official language minority communities in Canada? A recipient evaluation. *Linguistica Antverpiensia* 8: 123-55.

Castilho, Sheila, and Sharon O'Brien. 2017. Acceptability of machine-translated content: A multilanguage evaluation by translators and end-users. *Linguistica Antverpiensia* 16: 120-136.

Burgett, Will. 2015. Unmoderated remote usability testing of machine translation content. *TAUS Review of Language Business and Technology* IV.

Doherty, Stephen, and Sharon O'Brien. 2014. Assessing the usability of raw machine translated output: A user-centered study using eye tracking. *Intl. Journal of Human-Computer Interaction* 30: 40-51.

—— 2012. A user-based usability assessment of raw machine translated technical instructions. 10th Conference of the Association for Machine Translation in the Americas, San Diego, California, USA.

Fuji, M., N. Hatanaka, E. Ito, S. Kamei, H. Kumai, T. Sukehiro, T. Yoshimi, and H. Isahara. 2001. Evaluation method for determining groups of users who find MT “useful”. MT Summit VIII, 2001, Santiago de Compostela, Spain.

Flanagan, Mary. 1996. Two years online: Experiences, challenges and trends. Proceedings of the 2nd Conference of the Association for Machine Translation in the Americas "Expanding MT Horizons", Montreal, Canada.

Gaspari, Federico. 2007. The role of Online MT in Webpage Translation. Ph.D. thesis, University of Manchester.

——2006. The added value of free online MT services: Confidence boosters for linguistically-challenged internet users, a case study for the language pair Italian-English. 7th Conference of the Association for Machine Translation of the Americas, Cambridge, Massachusetts, USA.

—— 2004. Online MT services and real users' needs: An empirical usability evaluation. 6th Conference of the Association for Machine Translation in the Americas, AMTA 2014, Washington, DC, USA

Henisz-Dostert, Bozena. 1979. Users' evaluation of machine translation. *Machine Translation,* ed. Werner Winter, 149-244. The Hague: Mouton Publishers.

Hoshino, Sadao. 1995. Survey report: Korya eiwa: 100-dollar commodity has expanded the market. *Newsletter of the International Association for Machine Translation* (12).

Lawson, Veronica, and Muriel Vasconcellos. 1993. Forty ways to skin a cat: Users report on machine translation. Machine Translation Today: Translating and the Computer 15, London, UK.

Nuutila, Pertti. 2005. Rough Machine Translation in the Communication Process. Licentiate thesis, University of Tampere.

Nurminen, Mary. 2016. Machine translation-mediated interviewing in qualitative research: A pilot project. *New Horizons in Translation Research and Education* 4.

Ogura, Kentaro, Yoshihiko Hayashi, Saeko Nomura, and Toru Ishida. 2004. User adaptation in MT-mediated communication. The 1st International Joint Conference on Natural Language Processing (IJCNLP-04), Hainan Island, China.

Smith, Ross. 2003. Overview of PwC/Sytranet on-line MT facility. The 25th International Conference on Translating and the Computer, London.

Stewart, Osamuyimen, David Lubensky, Scott Macdonald, and Julie Marcotte. 2010. Using machine translation for the localization of electronic support content: Evaluating end-user satisfaction. Denver, CO, USA.

Turovsky, Barak. April 28, 2016. *Ten years of Google Translate* [Blog post]. Retrieved from https://blog.google/products/translate/ten-years-of-google-translate/.

Yamada, Setsuo, Syuuji Kodama, Taeko Matsuoka, Hiroshi Araki, Yoshiaki Murakami, Osamu Takano, and Yoshiyuki Sakamoto. 2005. A report on the machine translation market in Japan. The 10th Machine Translation Summit, Phuket, Thailand.

Yang, Jin, and Elke Lange. 2003. Going live on the internet. In *Computers and translation: A translator's guide,* ed. Harold Somers, 191-210. Amsterdam/Philadelphia: John Benjamins.