

DOI: <https://doi.org/10.24297/ijct.v19i0.8343>

Automated Translation Preference with a Bilingual Chatbot

Mina Park, Milam Aiken

¹School of Business, Southern Connecticut State University, New Haven, CT 06515²School of Business Administration, University of Mississippi, University, MS 38677

parkm2@southernct.edu maiken@bus.olemiss.edu

Abstract

Free, online language translation services are being used by people around the world to facilitate communication. However, it takes time and effort for a person to load the Web page in a browser, copy and paste the text into the site, and translate words. The process quickly becomes tiresome. Instead, some computer programs are providing an automated translation. However, no studies have been conducted to determine the efficiency or effectiveness of such an approach. In this study, we compare how students used an English-based chatbot with and without German automatic translation. Results show that students took nearly 1.5 times longer than their stated upper time limit to manually translate. In contrast, automated translation was at least 30 times faster. Also, the students were significantly more satisfied with the automated than the manual system.

Keywords: Chatbots, artificial conversation, Tutor Mike, machine translation, Google Translate, computer response time

Introduction

There are many instances of computer-based, foreign text needing to be translated, e.g., Web pages, email, documents, etc. For example, it is estimated that over a billion people use English on the Web, but over three billion use other languages [16]. In most cases, if users want to understand the text, they use a free, Web-based translation program rather than obtain the services of a human translator because of the relative ease of use and low cost.

Although the use of these online programs is free for small amounts of text, it still takes effort to copy and paste material with the translation service Web page. And, if several passages need to be translated, it can become exasperating, even though each transaction might take only a few seconds. If the process is perceived as too burdensome, the user might abandon the Web page or email message needing translation thinking the information contained is not worth the effort.

However, many programs are now incorporating automatic translation so that a user does not need to copy and paste the text into an online Web service. For example, a Web-based electronic meeting system uses Google to translate group comments [2], and Nolymit has developed a multilingual chatbot with the same automated translation service [19].

The purpose of this study is to determine how people perceive manual versus automatic translation with online Web services in conjunction with a chatbot. First, we discuss Google Translate and conversation agents. Then, we describe a new system that integrates the two types of software, eliminating the need for a user to process the text manually. The study consists of two parts: a response-time test and a chatbot satisfaction experiment. We conclude with directions for future research.

Google Translate

There are several free, online translation services, including:

- *PROMT* (<http://translation2.paralink.com/>)
- *Translator* (<https://www.bing.com/translator/>)
- *DayTranslations* (<https://www.daytranslations.com/freetranslation>)
- *DeepL* (<https://www.deepl.com/en/translator>)
- *Babylon NG* (<https://translation.babylon-software.com/>)
- *translator* (<http://translation.paralink.com/>)
- *Google Translate* (<https://translate.google.com/>)

Most people use *Google Translate* (GT), however, perhaps because *Google* is the most popular search engine with a market share of approximately 76% [3], and as of 2012, 200 million people used the free translation service each month [12]. Currently offering 103 languages in 10,506 combinations, the service is also fairly accurate, especially for Western European languages [1], though the accuracy can vary dramatically [15].

Chatbots

Chatbots or conversational agents are computer programs that enable people to communicate with the system naturally, communicating in sentences, as they would with other humans [4, 9, 11, 23]. Often, this software is used to provide a more intuitive interface for retrieval of information or learning [5, 6, 7], but it can also be used just for entertainment or provide companionship for lonely people [13, 20]. Many younger people have already interacted with chatbots and report positive experiences [21], and they may be more honest conversing with a chatbot than with other people because of its anonymous nature.

However, most chatbots support only a single language (typically English). A few chatbots support multiple languages [17], (e.g., Mondly supports 30 languages, Memrise supports 20, Watson supports 21, Eggbum supports 3), and as mentioned, Nolymit is integrated with Google Translate. Most people do not have access to these specialized systems, but multilingual support can be provided by using *Google Translate* (or another online translation service) in conjunction with a free, online chatbot. Users can copy and paste translations from their native languages into the conversational agents, but this is tedious. Alternatively, a program can be written that automatically links a high-performance chatbot with *GT* to provide the translation. We developed such a program to test the feasibility of multilingual artificial conversations for common use.

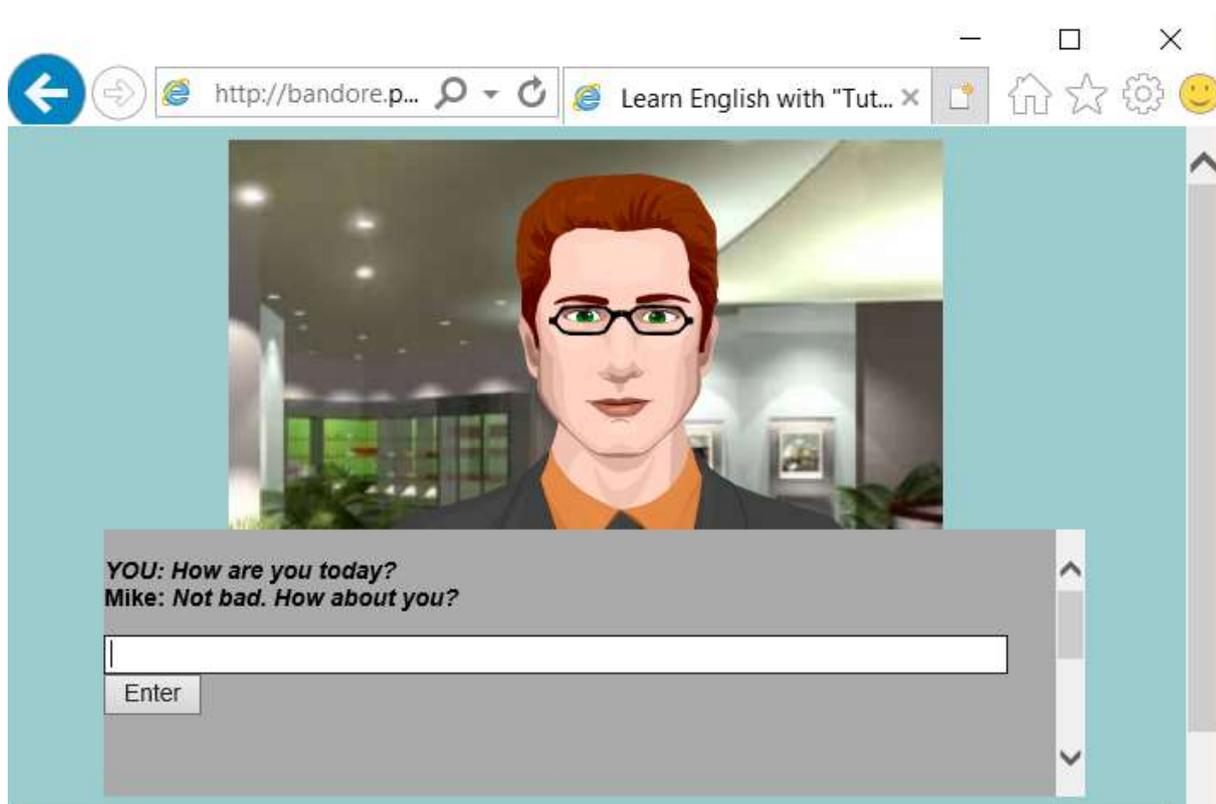
We chose *Tutor Mike* (<https://www.chatbots.org/chatbot/mike2/>) for our study because it is publicly available on the Web and according to the site, has achieved several honors demonstrating its ability to emulate human conversation:

- Second place in International Loebner Prize 2018 Contest in AI
- First place in International Loebner Prize 2018 Selection Contest
- Fifth place in International Loebner Prize 2017 Contest in AI
- Second place in International Loebner Prize 2016 Contest in AI
- Placed in semi-final of 2015 Chatterbox Challenge / one of the top four chatbots on the Web

- Fifth place in International Loebner Prize 2014 Contest in AI
- First place in International Loebner Prize 2013 Contest in AI
- First place in International Loebner Prize 2011 Contest in AI

Developed by Ron Lee, the goal of the chatbot is to support students learning English, serving as a conversational partner. The system can remember information the user types, can perform limited mathematical operations, and do some abstract reasoning. The program has also been trained extensively on several topics, including languages, cultures, geography, government, and history. Figure 1 illustrates the user interface of the system on the Web. The animated rendering of Mike's head serves to further the illusion that the user is talking to a real person rather than a computer program.

Figure 1. Tutor Mike Web Interface



Using a new hybrid system combining *Tutor Mike* with *GT*, a user types a comment in his or her native language in the top textbox and presses the 'Send' button. The text is translated by *GT* into English and sent to *Tutor Mike*, which generates a reply. The user presses the 'get a reply from chatbot' button to see the response in the original language or any other. That is, a person could type a comment in Japanese and receive an answer in French, for example, even though the chatbot's language is English. For this study, however, we limited the interface to German input and English output, as shown in Figure 2.

Methods

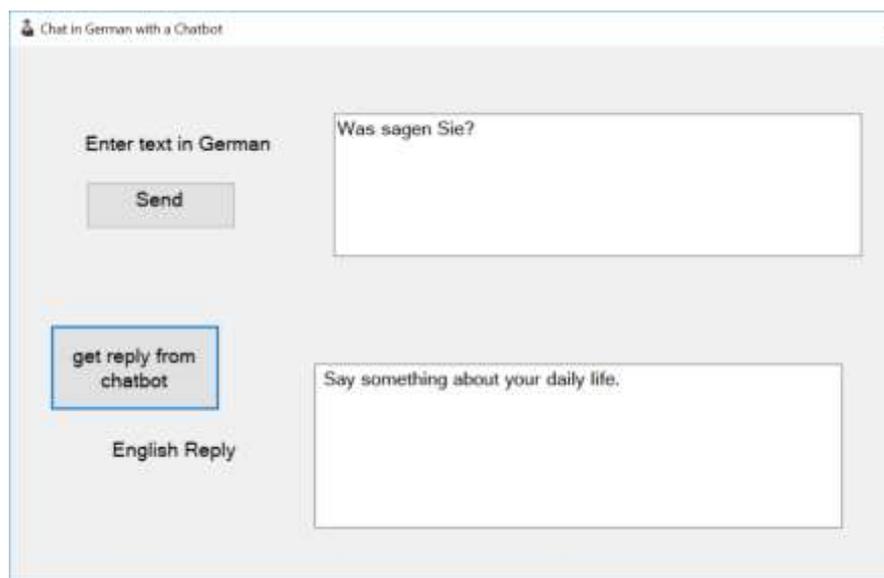
We wanted to determine how much of a burden copying and pasting text is for users of Web-based translation software in contrast with software that provides translations automatically. First, we conducted a study measuring how long it took users to copy and paste text; then we experimented with students comparing *Tutor Mike* with automated translation versus the same chatbot without the service.

Computer Response Study

We asked 58 undergraduate Business students (18 female) from a university in the northeastern region of the United States to participate in the study. None knew German, but all were proficient in English.

Fiona [8] reports that the tolerable waiting time for information retrieval is approximately 2 seconds, while Shneiderman [22] states that people are willing to wait 1 second for simple, frequent tasks, 2 to 4 seconds for common tasks, and 8 to 12 seconds for complex tasks.

Figure 2. Bilingual Chatbot with German Interface



Nielsen [18] states that 10 seconds is about the limit for keeping the user's attention focused on a dialogue such as might occur in an electronic meeting or a conversation with a computer chatbot. If it takes more than 10 seconds to send or read a message, users might abandon the program, and if they stay, they might rate their satisfaction with the program poorly [10].

We asked the students the first question shown in the Appendix. They stated that they would be willing to wait 9.6 seconds on average for a computer response such as a translation (min: 0.01, max: 60, std dev. 10.8. This was much longer than we were expecting they would be willing to wait, but it is not significantly different from the 10-second guideline mentioned by Nielsen ($p = 0.78$).

Students then used a program that presented three sentences in German sequentially. When each new sample of text appeared, the students went to the *Google Translate* Web site, copied and pasted the German text into the browser, and then copied the English translation from *GT* back to the testing software, as shown in Figure 3. The program recorded how long the complete transaction took.

The three German sentences and translations were:

German text: Wie Geht es Ihnen?

Google Translation: How are you?

German text: Meinen Sie dass, oder sagen Sie das nur so?

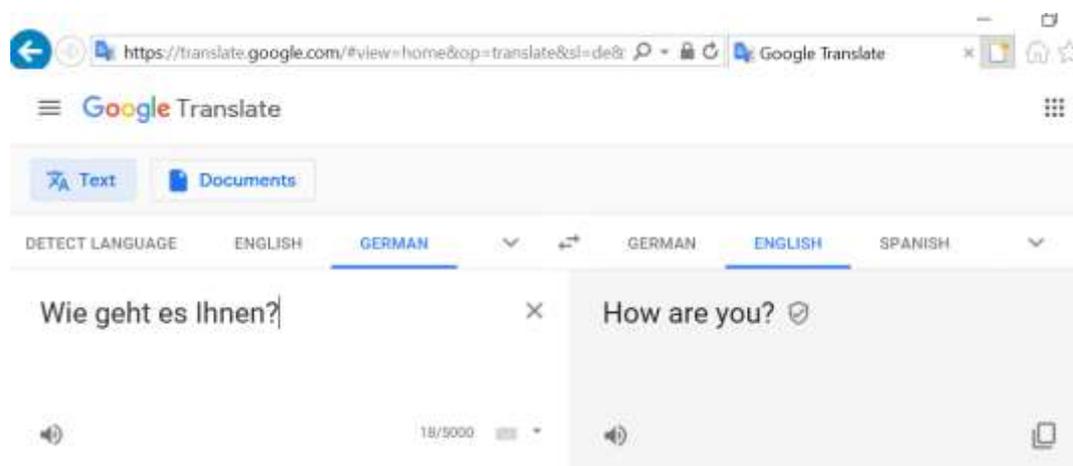
Google Translation: Do you mean that, or do you just say that? [A better translation is: Do you mean that, or are you just saying that?]

German text: Dieses Experiment ist jetzt vorbei.

Google Translation: This experiment is over now.

Results showed that the average roundtrip time for all three sentences was 25.5 seconds, (min: 8.9, max: 65.7, std dev: 13.7), significantly more time than they were willing to wait ($p < 0.001$). As expected, there was a practice or learning effect [14] as the time necessary to translate the third sentence was significantly less than it took for the first sentence ($p < 0.001$). The average time to translate, the first sentence was 41.5 seconds (min: 19.3, max: 65.3, std dev: 11.0), the average time for the second sentence was 19.6 seconds (min: 13.0, max: 44.8, std dev: 5.8), and the average time for the last sentence was 15.9 seconds (min: 8.9, max: 26.3, std dev: 3.2), a 61.7% reduction in time between the first and third translation. The time to translate the third sentence was still significantly longer ($p < 0.001$) than the students' were willing to wait, and we did not expect the time to decrease much more with subsequent translations.

Figure 3. Manual Translation with Google Translate



Bilingual Chatbot Study

Next, the students used *Tutor Mike* without automatic translation to converse in German, a language they did not know. We modified the chatbot so that only German input would be accepted, and an error message would appear if they tried to enter English text. That is, they had to think of something to say in their primary language of English, translate it to German with *Google Translate*, and then paste the German translation into the modified chatbot. We asked the students to enter four or five German comments in the five minutes allowed, and they were able to see the chatbot's responses in English, as shown in Figure 2.

Then, the students used a variation of the hybrid chatbot with automatic translation so that there was no need to copy and paste the text. With this system, they entered text and viewed responses in English, but the software also showed the automatic translations to German. Afterward, they answered questions 2 - 4 in the Appendix.

On average, the students rated the German system without automatic translation with a score of 3.93 on a 7-point scale (min: 1, max: 7, std dev 1.52), not significantly different from a neutral score of 4 ($p = 0.73$). They rated the English system with automatic translation as 6.33 (min: 2, max: 7, std dev 0.96), significantly higher than 4 ($p < 0.01$). Finally, they expressed a preference for the fully automated, English system with an average score of 6.43 (min: 1, max: 7, std dev 1.30), again, significantly above 4 ($p < 0.01$).

Discussion

Because of the extra time necessary to copy and paste text, we were expecting a lower score for the German system (non-automatic), but many students stated that they were impressed with the software and enjoyed chatting with it, perhaps influencing their satisfaction measure. Nevertheless, the system with automatic translation was rated significantly higher than the system without. Because the two chatbot systems were identical, we can attribute the different ratings solely to difficulty in translation. In a multilingual chatbot conversation without automatic translation, perhaps half the time might be consumed just by copying and pasting, clearly unacceptable. In this study, the average time to translate the last German sentence manually was about 16 seconds, 40% of the time it took for the first sentence, and 82% of the time for the second sentence due to the practice effect. We do not expect the translation time to decrease much more, and it probably won't go below the 10 seconds the students stated they were willing to wait. In contrast, automatic translation took less than 0.5 seconds. Also, in the manual process, there is the possibility that a person might copy text incorrectly, while the automated process does not make that error.

In general, however, outside the area of chatbots, the amount of time a user is willing to devote to translation depends upon the situation. If users expect a significant delay, i.e., they are warned, they might be willing to wait longer. Also, if the information to be obtained from a translation is important to the users, that might increase the amount of time they are willing to wait.

Conclusions

As people communicate more across the globe, there is a need for more translation. Currently, many are using free, online Web services to translate, but the process of copying and pasting text can be tedious. This study has shown that at least in the realm of multilingual chatbots, automated translation is far more efficient and satisfying for users. We believe similar results can be achieved by integrating automatic translation with other monolingual computer applications.

However, the study suffers from the limitation that only one language pair was tested (German and English). Other language pairs (e.g., Chinese and Hindi) might result in worse translations and consequently less satisfaction with the multilingual system. Further study should address this limitation and study uses with other software.

References

1. Aiken, M. and Balan, S. (2011). An analysis of Google Translate accuracy, *Translation Journal*, April, 16(2).
2. Aiken, M., and Park, M. (2009). Machine translation in a multilingual electronic meeting. *Journal of International Technology and Information Management*, 18(3/4), 395-408.
3. Chris, A. (2019). Top 10 search engines in the world. *Reliablesoft.net*, <https://www.reliablesoft.net/>
4. Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, 22(5), 811-817.
5. De Pietro, O., and Frontera, G. (2005). TutorBot: An application AIML-based for Web-Learning. In: *Advanced Technologies for Learning*.
6. Eynon, R., Davies, C., and Wilks, Y. (2009). The learning companion: An embodied conversational agent for learning. *Proceedings of the WebSci 2009: Society On-Line*.
7. Feng, D., Shaw, E., Kim, J., and Hovy, E. (2006). An intelligent discussion-bot for answering student queries in threaded discussions. In: *Proceedings of the International Conference on Intelligent User Interfaces IUI*.

8. Fion, F. (2004). A study on tolerable waiting time: How long are Web users willing to wait? *Behaviour & Information Technology*, 23(3), 153-163.
9. Floridi, L., Taddeo, M., and Turilli, M. (2009). Turing's imitation game: Still an impossible challenge for all machines and some judges—An evaluation of the 2008 Loebner Contest. *Minds and Machines*, 19(1), 145-150.
10. Hoxmeier, J., and Dicesare, C. (2000). System response time and user satisfaction: An experimental study of browser-based applications. *Proceedings of the Association of Information Systems Americas Conference*.
11. Hussain S., Ameri Sianaki O., and Ababneh, N. (2019) A survey on conversational agents/Chatbots classification and design techniques. In: Barolli L., Takizawa M., Xhafa F., Enokido T. (eds) *Web, Artificial Intelligence, and Network Applications. WAINA 2019. Advances in Intelligent Systems and Computing*, 927. Springer, Cham.
12. Kerr, D. (2012). Google Translate boasts 64 languages and 200M users. *CNET.COM*, April 26.
13. Matthews, K. (2018). Could chatbots make humans less lonely? *ChatbotNews*, September 27.
14. McCabe D., Langer K., Borod J., and Bender H. (2011). Practice Effects. In: Kreutzer J.S., DeLuca J., Caplan B. (eds) *Encyclopedia of Clinical Neuropsychology*. Springer, New York, NY.
15. McGuire, N. (2018). How accurate is Google Translate in 2019? *ARGO Translation*. www.argotrans.com
16. Miniwatts Marketing Group (2019). Internet world users by language. *Internet World Stats*, <https://www.internetworldstats.com/stats7.htm>
17. Mohammad, Z. (2018). Build a multilingual chatbots with Watson Language Translator & Watson Assistant. <https://medium.com/ibm-watson/>
18. Nielsen, J. (1993). *Usability Engineering*. Morgan Kaufmann.
19. Nolymit (2019). Nolymit's AI multilingual chatbot "InspireBot" gives you famous inspiring quotes based on your mood. <https://nolymit.com/Nolymit/multilingual-chatbot>
20. Pardes, A. (2018). The emotional chatbots are here to probe our feelings. *Wired*, January 31.
21. Park, M., Aiken, M., and Salvador, L. (2018). How do humans interact with chatbots?: An analysis of transcripts. *International Journal of Management and Information Technology*, 14, 3338-3350.
22. Shneiderman, B. (1998). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 3rd ed., Addison-Wesley, Reading, MA.
23. Zadrozny, W., Budzikowska, M., Chai, J., and Kambhatla, N. (2000). Natural language dialogue for personalized interaction. *Communications of the ACM*, 43(8), 116-120.

Conflicts of Interest. No conflict exists.

Funding Statement. No funding was provided.

Appendix

Please answer the following questions.

Sex M ___ F ___

1. How long are you willing to wait for the computer to do something, e.g., perform a translation for you?
_____ (seconds)

2. How satisfied were you with the system WITHOUT automatic translation (German program)?

1	2	3	4	5	6	7
Very			Neutral			Very
Dissatisfied						Satisfied

3. How satisfied were you with the system WITH automatic translation (English program)?

1	2	3	4	5	6	7
Very			Neutral			Very
Dissatisfied						Satisfied

4. Which system do you prefer?

1	2	3	4	5	6	7
Manual			Neutral		Automatic	
Translation						Translation
(German program)						(English program)