Usefulness and Usability of a Multilingual Electronic Meeting System

By William Pepper, Milam Aiken Ph.D., Bart Garner Ph.D.
University of Mississippi

Abstract- Communication in multilingual meetings is difficult, and interpersonal barriers increase with the number of languages used in the discussion. In this paper, we show how participants in pseudo-oral and electronic, multilingual meetings were able to comprehend previously prepared comments in up to five languages, but the electronic meetings took much less time. In addition, there were no significant differences between the two types of meetings in terms of ease of use or usefulness, which leads us to the conclusion that for certain sets of languages and topics, an electronic meeting with machine translation is a viable alternative to the more traditional, oral setting.

Keywords: Electronic meetings, group support systems, machine translation

GJCST Classification: JEL Code: O33
Usefulness and Usability of a Multilingual Electronic Meeting System

William Pepper, Milam Aiken, Bart Garner

Abstract - Communication in multilingual meetings is difficult, and interpersonal barriers increase with the number of languages used in the discussion. In this paper, we show how participants in pseudo-oral and electronic, multilingual meetings were able to comprehend previously prepared comments in up to five languages, but the electronic meetings took much less time. In addition, there were no significant differences between the two types of meetings in terms of ease of use or usefulness, which leads us to the conclusion that for certain sets of languages and topics, an electronic meeting with machine translation is a viable alternative to the more traditional, oral setting.

Keywords: Electronic meetings, group support systems, machine translation

I. Introduction

About one quarter of Americans can hold a conversation in a second language, and the incidence of multilingualism is even higher in some parts of the world, such as Europe (McComb, 2001). Yet, there is a growing need for meeting interpretation as international communication continues to increase (House & Rehbein, 2004). For example, the Directorate General for Interpretation at the United Nations provides interpreters for about 50-60 meetings per day, and meetings with up to 23 languages can be held with the aid of 69 interpreters (Fügen, et al., 2007). Interpreters and translators in the United States held about 31,000 jobs in 2004, and the demand has increased faster than the average for all occupations (Collegegrad, 2010).

In the past decade, technology has made the work of interpreters and translators easier (Sert & Açıkgöz, 2006). Now, people are using free, Web-based machine translation (MT) such as Google Translate and Yahoo!Babelfish to quickly obtain the gist of Web pages and email, and several multilingual Internet communities have arisen in which each participant communicates in his or her own native language (Yamashita & Ishida, 2006). In addition, groups are now engaging in face-to-face, multilingual discussions through electronic meeting systems that automatically translate among all the participants’ languages (Aiken, 2008).

However, much is unknown about how multilingual groups behave and how electronic meeting participants perceive machine translation. In particular, there has been no direct comparison of a traditional, human-interpreted, oral meeting with an equivalent, electronic discussion integrating MT. In this paper, we provide a background of oral and electronic multilingual meetings and then describe an experiment in which groups in simulated oral meetings are compared with groups in electronic discussions to determine the relative efficiency and effectiveness of each technique. The paper concludes with limitations and directions for future research.

II. Background

There are at least 11 million business meetings in the United States every day, and about 37% of the average employee’s time is spent in these sessions (Infocom, 2010). However, traditional, oral meetings have long been perceived as unproductive and unpleasant. For example, in one study (Wainhouse Research, 2010), 92% of attendees reported valuing meetings as providing opportunities to contribute, but 91% admitted to daydreaming during the discussions, and 39% have actually slept.

Electronic meeting systems have been developed to increase the productivity of group work. Using this technique, group members can exchange typed comments simultaneously and anonymously while all text is automatically recorded in a transcript, and numerous studies with this technology have demonstrated that people can generate more ideas in less time, participate more, and are more satisfied with it (Nunamaker, et al., 1991). However, the vast majority of research with electronic meetings has taken place using a single language, typically English, and multilingual groups have been relatively ignored in comparison (Aiken, et al., 2002; Briggs, et al., 1998). Traditional, oral, multilingual meetings have been supported by human interpreters who listen in one language and utter the equivalent in another, neither simultaneously or consecutively after the main speaker has verbalized a group of words or sentences. However, human interpretation can be expensive and fees have ranged from US $20 per hour (Ku & Flores, 2005), up to 300 to 400 Euros (US $423 to $564) per hour (Fügen, et al., 2001). Yet, there is a growing need for meeting systems that automatically translate among all the participants’ languages (Aiken, 2008).
In addition, scheduling an interpreter for a particular language pair (e.g., Vietnamese to German) at a specific time and location could be difficult.

In addition, human interpreters are not perfect, and errors rise steadily as time goes on (Al-Khanji, et al., 2000; Moser-Mercer, et al., 1998). For example, interpreter accuracy usually holds level for about 30 minutes, and then falls 10% for every 5 minutes afterward due to fatigue (Fügen, et al., 2007). As a result of these limitations, some researchers have investigated how electronic meetings can be used to assist groups that do not share a common language.

Although multilingual, electronic meeting systems were first developed in the early 1990s (Aiken, et al., 1994), only recently has there been a burst of growth in this field of research as MT quality has increased dramatically. Translation comprehension in these electronic discussions often suffers in comparison with that of oral meetings (Aiken & Ghosh, 2009), but the majority of MT users understand its limitations and are in general, impressed by the translation quality achieved (Yang & Lange, 2003). As a consequence, groups of up to 40 people at once using 40 different languages have used the technology successfully (Aiken, et al., 2010).

III. Theoretical Model

Many variables have been proposed over the years to measure an information system’s quality, including data currency, response time, turnaround time, data accuracy, reliability, completeness, system flexibility, and ease of use (Legris, 2003). However, no theoretical model has been specifically designed for evaluating a multilingual information system’s usefulness.

Prior studies of multilingual meetings have focused primarily on how well the group members comprehended the translations of comments, but few if any have compared this with the group’s required understanding (Aiken, 2008). That is, a measure of text comprehension is relatively worthless without some criteria for success, and required comprehension can vary based upon the importance, complexity, and urgency of the information (Aiken, et al., 2011). Thus, it is this relative comprehension success (or failure) which could influence a user’s perceptions of the meeting technique’s usefulness.

The time taken to translate or interpret meeting comments also affects the productivity and satisfaction of group members (Korth & Silberschatz, 1997). Even though machine translation of comments in an electronic meeting might be relatively worse than human interpretation in an oral meeting, the reduction in translation and comment submission time might offset the limitation of poor quality (O’Hagan & Ashworth, 2002). For example, because of the stress and time demands on human linguists in a meeting, some have suggested that an acceptable accuracy for interpretation is only 80%, while text translation needs to be at least 99% accurate (Fügen, et al., 2007). If translations are fast and accurate in a multilingual, electronic meeting, group members might be more likely to think the system is useful (Chuan-Chuan & Lu, 2000; Wixom & Todd, 2005).

A system’s ease of use is another factor that can influence its perceived usefulness (Segars & Grover, 1993). Oral meetings are natural and people are comfortable with speaking, but they might find the burden of taking turns and waiting for comments to be written on a board to be frustrating. On the other hand, in an electronic meeting, group members must learn how to use the software, and typing is less natural than speaking. But, if the meeting takes less time with acceptable translation quality, the technique might be perceived as easy to use and useful.

Based upon this prior research, we develop the theoretical model shown in Figure 1. Here, a comprehension difference is derived based upon group members’ required and actual comprehension, which in turn, influences the meeting technique’s perceived usefulness. Two other factors (translation time and ease of use) also affect this perception.

IV. Experiment

An experiment was conducted with small groups averaging about eight people, a size large
about 18 minutes to write the 35 sentences in the 5
groups about the meetings.

completed a questionnair e to express their feelin-
groups had more than 7 participants, some were idle.
member was assigned one comment to contribute
to electronic meetings because a previous  study
simulated, oral meetings were used in which the group
 facilitator simply wrote comments and translations on a
board at the front of the room, as if they were spoken
and then translated. Although these groups were only
simulated, they are designated ‘oral’ meetings
throughout the remainder of this paper. In the electronic
meetings, group members copied and pasted non-
English comments from a MS Word file into the Polyglot
user interface.

Seven random comments written in English
from a previous electronic meeting focused on ways to
solve the parking problem on campus were used:
1. Parking lots are too far from the business school.
2. Due to the parking problem, I was late to my morning
class.
3. Doctoral students should get preferred parking.
4. I wish the school could secure some space only for
PhD students.
5. The solution to the entire issue is to make everyone
ride bicycles.
6. We should prevent freshmen from parking on
campus.
7. Just assign me a parking place, and forget everyone
else.

These comments were translated, and groups
exchanged these comments in either three languages
(English, German, and Spanish) or five (English,
German, Italian, Spanish, and Swedish). The students
were randomly assigned to four electronic meetings
(sizes: 8, 7, 7, and 9), or six oral meetings (sizes: 9, 9, 8,
9, 9 and 10). In each electronic meeting, one group
member was assigned one comment to contribute
(translated to a foreign language). Because a few
groups had more than 7 participants, some were idle.

A total of five minutes was allocated for each of the
electronic meetings because a previous study
(Aiken, 2002) indicated that this time was more than
adequate for a person to contribute a comment with this
technique. After they submitted the comments, the
group members switched the language setting to
English and evaluated the automatic translations. In
each 3- language, oral meeting, the facilitator took about
11 minutes to write the 21 sentences on the board and
about 18 minutes to write the 35 sentences in the 5
language meetings. After meeting, the students
completed a questionnaire to express their feelin-
gs about the meetings.

Cronbach’s alpha values were calculated to test the
reliability of the two categorical variables, and ease of
use (0.911) and usefulness (0.935) each met the
minimum criterion of 0.7 (Cronbach, 1970; Nunnally,
1978).

Some of the means and standard deviations
(Mean/Std Dev) in the 3-language experiment varied
greatly between Electronic (E) and Oral (O). Examples of
this include: Actual comprehension - E (86.57/11.93) vs.
O (60.81/12.48); Required Comprehension – E
(80.28/5.52) vs. O (57.58/3.97); Compreh diff1 – E
(6.29/6.41) vs. O (3.23/6.51); Req. Comp(No Translation) – E
(60.28/32.58) vs. O (31.32/31.19); Req.
Comp (5 mins late) – E (85.22/17.74) vs. O (52.14 vs.
31.44).

Some of the 3-language experiments were very
similar in means between sessions. Examples include:
Ease of Use – E (4.87/1.06) vs. O (4.11/0.33);
Usefulness - E (4.75/1.06) vs. O (4.51/0.70); Time
preference – E (4.71/1.01) vs. O (4.57/0.28).

Similar results exist in the 5-language meetings for the
same measures.

There was a significant difference between the
electronic and oral meetings in regards to required
comprehension in the 3- language treatment. In
addition, there was a significant difference between the
electronic and oral meetings using three and five
languages in actual comprehension. Those in the oral
meetings indicated that they understood substantially
less; perhaps because they were confused by the text
written in so many different languages on the board.

Using three and five languages, the electronic
group comprehension was higher than the required
comprehension overall, required comprehension when
the alternative was no translation, and comprehension
when the alternative was information five minutes late.
However, the oral group members’ comprehension
using three languages was only slightly higher than the
three required comprehension criteria. Using five
languages, the oral groups’ comprehension was lower
than two of the required comprehension benchmarks.

Oral groups took significantly more time than the
electronic. Both types of groups wanted translations
quickly. In addition, there were no significant differences
between the oral and electronic groups or between the
3- and 5- language groups in terms of ease of use, and
there were no significant differences between the oral
electronic groups or between the 3- and 5-
language groups in terms of usefulness.

In addition to comparisons between the
meeting types, we conducted a correlation analysis to
investigate the relationships among the variables. Those
who comprehended the discussion more felt the
meeting technique was easy to use and useful, but

©2011 Global Journals Inc. (US)
surprisingly, they also required less comprehension in the discussion.

Students in the longer, oral meetings comprehended the discussions less, and required less understanding. Further, they thought the meeting technique was less easy to use and useful. As we expected, those who thought the meeting technique was easy to use also thought it was useful. Finally, students who expected translations quickly thought the technique was not easy to use or useful. This might be due to strong feelings among oral group members who were less satisfied with their technique.

VI. Conclusion

This research investigated simulated oral and electronic meetings using three or five languages. Results show that participants in the electronic groups were able to understand the foreign comments translated to English better than those in the simulated-oral groups, and better than that required. Oral group comprehension was lower than that required in some cases.

With five languages, the oral meeting technique was not perceived to be easy to use, but the electronic technique was easy to use and useful in all cases. Finally, the electronic meetings took significantly less time. Therefore, we conclude that for these groups, languages, and topic, electronic, multilingual meetings can be used effectively and efficiently.

The first limitation of this study is the fact that only a small subset of European languages was used (English, German, Italian, Spanish, and Swedish), and students evaluated only the translations to English. Comprehension of translations between other languages (e.g., Croatian to Chinese) could be different.

Second, group members in the pseudo oral meetings did not actually say anything because the text needed to be identical between the treatments and participants might have behaved differently in this simulated environment.

Fourth, some results could have been affected by members’ dissatisfaction with the overall process. For example, oral-group members’ reported comprehension was lower than expected, perhaps because of their frustration with the long meeting time.

Finally, only one facilitator was used to write comments on the board in the simulated oral meetings. More facilitators writing simultaneously in different languages could reduce the amount of time needed in these meetings, but there also could be more confusion.

Future research should focus on a comparison of electronic groups with actual, oral groups using a variety of languages and topics to determine in which cases the technology is most beneficial.

References Références Referencias
