A Comparison of two Multilingual Meeting Systems

Milam Aiken, Mina Park

[1]School of Business Administration, University of Mississippi, University, MS 38677

[2]School of Business, Southern Connecticut State University, New Haven, CT 06515

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

Abstract

Multilingual meeting software is still somewhat rare, but at least three systems can provide support to large groups in multiple languages. This paper compares two (Microsoft Live Translator and an academic system called Polyglot). Results show there were no significant differences between the two in terms of perceived translation accuracy, usefulness, ease of use, and design quality. However, Polyglot provides support for more languages (103 versus 66). These and other results are discussed along with directions for future research.

Keywords: machine translation, multilingual meetings, Microsoft Live Translator, Google Translate

Introduction

The idea for multilingual meetings supported by automatic language translation was first proposed in the late 1980s (Gray and Olfman, 1989), and the first system was developed in 1991 using Spanish and English only (Aiken, et al., 1992). Several companies developed other multilingual meeting software in the following years. For example, Multicity.com introduced a system that translated text among six languages (English, French, German, Spanish, Portuguese, and Italian) (Lamont, 2000). Other systems, e.g., AmiChat, (Flournoy & Callison-Burch, 2001; Shigenobu, et al. 2007), Helpmate (Curran, 2002), and IM Translator (Smart Link Corporation, 2007), also provided multilingual support to meetings.

In 2008, researchers developed a Web-based meeting system based upon Google Translate called Polyglot (Aiken & Ghosh, 2009; Aiken & Vanjani, 2009). Later, Google developed Clickmeeting and Microsoft developed Live Translator that also supported multilingual meetings with completely automatic translations (Gottfried, et al., 2015). This paper describes these three in the following section, and then compares Polyglot and Live Translator for effectiveness and efficiency.

Current Multilingual Meeting Systems

Three Web-based multilingual meeting systems can be used to translate languages, and each product has unique characteristics.

1.   Clickmeeting (https://clickmeeting.com/tools/simultaneous-translation) Although developed by Google, Clickmeeting uses only about 52 of the 103 languages supported by Google Translate. The software detects the source language automatically (Figure 1), and both the original and translated chat are displayed immediately. A 30-day free trial is available, and the service costs US $40 per month afterwards. The website claims that about 150,000 customers use the system regularly, with 17 million in the last year in 111 countries in over 836,000 hours of meetings. However, it is unknown how many of these meetings are multilingual.

2.   Microsoft Live Translator (https://translator.microsoft.com/) This free service supports 66 languages, and Figure 2 illustrates a sample conversation with original content and translations shown in German. Using the system, a host initiates the meeting and provides a key code to the participants to join. Each member can either type or speak their comments to be translated and shared with others. Like Clickmeeting, Live Translator

does not provide complete anonymity to the meeting. Instead, users enter their names or pseudonyms, which are shown for each of their comments. Finally, the design is very modern emulating many instant messaging applications.

3.   Polyglot This system is a research prototype and is not freely available to the public for use. Based upon Google Translate, it provides automatic translation among all of its 103 languages in 10,506 combinations. All comments are completely anonymous. Although Clickmeeting and Live Translator have recorded transcripts, Polyglot also provides time stamps of when each comment was submitted, the language that was used, and other customizable options for researchers. Figure 3 shows a Lao speaker viewing public comments with the system.

Figure 1: Clickmeeting multilingual translation (English speaker's screen)

(source: https://clickmeeting.com/tools/simultaneous-translation#browse_all_features)

Figure 2: Live Translator multilingual translation (German speaker's screen)
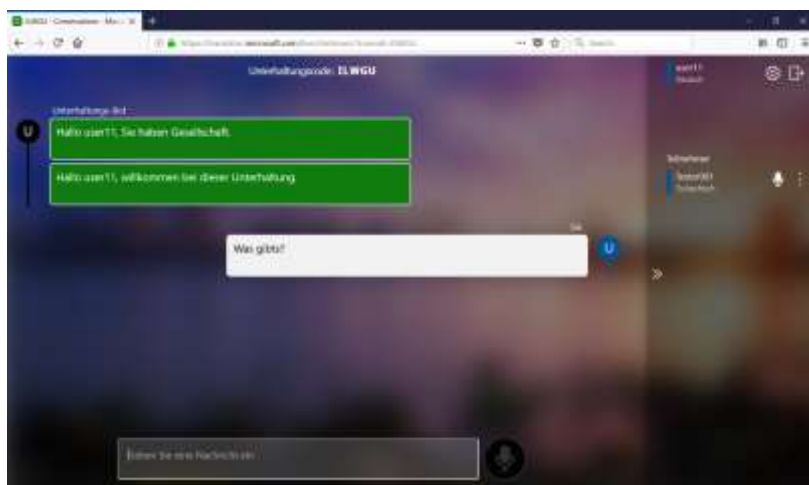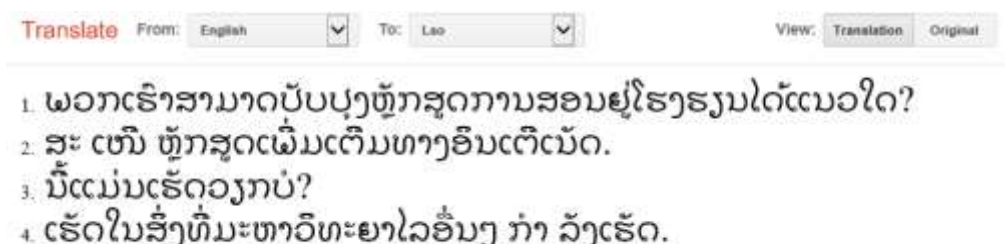

Figure 3: Polyglot multilingual translation (Lao speaker's screen)

Because of the limited 30-day trial for Clickmeeting, we chose to focus only on Polyglot and Live Translator. Table 1 shows a more detailed comparison between these two systems.

Table 1: Comparison of Live Translator and Polyglot

	Live Translator	Polyglot
Access	Free, publicly available	Free, but restricted access
Languages	66 – Microsoft	103 – Google Translate
Anonymity	no - names or pseudonyms	yes
Transcripts	optional	yes
Speech input	Optional 11 languages	no
Group size limit	500	unlimited
Time limit	4 hours	unlimited
Lock out users	optional	no
Prior comments viewable	no	yes

A Multilingual Meetings Comparison Study

We asked 67 students from a university in the northeast region of the United States to participate in a comparison study. One group of 21 students used Live Translator, and 46 used Polyglot. Of the total sample, 19 were female and 48 were male. The students typed comments in English about the topic "How can we improve the curriculum at this University?" for approximately five minutes, a time we thought was enough for the groups to learn the characteristics of each system. The meeting host also entered comments in Korean which were automatically translated to English for the group members to read. Because the text was anonymous (using Polyglot) and pseudonymous (using Live Translator), the participants could not be sure which comments were written in English (untranslated), and which were typed in Korean (translated to English).

After the meetings, the students answered the following questions using a scale of (1=disagree, 4=neutral, 7=agree).

1.   The translations from Korean to English were good enough for this conversation.

2.   This software was easy to use.

3.   I like the design of the user interface.

4.   I believe this software would be good for a multilingual meeting.

5.   My comments were anonymous to the other users.

Table 2 shows the results from the questionnaire. All means were significantly higher than the neutral value of 4 at α = 0.05, and there was no significant difference between the two systems on any measure. Finally, there was no significant difference on any variable based upon the sex of the respondent. Thus, users thought both systems were useful for multilingual meeting and easy to use. The systems also provided accurate translations from Korean (although the users were not sure which comments were written in that language) and good anonymity. The group members thought the design quality for both systems was good, although one user mentioned that Polyglot's design was "a little bland, but ok." Another participant using Live Translator stated that it was difficult to keep up with the many messages appearing on the screen. With this system, comments appear on the user's single screen as soon as they are written, possibly causing a distraction. Using Polyglot,

there are two different screens, one for writing text and another for viewing others' comments. One study (Calefato, et al., 2016) suggests that real-time machine translation is not disruptive of the conversation flow and, therefore, might be preferred by participants. However, with many participants in a multilingual meeting, the popup bubbles showing translations with Live Translator might become annoying.

For the Polyglot system, these results are comparable with an earlier study (Posey & Aiken, 2015) that found group members believed the meeting system was easy to use and the functionality was clear, and they learned how to use it quickly. In addition, they believed a multilingual meeting would benefit from the system, and they would be willing to use the system in such a meeting.

Table 2: Comparison of Live Translator and Polyglot

	All		Live Translator		Polyglot	
Question	**Mean**	**Std Dev**	**Mean**	**Std Dev**	**Mean**	**Std Dev**
Accuracy	6.58	0.81	6.67	0.78	6.54	0.83
Ease of use	6.76	0.85	6.86	0.35	6.71	0.99
Design quality	5.33	1.96	5.90	1.60	5.07	2.05
Usefulness	6.64	0.73	6.52	0.66	6.70	0.75
Anonymity	6.96	0.21	6.90	0.29	6.98	0.15

Table 3 shows strong, significant correlations between several variable pairs. As indicated, many pairs were highly positively correlated and statistically significant. Design quality and anonymity had no significant correlation, and design quality correlations with usefulness and anonymity were very low. Design quality and ease of use should naturally be positively correlated, as well as translation accuracy and usefulness.

Table 3: Variable correlation analysis (probability and significance)

	Ease of use	Design quality	Usefulness	Anonymity
Accuracy	0.438 < 0.001	0.260 0.031	0.580 < 0.001	0.645 < 0.001
Ease of use		0.456 < 0.001	0.431 < 0.001	0.615 < 0.001
Design quality			0.282 0.019	0.163 0.182
Useful				0.662 < 0.001

Polyglot and Live Translator Accuracy Evaluation

In one study of accuracy (MT-Qualifier, 2015), Live Translator achieved an accuracy score of 68.7 out of 100 and Google Translate 67.7 when translating from English to Spanish, and 63.3 versus 62.8 when translating from English to Portuguese. Another study (Shen, 2010) found that Google Translate was superior when translating long passages, but Live Translator often produced better translations for phrases below 140 characters. Papula

(2014) found Live Translator was better when translating English to Spanish using a point system, with a value of 68.4 to 67.3, but Google was better when translating Portuguese to Spanish (60.9 to 60.6). Finally, a comparison using seven languages translated in all combinations showed that Google Translate was superior with an average BLEU score (Papineni, et al., 2002) of 58.9 out of 100 versus 57.1, but the difference was statistically insignificant.

To investigate the differences in translation accuracy between the two systems more thoroughly, we obtained equivalent text from Omniglot (http://www.omniglot.com/) for the following sentences in six of the most-spoken, non-English languages in the world:

1. Pleased to meet you.

2. My hovercraft is full of eels.

3. One language is never enough.

4. I don't understand.

5. I love you.

6. All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.

BLEU scores were calculated using Tilde Custom Machine Translation's Interactive BLEU score evaluator (https://www.letsmt.eu/Bleu.aspx) with the 1-gram score option selected for the text translated to English. In addition, a human evaluated the translated text for comprehension, giving the scores shown in Table 4.

Table 4: Comparison of Google Translate and Live Translator (languages translated to English)

	Google Translate		Live Translator	
	BLEU	Comprehension	BLEU	Comprehension
Chinese	78	95	82	98
Hindustani	55	80	64	70
Spanish	80	98	86	98
Arabic	76	85	68	83
Malay	76	95	73	88
Russian	84	92	90	84
Average	74.8	90.8	77.2	86.8

There was no significant difference at $\alpha = 0.05$ in the BLEU ($p = 0.418$) or comprehension scores ($p = 0.113$) between the two translation services on average. Both systems translated Chinese and Spanish well, but they had problems with Hindustani.

For a further evaluation of accuracy, a native Korean speaker evaluated comments translated from English to Korean using Google Translate and determined that they were very accurate, but longer English sentences tended to be translated more correctly than shorter, incomplete sentences. Comments translated from Korean

to English had a few more errors. Live Translator was also very accurate when using Korean and English, even when mistakes were made in the source comment. As an example of the differences between the two systems, Korean comments entered by the host during the meetings are shown in Table 5 along with the English translations provided by Polyglot and Live Translator. It is interesting to note that most of the translations were worded differently, but usually had the same meaning.

Conclusion

At least three systems are available to support multilingual meetings: Google's Clickmeeting, Microsoft's Live Translator, and a research prototype called Polyglot. This study compared the latter two and found no significant differences in perceived design quality, ease of use, and usefulness in a bilingual meeting. A detailed examination of translation accuracy also showed no significant differences. However, Live Translator might appeal to more users because of its modern design, and it is freely available for use. Polyglot is restricted, but it supports more languages, provides complete anonymity, and has no time limit for the meeting.

Future research should investigate how Clickmeeting compares with these two systems. In addition, the question of group size limits should be explored with Live Translator. Although the software can support up to 500 users, such a large meeting might become unwieldy with constant messages popping up from other users.

Table 5: Examples of translations from Korean to English using both systems

Comments entered in Korean	Comments translated to English	
	Polyglot	**Live Translator**
커리큘럼에 관해서는 학생들의 의견을 들을 필요가 있어. 그런데 학생들이 의견을 제시 할 수 있는 기회가 없어.	I need to hear from students about the curriculum. But there is no opportunity for students to comment.	When it comes to the curriculum, you need to hear from students. But there's no opportunity for students to give their opinions.
온라인 강의를 선호하지 않는데 점점 더 많은 과목들이 온라인 강의로 바뀌고 있어. 온라인 강의를 좋아하는 학생들은 선택의 폭이 넓어지고 온라인 강의를 싫어하는 학생들은 선택의 폭이 줄어들고 있어.	I don't like online classes, but more and more subjects are changing to online classes. Students who love online classes have more choices, and students who don't like online classes have fewer choices.	I don't prefer online classes, but more and more subjects are turning into online courses. Students who like online teaching have more choice, and students who don't like online classes have fewer choices.
학사 학위를 받기 위해서는 4년 이상이 걸릴 수 도 있어. 4년 내에 학위를 끝내기 위해서는 여름학기나 겨울학기를 들어야 하는 경우도 있어.	It might take more than four years to get a bachelor's degree. You may have to take the summer or winter semester to complete your degree within four years.	It can take more than four years to get a bachelor's degree. In order to complete your degree within four years, you may need to take a summer or winter semester.

References

1. Aiken, M. and Ghosh, K. (2009). Automatic translation in multilingual business meetings. Industrial Management & Data Systems, 109(7), 916-925.

2. Aiken, M., Martin, J., Reithel, B., Shirani, A., and Singleton, T. (1992). Using a group decision support system for multicultural and multilingual communication. Proceedings of the 23rd Annual Meeting of the Decision Sciences Institute, November, San Francisco, CA, 792-794.

3. Aiken, M. and Vanjani, M. (2009). Polyglot: A multilingual group support system. Issues in Information Systems, 10(2), 101-106.

4. Calefato, F., Lanubile, F., Conte, T., and Prikladnicki, R. (2016). Assessing the impact of real-time machine translation on multilingual meetings in global software projects. Empirical Software Engineering, 21(3), 1002-1034.

5. Flournoy, R. and Callison-Burch, C. (2001). Secondary benefits of feedback and user interaction in machine translation tools. In Workshop paper for "MT2010: Towards a Roadmap for MT" of the MT, Summit, 8, 2-3.

6. Gottfried, J., DeLancey, L., Watwood, C., and Hardin, A. (2015). Virtual conferencing and meeting systems: Resources for online connections. College & Research Libraries News, 76(2), 98-101.

7. Gray, P., and Olfman, L. (1989). The user interface in group decision support systems. Decision Support Systems, 5(2), 119-128.

8. Lamont, I. (2000). Parlez-vous Multicity? Network World, June 26, 2000.

9. MT-Qualifier (2015). Machine Translation Quality Statistics. Retrieved January 13, 2020 from http://mt-quality.multilizer.com/machine-translation-quality-statistics

10. Papineni, K., Roukos, S., Ward, T., and Zhu, W. (2002). BLEU: A method for automatic evaluation of machine translation. ACL-2002: 40th Annual Meeting of the Association for Computational Linguistics, 311-318.

11. Papula, N. (2014). MT quality comparison: Google Translate vs. Microsoft Translator. Multilizer. Retrieved January 13, 2020 from http://translation-blog.multilizer.com/mt-quality-comparison-google-translate-vs-microsoft-translator

12. Posey, J. and Aiken, M. (2015). Large-scale, distributed, multilingual, electronic meetings: A pilot study of usability and comprehension. International Journal of Computers and Technology, 14(3), 5578-5585.

13. Shen, E. (2010). Comparison of online machine translation tools. Tcworld. Retrieved January 13, 2020 from http://www.tcworld.info/e-magazine/translation-and-localization/article/comparison-of-online-machine-translation-tools

14. Shigenobu, T., Fujii, K., and Yoshino, T. (2007). The role of annotation in intercultural communication. In International Conference on Usability and Internationalization, 186-195. Springer, Berlin, Heidelberg.