



Effect of Cultural Misunderstanding Warning in MT-Mediated Communication

Mondheera Pituxcoosuvarn¹(✉) , Yohei Murakami¹ , Donghui Lin² ,
and Toru Ishida³

¹ Faculty of Information Science and Technology,
Ritsumeikan University, Shiga, Japan
{mond-p,yohei}@fc.ritsumei.ac.jp

² Department of Social Informatics, Kyoto University, Kyoto, Japan
lindh@i.kyoto-u.ac.jp

³ School of Creative Science and Engineering, Waseda University, Tokyo, Japan
toru.ishida@aoni.waseda.jp

Abstract. Thanks to today's technologies, the world's borders have been fading away and intercultural collaboration has become easier and easier. Language and cultural differences are common problems in intercultural collaboration. Machine translation (MT) is now available to overcome the language barrier, so people can easily express and understand messages in different languages. However, misunderstandings often plague users from different cultures, especially in MT-mediated communication. To communicate productively, it is important to avoid such misunderstandings. One existing work proposed the idea of using automated cultural difference detection to warn the users of misunderstanding. However, no study has examined how such warnings affect the communication. To eliminate this gap, we conduct a controlled experiment on how users react to the warnings and what are the results in terms of communication. The results show that, with the data from cultural difference detection, warning the user of cultural misunderstanding can help reduce misunderstandings and increase awareness of cultural differences. The results of this experiment confirm the effectiveness of cultural misunderstanding alerts and suggest new directions in multilingual chat design.

Keywords: Intercultural collaboration · Machine translation · Cultural misunderstanding

1 Introduction

Given the advances in transportation and technology, we have more chance to communicate across cultures than before. Intercultural collaboration and cultural diversity provide societies with vast benefits [5]. Nevertheless, communication is challenged by many difficulties. In the past, people needed to learn

a foreign language or needed an interpreter to communicate smoothly across languages. Now, communication has been made easier through the support of machine translation (MT). There are various tools and services available to choose from. MT can be easily used by general users without any expert knowledge to translate documents, conversations, and messages. It has also been embedded in chat systems so that users without a shared language can communicate. Moreover, there are various web services that can be used by both general users with more technical knowledge to create their own resource [8].

However; MT is still not perfect and it can cause various difficulties, for example, misunderstanding due to mistranslation, conversation breakdown [13], and gaps in mutual comprehension [16].

Some difficulties, i.e. mistranslation, can be solved by improving MT quality. Even with improvements in quality, there are some situations where MT output hinders successful communication. For instance, a group of researchers [14] conducted a field study at a children’s workshop where the children used an MT-embedded chat system communicate. They reported that communication became difficult when an adult facilitator showed a block of brown play dough to the children and asked “what does this looks like?”. A Japanese participant answered it looks liked ‘*餡子 (Anko)*’ which is can be translated as ‘red bean paste’. The children from different cultures did not understand reference made by the Japanese participant. Later they used image browser to find pictures of *Anko* and they came to understand that it is a block of stiff red bean paste. Even with perfect MT quality, this kind of cultural problem still occurs and creates a barrier to achieving mutual understanding. For effective collaboration, it is important to establish mutual understanding but the current MT-embedded chat systems sometimes fail and actually cause cultural-based misunderstanding. Based on this field study, one study proposed automated cultural difference detection [15]. Their method detects cultural difference by comparing images in databases linked to each language. They also suggested that the result of cultural difference detection be used to warn the users of possible cultural differences. However, the impact of warning the users of these differences was not confirmed.

To fill this gap, we conduct a controlled experiment based on our research question: *how warning the user of possible cultural differences and cultural misunderstandings can affect MT-based communication?*. Our hypothesis is that warning the user of cultural misunderstanding will significantly help the user in reducing misunderstanding and thus support mutual understanding. We designed a collaborative task and asked our participants to complete the task together by chatting on an MT embedded chat system with cultural misunderstanding warnings for the experimental group and without warning for the controlled group. We interviewed the participants to find out if each participant understand correctly or not, then conduct a t-test to examine if there is a significant difference between the experimental group’s understanding and the controlled group’s understanding.

In the next section, we introduce studies related to our work. Section 3 reviews a key component of our experiment, a key method to detect cultural

differences. Next, Sect. 4 details our experiment. The results of the experiment are shown in Sect. 5 and discussed in Sect. 6, which is followed by our conclusion of this paper.

2 Related Work

2.1 Misunderstanding in Intercultural Collaboration

Because people with different language backgrounds sometimes perceive things differently [3], misunderstanding can readily occur in intercultural collaboration. Because of this, many studies have tackled cultural misunderstanding.

Grounding a conversation or establishing mutual understanding is difficult, especially when communication is carried out via a chat system or MT. Yamashita et al. [16] studied why and how conversation grounding is problematic in MT-mediated communication. Their experiment found three problems. First, the users were not aware of which conversation content was or was not being shared. Second, the users were not aware which concepts they could or could not share with others. Third, users faced difficulties in constructing efficient utterances when using MT-mediated communication because of the first problem.

2.2 Cultural Difference Identification and Detection

In order to prevent misunderstanding, it is necessary to be able to detect it. Various works have tackled detecting and identifying cultural differences. Most studies collected and analyzed data from cross-national surveys. One of the most well-known works is Hofstede's cultural dimension [6]. He identified cultural differences in different regions. Yoshino et al. [17] also conducted a cross-national survey but compared some aspects of culture, such as social values and ways of thinking. The results from surveys are interesting, however, it is difficult to apply them to computer-mediated communication.

Other researchers have worked on cultural differences related to computer-mediated communication. In 2007, Cho et al. [2] published a study on the cultural differences found in pictogram interpretations. They conducted a web survey to understand the differences in pictogram interpretations between Japanese and Americans. Their report found that 19 of 120 pictograms were judged to have cultural differences.

Later, Yoshino et al. [18] proposed a method for cultural difference detection in Wikipedia. Japanese students and Chinese students were asked to examine words and phrases with different meanings and the results were used to create an initial dataset. Based on the dataset, they proposed a process for judging whether cultural differences existed or not in certain words or phrases.

Yet, the cultural difference detection methods mentioned above cover only specific areas and usages, i.e. pictograms, Wikipedia and all require human intervention. In 2019, a group of researchers [15] proposed a method to automatically

detect cultural differences in words when they were translated into another language. This method can be applied to various languages and can cover broad area, as long as there lexical databases and image libraries are available. The authors also proposed that detection results can be used to warn the users of potential cultural differences. Base on this automated detection, Nishimura et al. [12] proposed a method and conducted an experiment to find the threshold that serves as a basis for confirming cultural difference.

3 Cultural Difference Detection (CDD)

To prevent and warn users of cultural misunderstanding in MT-mediated communication, it is important to detect possible misunderstandings. This work adopts a method from our previous work [15] that can automatically detect the words that might cause misunderstanding when they are used and translated into another language. This section briefly reviews how cultural difference detection (CDD) works.

To investigate if using a word W in language L_1 could cause misunderstanding due to cultural difference when it is translated into language L_2 , the following procedure should be performed.

1. Translate word WL_1 (language L_1) into WL_2 (language L_2).
2. Search for images using WL_1 and WL_2 as *keywords*.
3. Extract image vector features of each image.
4. Compare the two images by computing their vector features.
5. If the similarity is low, the possibility of misunderstanding is high.

To apply this CDD concept, several variables must be considered, including, language (word) resource, number of images for each keyword, tools for feature extraction and comparison.

Here is an example of finding a list of words that have high possibility of causing misunderstanding when they are used in multilingual communication between Japanese and English.

Base on Fig. 1, first, from Japanese WordNet [7] which is a Japanese-English lexical database created from the original English WordNet [11], a *synset* is selected. A *synset* in Japanese Wordnet is a set of synonym containing words in English and Japanese under the same concept with similar meaning. Here, in Fig. 1 the *synset* randomly selected is the synset that contains *william_cowper* and *cowper* in English and *クーパー* (*Kuupaa*) in Japanese. Then search for 30 images for each language: 15 images for *william_cowper*, 15 images for *cowper* and 30 images for *クーパー* (*Kuupaa*). Next the vector features of the images are extracted. Both the original paper and our paper used VGG16¹ from Keras. The feature values of each language are averaged. Next, averaged vector features of each language are compared, (here we use Cosine similarity).

A list of words that might cause misunderstanding can be made by repeating this process a few thousand times or more. Words whose similarity is lower than

¹ <https://keras.io/applications/#vgg16>.

a threshold, the original work suggested 0.6, are entered in a list and the user is warned when they use a word present in the list.

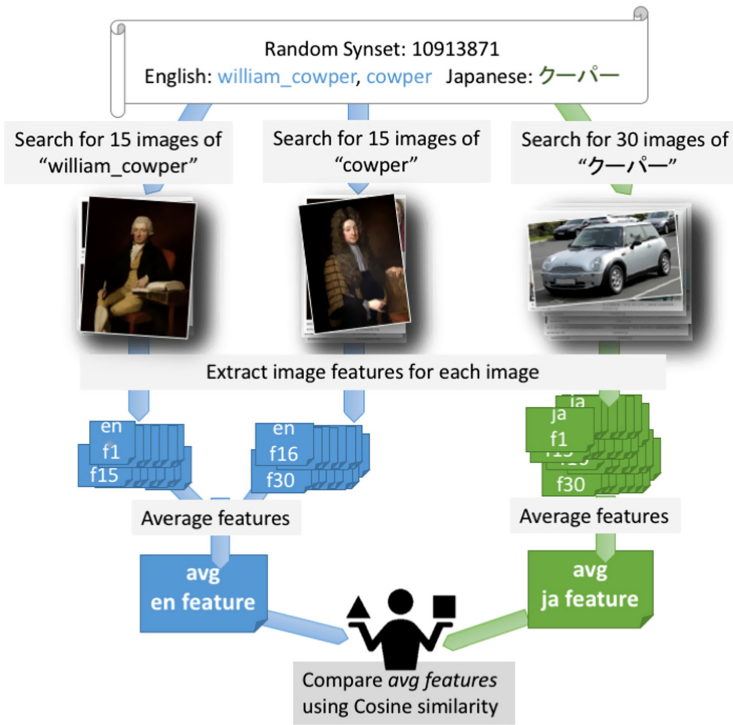


Fig. 1. An example of a process based on CDD

4 Experiment

To study how cultural misunderstanding warnings might impact communication we conducted an controlled experiment. Our hypothesis is that users who are warned of cultural-misunderstanding achieve better communication when using MT-mediated communication.

4.1 Participant

We asked 18 volunteers with various language and cultural back-grounds to participate in MT-mediated conversations. The participants were in their 20s and 30s. They were separated into six groups of three people. Details are as follows:

- A native Japanese or a person with native-level Japanese who has been educated in Japan or currently lives in Japan.
- A native Chinese or a person with native-level Chinese who has been educated in China or currently lives in China or a Chinese speaking environment.
- A native English speaker or a person with native-level English who has been educated in England or currently lives in an English speaking environment.

The six groups of participants were divided into three experimental groups (*E1*, *E2*, and *E3*) and three control groups (*C1*, *C2*, and *C3*).

4.2 Communication Tool

The tool used in this experiment was a web application developed around translation services from the Language Grid [8]. This application is an MT-embedded chat system that allows users to communicate in their preferred language. When a user logs-in to the application using a given link, he/she can select her/his preferred language on the right-top of the page. If a user chooses to chat in English, all the message from the other users, who might be accessing the system in different languages, are shown in English. When he/she enters and sends messages in English, the other users will see those message in their selected languages.

4.3 Task Design

Designing the task given the participants was challenging. In normal chat conversation or normal collaborative tasks, there is no guarantee that a cultural-misunderstanding will occur. To test our hypothesis, we designed a game that led the users to communicate using words that might cause misunderstanding. Because this game was designed only to create and lead the conversation, there is no evaluation of the game result nor the correct answer.

Our game was inspired by the Desert Survival Problem (DSP) [10]. DSP is widely used in team building and collaboration practice. The conventional DSP asks the players to collaboratively rank items by its important to their survival in the desert. Many variations of the game can be created by giving different situations and items. We create our variation and indirectly force the user to talk about things that might easily be misunderstood by adding words in the CCD-derived list to the item choices. To encourage participation in the conversation, the choices given to each group member were different, so everybody had to speak up and share. Every group member was given three choices; they were instructed to share and collaboratively select one of the most important choices from each list. Only three of the nine choices could be chosen. The collaboration ended when all members agreed on the three choices.

Examples of the lists given to the participants are shown in Fig. 2. Words that could cause misunderstanding were emphasized (red underlined) for clarity. For the control group, the lists given to them did not contain any text emphasis.

The list in Japanese translated into English reads

You found a train container that has not been destroyed yet. You can take one of the listed items.

Chinese Choices	Japanese Choices	English Choices
<p>后来在车祸中，你们找到了另外三个乘客活了下里。在他们3人中，你只可以帮一个。剩下的两个会跟着另外一组，向南去试一试。这3个人当中，你不懂他们的工作或拿手的艺术。你唯一的方法是看着他们穿着的衣服和形象。</p> <ol style="list-style-type: none"> 1. <u>修行者</u> 2. 青少年 3. 电气技师 <div style="border: 1px solid black; padding: 5px; width: fit-content;"> <p><i>Translation</i></p> <ol style="list-style-type: none"> 1. A religious practitioner 2. A teenager 3. An electrician </div>	<p>あなたはまだ壊れていない電車のコンテナを見つけた。以下の中から一つだけ手に入れることができる。</p> <ol style="list-style-type: none"> 1. 2台の<u>ママチャリ</u> 2. 2つの<u>花冠</u> 3. 一本の刀 <div style="border: 1px solid black; padding: 5px; width: fit-content;"> <p><i>Translation</i></p> <ol style="list-style-type: none"> 1. Two bicycles 2. Two corollas 3. A sword </div>	<p>While you are walking, you found a cute plastic box with the following items in it, you can use one from the list.</p> <ol style="list-style-type: none"> 1. A <u>mud pie</u> 2. Three cans of <u>pop</u> 3. One pack of family size crackers

Fig. 2. Lists given to participants in each language. (Color figure online)

1. Two bicycles (ママチャリ - Mamachari)
2. Two corollas (花冠 - Kakan)
3. A sword

The list in Chinese can be translated into English as
You found the other three people who took the same train. They are alive but wounded by the wreck. You can choose to help one person and take that person with you. The rest will be helped and carried by the other group who are heading to the south. You don't know them, but you can guess who they are from how they look and dress.

1. A religious practitioner (修行者 -Xiuxing Zhe)
2. A young teenager
3. An electrician

The underlined choices are expected to cause cultural misunderstanding. Most underlined words were taken from the list of cultural differences made using CDD; some words were added manually to create difficulties, including ママチャリ (*Mamachari - Bicycle*) and 修行者 (*Xiuxing Zhe - Religious practitioner*). The word 修行者 (*Xiuxing Zhe*) was detected when the CDD was run using Japanese Wordnet, as a Japanese word with Chinese character (Kanji); this word also exists in Chinese language but we did not run CDD on any Chinese resource.

4.4 Expectation

We expected that the participants would have communication difficulties when using the underlined words in Fig. 2 because of the cultural differences and translation problems, if not warned.²

² Note that the translation output is likely to change from time to time since MT is always developing and the output also depends on the services selected.

The choices exhibited significant cultural differences and thus problems in understanding. First of all, on the English list, mud pie is definitely problematic because there are two meanings. The original meaning is a pie made of mud by children. The other meaning emerged later as an edible pie that resembled a mud pie. People from different cultures might not be sure if mud pie is edible or not, regardless of translation. The second choice on English list, *pop*, is a slang but well known for a carbonated drink. If MT does not know the content of conversation, it might translate pop into different word, such as “pop music”.

On Japanese list, ママチャリ (*Mamachari*) is a word that not only means bicycle but also information about size and use. It is possible to carry children since it usually has space for luggage or child seats. The list choice was two bicycles since the team of three could fit onto 2 Japanese style bicycles. People who did not understand ママチャリ (*Mamachari*) would not know this fact.

For, 花冠 (*Kakan*), the MT output was “corolla” which is a wreath. This usage is very archaic and rarely used nowadays. In many regions, corolla is recognized as a car since it is a famous car model.

We gave them two bicycle for the team of three which is enough to ride. People who does not understand ママチャリ (*Mamachari*) would not know this information. For 花冠 (*corolla*), the translation in English is “corolla” which is a headgear, however not popularly used nowadays. In many regions, *corolla* is recognized as a car since it is a famous car model.

On Chinese list, 修行者 (*Xiuxing Zhe - Religious practitioner*) is the most difficult to explain. MT usually output the English word “practitioner” which most people understand to be a medical doctor. However, it actually means a religious practitioner or a monk who often goes on pilgrimages and so might be useful in helping the group to survive since he has experience in traveling.

4.5 Method

We conducted the experiment using the *Wizard of Oz* [9] technique which is often used in human-computer interaction studies. In our experiment on how communication is effected by the warning, the *Wizard* is a human who warns the participant instead of the computer system. The experiment group members were given the situation and their choices with the suspected words indicated by red underlining while the control group members were given the same situation and choices without any emphasis.

To evaluate the effect of warnings on communication, after the collaborative task, we asked each participant to explain the six choices the other participants had and recorded how many choices each participant actually understood. Then a t-test was conducted to examine the significance between two independent samples including the percentage of understanding from the experimental group and from the controlled group.

4.6 A Preliminary Experiment on Number of Languages Used in MT-Embedded Chat

Besides the main experiment, we also designed a preliminary experiment to study if number of languages used impacted the participants' understanding of the choices. To conduct this preliminary experiment, we instructed experimental group *E1* to communicate using only two languages: English and Japanese. In this case, the Chinese speaker who was also fluent in English used English to communicate, but the given choices were written in Chinese, as in the main experiment.

5 Result

5.1 Cultural Misunderstanding

After asking the participants to explain the choices the others had been given, and quantitatively analyzing the chat log, we divided participant understanding of the six other items into three groups.

- *U*: The user understood right after the choice was first mentioned
- *L*: The user understood after the choice was introduced but before the game ended
- *M*: The user could not understand or misunderstood the choice

The results are shown in Table 1. Asterisk marks are used to indicate incomplete understanding of the detailed characteristics of the choice, for example, knowing pop is a drink but not that it is carbonated, and knowing that ママチャリ (*Mamachari*) is a bicycle but not that it is often used by mother so it often has enough space to carry things or has extra seat(s) for kid(s). In addition this table also displays the number of turns to show how many turns among three users were taken to complete the task. Time of interruption shows the number of times when the game flow was interrupted by questions about the choices the participants wanted to confirm or could not understand. However, in this experiment, there is no correlation between the time of interruption and understanding (t-test p -value = 1).

By the end of the game, all the experimental groups had successfully established mutual understanding. They successfully shared and understood all the given choices. On the other hand, none of control group successfully shared or understood those choices.

The experiment showed that when the being choice introduced did not cause misunderstanding, usually the other participants could easily understand it right away (tagged *U*). If the participant felt that the word was difficult to understand, usually someone would ask for an explanation which would allow the group members to finally understand the choice. The words from the CCD list were frequently misunderstood, especially by members of the control group.

Figure 3 shows the rate at which each choice was understood by each participant. The percentage is calculated by summing the choices tagged *U* and *L*

Table 1. Understanding result of each participant in each group

Group	Participant	English choice 1 (mud pie)	English choice 2 (pop)	English choice 3 (cracker)	Japanese choice 1 (bicycle)	Japanese choice 2 (corolla)	Japanese choice 3 (sword)	Chinese choice 1 (religious practitioner)	Chinese choice 2 (teenager)	Chinese choice 3 (electrician)	Number of turn	Time of interruption during the explanation	Mutual understanding established
Experiment E1**	English speaker	-	-	U	U	U	U	U	U	U	60	5	Yes
	Japanese speaker	U	U	-	-	-	-	U	U	U			
	Chinese speaker	U	U	U	U	U	U	-	-	-			
	English speaker	-	-	U*	L	U	U	L	U	U	43	5	Yes
Experiment E2	Japanese speaker	U	U	U	-	-	-	L	U	U			
	Chinese speaker	U	U	U	U	L	U	-	-	-			
	English speaker	-	-	U*	U	U	U	L	U	U	48	1	Yes
	Japanese speaker	U	U	U	-	-	-	L	U	U			
Experiment E3	Chinese speaker	U	U	U	U*	U	U	-	-	-			
	English speaker	-	-	U	U	M	U	M	U	U	33	3	No
	Japanese speaker	M	L	U	-	-	-	M	U	U			
	Chinese speaker	M	L	U	U	M	U	-	-	-			
Control	English speaker	-	-	U	U	M	U	M	U	U	49	6	No
	Japanese speaker	L	L	U	L*	M	U	M	U	U			
	Chinese speaker	L	L	U	-	-	-	M	U	U			
	English speaker	L	L	U	L*	M	U	-	-	-	54	2	No
Control	English speaker	-	-	U	U	U	U	M	M	U			
	Japanese speaker	M	U	U	-	-	-	M	M	U			
	Chinese speaker	M	U	U	U*	U	U	-	-	-			
	English speaker	U	U	U	U*	U	U	-	-	-			

Note

U The user understood right after the choice was introduced.

L The user understands later, after the choice was introduced but before the game ended.

M The user could not understand or misunderstood the choice.

*The user understood but not completely. The detailed characteristic of the choice was missed.

**This group used 2 languages in the experiment.

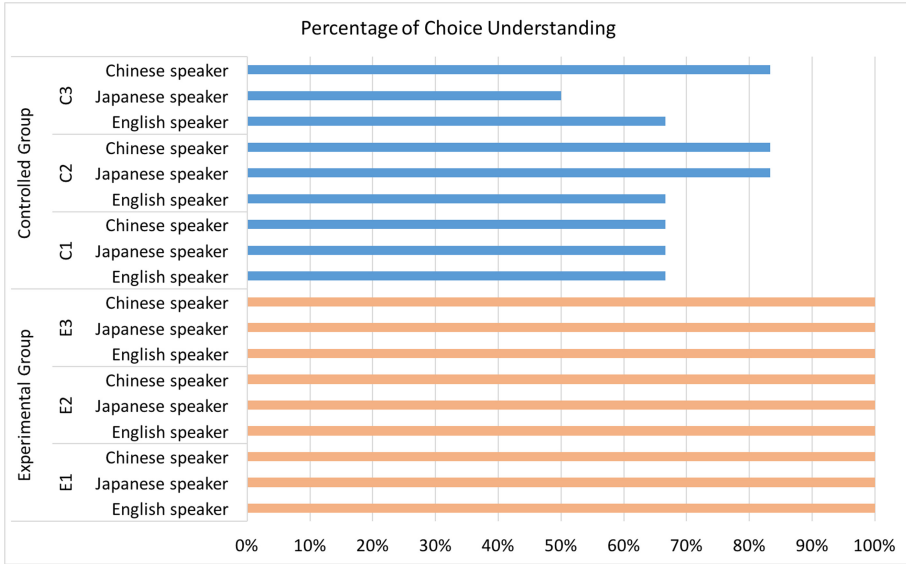


Fig. 3. Percentage of understanding of choice shared by the other group members

and dividing by six, the number of choices introduced by the other participants. The graph shows that the experimental group had full understanding (100%) by the end of the game while the control group had less understanding (average of 70%). Warning the users of cultural misunderstanding significantly improved understanding in MT-mediated communication, especially when using words that might cause misunderstanding. We conducted a t-test with independent samples, including the percentage of the correct understanding of the experimental group and the percentage of the correct understanding of the controlled group. From the test, with p -value equals to $5.54545E-07$, the null hypothesis is strongly rejected as $p < 0.001$ and we conclude that warning the user of cultural misunderstanding and cultural differences can improve understanding and reduce misunderstanding in MT-mediated communication.

5.2 User Behavior

The results detailed in Sect. 5 indicate that the experimental groups had better understanding than the control groups. To understand the reasons behind this, we qualitatively analyzed the chat log.

Explanation of Choices. Every participant in the experimental groups tried to explain the items in some detail when they were warned that those words might cause misunderstanding. Some examples are displayed in Fig. 4 and Fig. 5. The control group participants who were not warned seldom explained details of the items except when he/she was asked by their teammates, whereas the

experimental group members were more careful in explaining the word being introduced or explained it soon after.

User	English (Original)	Japanese (Translation)	Chinese (Translation)
English Speaker [E1]	for me i have a mud pie, a pie made from mud. three can of soda and a 1 pack of family size cracker	私にとっては、泥のパイ、泥から作られたパイがあります。 3缶のソーダと1パックのファミリーサイズのクラッカー	对我来说，我有一个泥饼，是用泥做的。三罐苏打水和一包家庭装饼干

Fig. 4. A chat message from English speaking user when is warned of cultural difference.

User	English (Translation)	Japanese (Original)	Chinese (Translation)
Japanese Speaker [E3]	Two corollas. It is a crown of flowers.	花冠がふたつ。花の冠です。	两个花冠。它是一朵花的冠冕。

Fig. 5. Chat messages from Japanese speaking user when is warned of cultural difference.

From the example in Fig. 4, the English speaker of the experimental group explained the word ‘mud pie’ and switched the word ‘pop’ to ‘soda’.

In Fig. 5, the Japanese participant from another experimental group did not only introduce the word 花冠 (*Kakan-Corolla*), but also explained that are the corollas.

However, some experimental group members (very few) did not explain the choice. In the interview, a Chinese participant stated that “*I just use it normally. I didn’t think too deeply*”. He also commented that “*I might be more careful if you add a warning message after the red letters*”. This opinion suggests that the warning was not obvious enough. Implementing the proposal in a real application would need stronger alerts.

Word Substitution. Sometimes a participant would use a word different from choice written in their documents. In the real world, especially when writing, we often use synonyms to provide variety and catch the reader’s interest [1]. Synonyms are selected to best describe the matter being raised [4]. These experimental group members switched more words than the control group members as they were more aware of the cultural differences. We found that using alternate words could yield immediate understanding. For example, in the experiment, replacing the word *pop* with *soda* raised translation accuracy and allowed the other participants to understand more easily.

Skipping the Discussion of Incomprehensible Topic. Many times, the participants skipped the choices that were incomprehensible to them. For example, in control group *C1*, when the Japanese speaker introduced 花冠 (*Kakan-Corolla*), the English speaker was presented with *corolla* which yielded uncertainty. Accordingly, the English speaker did not discuss the choice and followed the flow of the conversation, especially when the other two members agreed to settle on the two *bicycles* without further discussion of the other choices.

5.3 Preliminary Experiment Result

When we compare the understanding exhibited by the experimental groups from the data shown in Table 1, the experimental group that had two languages (*E1*) understood all the choices right after they were introduced, while the other experimental groups, *E2* and *E3*, understood 72.22% and 77.78% of the choices immediately upon introduction.

The reason for the sudden understanding of group *E1* compared to groups *E2* and *E3* could be for the following reasons:

First, the message is only translated once in two-language communication but the message is translated twice in three-language communication. Moreover, the two translation outputs in three-language communication could be different. There is more chance of misunderstanding when the message is translated into many languages. In the case of using foreign language (i.e. the Chinese speaker using English), if the foreign language skill of the participant is good enough, the resulting communication is likely to be superior to that achieved when using MT.

Second, if fluency in the use of the foreign language is achieved (Chinese speaker using English), it is possible that the participant will be more aware of cultural differences. Since the participant had experience in using English, he instinctively tried to make his message understandable in English.

6 Discussion

This section discusses the pattern of failure to communicate via MT-embedded chat, the lost in translation, limitation, and the future direction of this work.

6.1 Pattern of Failure to Communicate

After the end experiment we analyzed the choices that yielded failures in establishing mutual understanding and found two patterns.

Surface Failure to Establish Mutual Understanding. Participants simply could not understand the words used. In this case, the participants acted in two different ways: ignoring the choice and trying to understand the choice. When the participants tried to understand the choice they would ask questions and

they would often reach an understanding after some explanations. This does not negatively impact the understanding result but the questions and explanations might interrupt the conversation flow. However, when the participants chose to ignore what they did not understand, they would fail to establish mutual understanding. We categorized this problem as *surface failure* as it is obviously known to the participants themselves that they failed to understand and mutual understanding could not be established. However, the speaker would not know if the recipient understood the choice or not.

Underlying Misunderstanding. Sometimes, the discussion seems to be proceeding smoothly, without any problem; however, the participants actually misunderstood the conversation. We discovered this problem from the interview when we asked the participants about the choices shared by the other group members.

For example, 修行者 (*Xiuxing Zhe - Religious practitioner*) was translated as a practitioner. The groups that failed to understand this word thought they understood correctly, but they were wrong. In that situation, no questions were raised, and the conversation seemed to be simple and short. This problem is deeper than the surface failure and it is important that this kind of problem be detected and ameliorated.

6.2 Lost in Translation

Some misunderstanding occurred because of translation failures. A translated word could miss some or all of the full subtlety of meaning or significance. In the given list for Chinese speaker, 修行者 (*Xiuxing Zhe - Religious practitioner*) is translated as *practitioner* in English and 開業医 (*Kaigyoyi - Medical practitioner*) in Japanese. But the term 修行者 (*Xiuxing Zhe*) actually means a practitioner of a religion, usually Buddhism, who often undertakes extensive pilgrimages. The English translation was too vague because practitioner has several meanings, and the translation in Japanese was wrong, as shown in Fig. 6.

User	English (Translation)	Japanese (Translation)	Chinese (Original)
Chinese Speaker [C2]	The first is a practitioner, the second is a teenager, and the third is an electrical technician	1人目は開業医、2人目は1 0代、3人目は電気技術者	第一个是修行者，第二个 是青少年，第三个是电气 技师

Fig. 6. Chat log of the control group when the word ‘practitioner’ was mentioned.

Here, people who do not understand Asian culture might not be able to draw on the knowledge of the extensive travels undertaken by monks who travel around and might instead recall their minister who often stays in her/his church.

The translation gaps are inherent across cultures, and it is difficult to deal with them.

6.3 Limitation and Future Work

Even though we found that the participants were more careful to explain the choices, it is also possible that the participants did so because they are told to be aware of potential misunderstandings. We would like to find this out in the future by conducting another experiment where the experimental groups are warned of the words with the potential to cause misunderstanding while the controlled group will be warned of the different words that have low potential to cause misunderstanding. In addition, the preliminary experiment to study the number of languages used here can be extended as a full experiment in the future.

7 Conclusion

Intercultural collaboration depends on establishing mutual understanding and minimizing as much misunderstanding as possible. MT can help with the language barrier but cultural misunderstanding still happens. To solve this misunderstanding problem, an existing work previously suggested warning the user of the possible cultural differences when using MT-mediated communication. The main contribution of this paper is the experiment conducted to validate the suggestion. Our research results are useful for multilingual chat tool design.

In this research, we designed and conducted an experiment comparing experimental groups that received the warnings and control groups that had no warning. We found that the experimental groups had successfully established mutual understanding, while the control group did not and encountered many misunderstandings. We conclude that warnings of cultural misunderstanding significantly improve understanding in MT-mediated communication.

In addition to our main experiment, we conducted a preliminary experiment to study if the number of languages used in multilingual chat affects the degree of understanding. The preliminary results showed that a group that used only two languages established mutual understanding earlier than the group that used three languages.

Acknowledgments. This research was partially supported by a Grant-in-Aid for Scientific Research (A) (17H00759, 2017–2020), a Grant-in-Aid for Scientific Research (B) (18H03341, 2018–2020), and a Grant-in-Aid Young Scientists (A) (17H04706, 2017–2020) from the Japan Society for the Promotion of Science (JSPS).

References

1. Bailey, S.: *Academic Writing: A Handbook for International Students*. Taylor & Francis (2003). <https://books.google.co.jp/books?id=SUeRAgAAQBAJ>
2. Cho, H., Ishida, T., Yamashita, N., Inaba, R., Mori, Y., Koda, T.: Culturally-situated pictogram retrieval. In: Ishida, T., Fussell, S.R., Vossen, P.T.J.M. (eds.) *IWIC 2007*. LNCS, vol. 4568, pp. 221–235. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74000-1_17

3. Deutscher, G.: *Through the Language Glass: Why the World Looks Different in Other Languages*. Metropolitan Books, New York (2010)
4. Gookin, D.: *Word 2013 For Dummies*. For Dummies, Wiley (2013). <https://books.google.co.jp/books?id=gxd6bMou76EC>
5. Herring, C.: Does diversity pay?: race, gender, and the business case for diversity. *Am. Sociol. Rev.* **74**(2), 208–224 (2009)
6. Hofstede, G.: Cultural dimensions in management and planning. *Asia Pac. J. Manage.* **1**(2), 81–99 (1984)
7. Isahara, H., Bond, F., Uchimoto, K., Utiyama, M., Kanzaki, K.: Development of the Japanese wordnet. In: *Sixth International Conference on Language Resources and Evaluation* (2008)
8. Ishida, T., Murakami, Y., Lin, D., Nakaguchi, T., Otani, M.: Language service infrastructure on the web: the language grid. *Computer* **51**(6), 72–81 (2018)
9. Kelley, J.F.: An iterative design methodology for user-friendly natural language office information applications. *ACM Trans. Inf. Syst. (TOIS)* **2**(1), 26–41 (1984)
10. Lafferty, J.C., Pond, A.W.: *The Desert Survival Situation: Problem: a Group Decision Making Experience for Examining and Increasing Individual and Team Effectiveness*. Human Synergistics (1974)
11. Miller, G.: *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge (1998)
12. Nishimura, I., Murakami, Y., Pituxcoosuvarn, M.: Image-based detection criteria for cultural differences in translation. In: *International Conference on Collaboration and Technology*. Springer (2020)
13. Pituxcoosuvarn, M., Ishida, T.: Multilingual communication via best-balanced machine translation. *New Gener. Comput.* **36**(4), 349–364 (2018)
14. Pituxcoosuvarn, M., Ishida, T., Yamashita, N., Takasaki, T., Mori, Y.: Machine translation usage in a children’s workshop. In: Egi, H., Yuizono, T., Baloian, N., Yoshino, T., Ichimura, S., Rodrigues, A. (eds.) *CollabTech 2018*. LNCS, vol. 11000, pp. 59–73. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-98743-9_5
15. Pituxcoosuvarn, M., Lin, D., Ishida, T.: A method for automated detection of cultural difference based on image similarity. In: Nakanishi, H., Egi, H., Chounta, I.-A., Takada, H., Ichimura, S., Hoppe, U. (eds.) *CRIWG+CollabTech 2019*. LNCS, vol. 11677, pp. 129–143. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28011-6_9
16. Yamashita, N., Inaba, R., Kuzuoka, H., Ishida, T.: Difficulties in establishing common ground in multiparty groups using machine translation. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 679–688. ACM (2009)
17. Yoshino, R., Hayashi, C.: An overview of cultural link analysis of national character. *Behaviormetrika* **29**(2), 125–141 (2002)
18. Yoshino, T., Miyabe, M., Suwa, T.: A proposed cultural difference detection method using data from Japanese and Chinese wikipedia. In: *2015 International Conference on Culture and Computing (Culture Computing)*, pp. 159–166. IEEE (2015)