BCT Journal

Journal of Business, Communication & Technology

The Characteristics of Face-to-Face Communication Using Consecutive Speech Machine Translation Technology

Kazuma Hatano[1]*
[1]*Soka University, Japan*

Abstract The growing role of technology in society, particularly machine translation, has grown due to its enhanced accuracy. Yet, communication quality depends not only on translation precision but also on user factors. It is vital to understand communication characteristics when using machine translation for its future application in daily life, business, and education. This study investigates these characteristics in face-to-face communication using consecutive speech machine translation technology. Participants were international students at a Japanese university who conversed using speech machine translation. Despite mistranslations, they continued to communicate, either understanding most of the message or giving up due to incomprehension. Notably, the turn-taking order was alternated mechanically, even when there was a long wait for the other person to speak. Unlike previous studies, gestures were not used to overcome communication barriers due to mistranslations; rephrasing was preferred. This study proposes the concept of "speech production load," the burden of device operation and translation waiting time, as a determinant of the relationship between a speech machine translation system and user communication behavior in the context of modern technology.

Keywords: *Communication, Machine translation, Nonverbal behavior, Technology, Communication strategies*

1. Introduction

Machine translation is becoming an important part of the social infrastructure as a technology that provides multilingual support in hospitals and public institutions due to the dramatic improvement in translation accuracy in recent years. For example, in the U.S., where there are many immigrants, sales of "Pocketalk," a portable machine translation device, tripled in the April-September period of 2021 compared to the previous year, and sales to educational, medical, and public institutions are also increasing (Ikuta, 2021). In Japan, against the backdrop of a shrinking workforce, the government is leveraging technology by promoting the introduction of machine translation applications in public services, in addition to the enhancement of Japanese language education, as a multilingual response to the growing number of foreign workers and their families (Ministry of Justice, 2023). The Japanese government is also embracing technology beyond

***Corresponding Author:**
Kazuma Hatano
khatano@soka.ac.jp

Received: January 2024
Revised: March 2024
Accepted: March 2024
Published: April 2024

http://doi.org/10.56632/bct.2024.3104

the scope of daily life to the use of machine translation in business and is developing a simultaneous speech translation system by 2030 that will allow for complex business negotiations (Ministry of Internal Affairs and Communications, 2020). Thus, machine translation is expected to be a useful tool to support multilingual communication in a wide range of areas, from daily life to business.

It goes without saying that a major contribution to the success of communication via machine translation is the qualitative improvement of translation accuracy. In the sense that the quality of communication is greatly affected by the degree of accuracy with which the intended message is translated, it is only natural that research on translation accuracy has been, and will continue to be, the most important research issue that will lead to successful communication.

However, there are other factors besides translation accuracy that affect the quality of communication via machine translation. For example, if a speaker uses inappropriate vocabulary due to a misstatement or misunderstanding, the resulting communication may be negatively affected because the original utterance itself is inappropriate, no matter how high the translation accuracy is. In such cases, the machine translation user may attempt to repair communication by restating what was said, but this repair capability may be affected by the speaker's keen lexical sense and proficiency in using the machine translation device. Conversely, even if machine translation provides a mistranslation, the user may be able to flexibly interpret it and avoid the communication obstacle that could have been caused by the mistranslation. Therefore, we cannot ignore the possibility that there are factors that affect the quality of communication on the human side, not only the accuracy of translation.

Thus, it is important to investigate the characteristics of communication using machine translation in terms of user behavior, thereby acknowledging the influence of technology on it. When machine translation is used to provide language support for medical, welfare, and public services, communication at the counter or office is mainly face-to-face. In addition, since it is not known which language the service recipient will be using, multilingual support is necessary. The machine translation services that are currently mainstream in such situations are generally consecutive translation systems operated on dedicated devices or smartphones. Therefore, in considering the demand for the use of machine translation in society, an important question here is what characterizes multilingual face-to-face communication using consecutive speech machine translation. Based on this research question, the objective of this study is to identify the technological influences on the characteristics of communication in such situations by observing conversations between individuals speaking different languages.

2. Theoretical Framework

2.1. Technology-Enhanced Communication via Machine Translation

Researchers in the field of second language communication, in which people from different languages and cultures use a common language other than their native language, have studied various aspects of communication, including communication strategies (e.g., Kasper & Kellerman, 1997; Tarone, 1980), turn-taking (e.g., Sidnell, 2007; Stivers et al., 2009), and willingness to communicate (e.g., MacIntyre et al., 1998), and the findings from these studies have provided meaningful suggestions for second language learning and education. These findings will be helpful in considering the effective use of machine translation since communication via machine translation also takes place between people with different linguistic and cultural backgrounds. However, the premise of communication via machine translation is very different from that of second language communication since language production and comprehension are done primarily in the native or most fluent language when using machine translation. In addition, communication behavior may differ from normal conversation due to some unusual actions and procedures, such as operating translation devices. Therefore, in light of the advent of a society in which machine translation will be widely used, it is important to promote research on communication via machine translation as a research field different from second language communication.

However, findings from academic research on the characteristics of communication using machine translation have not accumulated much to date. The main objective of many studies on communication using machine translation has been to evaluate the performance of translation systems, including

translation accuracy, fluency, and usability (e.g., Seligman & Dillinger, 2015; Wu et al., 2011), and thus, there is limited understanding of the characteristics of communication via machine translation.

2.2. Telecommunication via Text-based Machine Translation

Many of the few studies on the communication characteristics of machine translation users are mainly concerned with telecommunication via text-based machine translation. For example, Ogura et al. (2005) conducted a study on how users of multilingual machine translation adapt to the translation system in text-exchange communication. In this study, they found that language pairs have a significant impact on adaptation to the translation system, as different language pairs use different methods of rewriting in case of mistranslation. As another example, the Language Grid Project (Ishida, 2006), led by Kyoto University, aimed to make machine translation a more accessible system by creating a platform that allows multiple translation services to be used in a combined manner. A series of studies in this project discussed the finding that communication using machine translation is more difficult than text-based chat communication in a second language (Yamashita & Ishida, 2006) and evaluated the usefulness of translation tools to support multilingual communication (Inaba, 2007). Other studies on text-based telecommunication via machine translation in multilingual environments have provided insights into performance evaluations such as the convenience of translation (Gaudio et al., 2016) and the usefulness of back-translation (Jain et al., 2021; Shigenobu, 2007), as well as strategies to overcome mistranslation (Nose & Hishiyama, 2014).

Some studies, including the above, attempted to find factors that influence communication via machine translation. In addition to language pairs (Ogura et al., 2005; Pituxcoosuvarn et al., 2018; Taira et al., 2021) we mentioned above, factors such as cultural differences (Hill et al., 2022; Pituxcoosuvarn et al., 2018) and proficiency in using translation devices (Hill et al., 2022) have been discussed as some factors that influence communication. Furthermore, Yamashita et al. (2009) and Yasuoka and Bjørn (2011) discussed the importance of building a common ground for mutual understanding in communication via machine translation. In another study observing children (Pituxcoosuvarn et al., 2018), when there were obstacles to communication via machine translation, the participants showed behaviors to facilitate understanding by using common tools such as a common language (English), gestures, and pictures that they knew each other. Thus, it has been suggested that using a common ground influences communication positively when the user uses it to overcome barriers such as mistranslation. Using such tools, machine translation users seem to be aware that they can continue to communicate even if the translation accuracy is not all accurate (Calefato et al., 2015).

2.3. Telecommunication via Speech Machine Translation

There are even fewer studies on the characteristics of communication using speech machine translation. According to a meta-analysis by Dew et al. (2018), of the 27 peer-reviewed articles in English on machine translation in the health field published between 2006 and 2016, 8 of these papers are about speech translation used between healthcare professionals and patients, but all of them are about evaluating the convenience of translation systems or translation accuracy; they do not have a central discussion about communication characteristics.

Among the few studies on communication via speech machine translation, Hara and Iqbal (2015), which investigated 23 pairs of multilingual communication via remote video communication, is of particular importance. In this study, 8 French-German pairs and 15 English-German pairs were tasked to collaborate on a story by telecommunication using a simultaneous machine translation system (in addition to speech translation, the user could choose to have the translation results displayed as text on the screen). The results showed that even when the translation accuracy was not good, the degree of usability of the machine translation system increased as the machine translation user adjusted the speed and volume of speech and used simpler vocabulary and sentences to speak more clearly, indicating that the machine translation user adapts to the translation system. Besides that, the speakers also tried to overcome communication obstacles by repeating or rephrasing their statements in case of mistranslation and by using nonverbal means such as gestures. They also reported that while the translated text displayed on the screen, in addition to the voice translation, was helpful for understanding, the slow translation speed, the delay in the screen display, and the fact that the translation results were translated

while the speaker was speaking made communication difficult. Shi et al. (2014) advocate the importance of machine translation users adapting to the translation system in order to prevent communication failures due to mistranslations, and the results of the above study presented by Hara and Iqbal (2015) suggest that such a point is appropriate.

The above studies discuss characteristics of remote communication using speech machine translation; however, little academic research has been conducted on the case of face-to-face communication. To fill the gap, the current study focuses on investigating the characteristics of face-to-face communication using speech machine translation.

3. Methodology

3.1. Participants

International students attending a private university in an urban area of Japan were invited, based on convenience sampling, to participate in the study, and 22 participants were paired up with those whose first language was different (11 pairs). Although Japanese was not necessarily the common language because some students belonged to programs that allowed them to graduate only in English, Japanese functioned as the lingua franca for some pairs. In addition, although to varying degrees, English functioned as the common language in all pairs because all students understood English. Since the purpose of this study was to explore common behaviors in any language pair, no particular consideration was given to language combinations, but in order to eliminate the possibility of guessing meanings from utterances in the partner's language before translation, close language combinations that would allow participants to infer meanings in the language without learning efforts were avoided. At the start of the observation, the investigator also verbally confirmed with each participant whether s/he had not learned the first language of the conversation partner. The language pairs observed are shown in Table 1.

Table 1
Research Participants

Session	Language Pair			
1	Participant 1	Korean	Participant 2	Burmese
2	Participant 3	Italian	Participant 4	Chinese (Simplified)
3	Participant 5	Portuguese (Brazil)	Participant 6	Chinese (Simplified)
4	Participant 7	Amharic	Participant 8	Cantonese (Hong Kong)
5	Participant 9	Mongolian	Participant 10	Portuguese (Brazil)
6	Participant 11	Russian	Participant 12	Spanish (Mexico)
7	Participant 13	Czeck	Participant 14	Chinese (Singapore)
8	Participant 15	Chinese (Malaysia)	Participant 16	Japanese [investigator]
9	Participant 17	Ukrainian	Participant 18	Hindi
10	Participant 19	Swahili	Participant 20	German
11	Participant 21	Napali	Participant 22	Japanese [investigator]

Two of the participants were unable to participate in the face-to-face conversation sessions due to illness and other reasons, resulting in a total of 20 participants. Since each conversation session was held on a separate date for each pair, the investigator of this research (a native Japanese speaker) had to rush to be the conversation partner for the two sessions in which there were absent participants, and the final total of 21 participants and 11 pairs of conversations, including the investigator, were recorded. However, the analysis was made based on the observations of 20 participants, excluding the investigator himself, in order to eliminate the possibility of mixing data on specific communication behaviors that the investigator unintentionally caused. During the analysis, the investigator provided only supplemental insights from the perspective of the "insider" who participated in the conversation.

3.2. Instruments

The machine translation device used in the conversation was "Pocketalk W," a portable translator that is widely used by general consumers and is capable of voice and text translation in 74 languages. The translator works in such a way that the speaker presses a button on the terminal as he or she speaks, and

when the button is released, the original utterances and their translation results are displayed on a small terminal screen, while the translated result is played back audibly from the terminal. This allows the conversation partner to hear the translation results in the language he or she understands, but since the translation is played after the speech is finished, it functions as a consecutive translation rather than a simultaneous translation. In addition, the original utterances and their translations are stored in the translator and in an online database.

The seven conversation topics listed in the Appendix were developed based on the following criteria, with the use of machine translation as a social infrastructure in mind.

a) Since previous studies (e.g., Pituxcoosuvarn et al., 2018) have found behavior using gestures as a common ground to assist communication, topics should include both concrete themes (Topic 2), which are easy to express with gestures and abstract ones (Topic 6), which are difficult to express with gestures alone.

b) Topics should include those that require an explanation of culture (Topic 3) because previous studies (e.g., Hill et al., 2022) have found cases where barriers to communication were caused by cultural differences.

c) Assuming that machine translation will be used in places such as hospitals where doctors use medical terminology, topics should include those that elicit conversations using technical terms (Topic 5).

d) Assuming that machine translation will be used in travel information offices, topics should include those that require an explanation of time and place (Topic 7).

e) Topics should include those that elicit the use of proper nouns, which are considered difficult to machine translate (Topic 4).

In addition, to familiarize the participants with the operation of the translation device, they were asked to introduce themselves using the translator in Topic 1.

3.3. Procedure

In Hara and Iqbal (2015), participants were given the task of creating a story, but in light of the increasing use of machine translation as part of language support in hospitals, city halls, etc., it is assumed that conversations will be primarily topic-based, such as giving directions, explaining a problem, or describing oneself, rather than jointly performing tasks. Therefore, in this study, pairs of two participants were asked to engage in conversation using speech machine translation on the topics listed in the Appendix, which were created based on the criteria outlined in the previous section.

Each participant was called into the investigator's office, where a quiet environment was ensured, and his/her partner was met. Before conversations started, each of them was given a portable translator and instructions on how to use it. Participants were then instructed to continue the conversation on the same topic until the investigator asked them to stop (approximately 3 minutes per topic), after which they were given the next topic; the conversation session ended when all seven topics were discussed. During the session, with the participants' consent, their conversations were video-recorded for later analysis. To control for the possibility of using English or Japanese as a common language when communication was not going well, 4 of the 11 pairs were instructed, before the conversation began, not to use English or Japanese during the conversation.

Both the original statements and the translated results recorded in the translation logs were translated into Japanese and English using DeepL and Google Translate to see if there were any differences between the original utterances and their translations. However, since the main purpose of this study was not to evaluate the accuracy of the translation, the details of the linguistic differences were not analyzed; rather, our main purpose of the observation was to determine where the communication problems were and what strategies were used to deal with them. The video data provided information about nonverbal behaviors and the order and time course of turn-taking, etc., that could not be determined from the translation logs alone. In addition, after the observation was over, participants were

asked about their impressions of the conversation using machine translation, which was included in the analysis as supplementary data.

Most importantly, since there is little research on the characteristics of face-to-face communication via speech machine translation, the main focus of the observations and analysis in this study was to discover communication characteristics in a broad and exploratory manner, leading to focused and detailed studies in the future.

4. Results

The observations in this study yielded important findings regarding face-to-face communication using consecutive speech machine translation. The main findings are summarized in Table 2 and discussed in detail in the remainder of this section.

Table 2
Main Findings

Category	No.	Main findings
translation accuracy	Finding 1	Translation accuracy depended on the language pair.
	Finding 2	Proper nouns were difficult to translate.
	Finding 3	The use of foreign words sometimes caused misunderstanding, especially when used as technical terms or cultural expressions.
continuation of conversation	Finding 4	Participants rarely pointed out mistranslations and continue conversation.
	Finding 5	Participants continued the conversation by switching topics when they give up understanding mistranslations.
turn-taking	Finding 6	Participants did not steal a turn from their conversation partner even when there was a long waiting time.
	Finding 7	Turns were alternated in an orderly fashion, which is unnatural compared to normal conversation.
	Finding 8	Speakers tended to follow the same turn-taking patterns as their interlocutor.
	Finding 9	The timing of turn-taking was inferred by the silence after the translation and by the pointing gestures of the partner.
	Finding 10	The timing for expressing emotion tended to occur when the partner understood the translation.
common ground	Finding 11	Common languages and gestures were rarely used to remove communication barriers caused by mistranslations; rather, rephrasing was preferred.
	Finding 12	Most gestures used were those that express understanding/agreement or incomprehension/denial/questioning.

4.1. Evaluation of Translation Accuracy

Since the purpose of this study is not to evaluate translation accuracy, we will not focus on discussing whether each translation is appropriate or not. However, since we cannot discuss translation accuracy and communication behavior completely in isolation, we would like to briefly evaluate translation accuracy overall. For example, there were few communication problems for the pairs of Participant 1 (Korean) and Participant 2 (Burmese), Participant 3 (Italian) and Participant 4 (Chinese), and Participant 11 (Russian) and Participant 12 (Spanish), but there were considerable mistranslations and incomprehensible situations for the pairs of Participant 7 (Amharic) and Participant 8 (Cantonese) and Participant 17 (Ukrainian) and Participant 18 (Hindi). In addition, proper nouns were sometimes difficult to translate. For example, "Kyoto" was recognized as a proper noun for a Japanese city, but "Hachioji," which has a relatively low recognition level, was not recognized correctly; it was recognized as a word similar to the sound "Hachioji" in the speaker's language, which was translated completely out of context. Overall, however, communication was better than expected for the first conversation using the translator, although some language pairs had considerable difficulty communicating in some situations, and some proper nouns were not translated well. It is expected that repeated use of the translator and familiarity with the translation system will help facilitate smoother communication.

In addition, from the feedback obtained after the observation session, we learned that in some cases, when the participants wanted to talk about their specialized fields in Topic 5, they did not have the appropriate expressions for certain specialized terms in their first languages, or they did not know how to express the relevant specialized terms because they had studied the subject in English. It was also pointed out that it was difficult to have a conversation about spring break in Topic 2 because there is no word for "spring" in Swahili. In such cases, it is assumed that in actual conversation, words that cannot be expressed in the first language will be replaced by foreign words, but when machine translation is used, such foreign words may be recognized as words with similar pronunciation in the language, instead of being translated in their original meaning, as in the case of "Hachioji" mentioned above. This issue may be improved in the future for frequently used vocabulary, but it must be noted when evaluating translation accuracy that in some cases, the translation does not take into account the use of foreign words.

4.2. Continuation of Conversation in the Event of Mistranslation

One thing we learned from our observations of communication behavior is that even when there was some mistranslation, participants rarely pointed it out and did not stop communicating. For example, when Participant 19 (Swahili) said "Nagano resort ...," which included the foreign English word “resort,” it was recognized as "Nagano results arsenal." This was incomprehensibly translated into German as "Nagano-Ergebnisarsenal" (Nagano results arsenal). Participant 20 (German) tilted his head for a moment when he heard this, but continued speaking as if he got the general message from the back and forth. This was the case for 10 of the 20 participants, and if we include the discrepancies in minor linguistic nuances, we can say that this was the case for almost all of them. In face-to-face speech communication, the conversational partner is right in front of you, so you can respond immediately. Therefore, it is considered easier than in text-based communication or remote speech communication to point out mistranslations to gain understanding. However, unless the translation was critically incomprehensible, the behavior rarely pointed out mistranslations but tried to understand the translation with the given translation results.

However, in some cases, the act of continuing the conversation involved switching to a different topic as a result of giving up because they could not understand the turn immediately preceding it. For example, in the pair of Participant 17 and Participant 18, there was a turn where both participants could not understand each other due to a mistranslation, and both participants appeared to be tilting their heads, but instead of making corrections, they gave up and moved on with the conversation. In addition, the translator recognized Participant 14’s original statement in Chinese as an expression that did not fit the context, so he reiterated it several times while shaking his head and laughing a little, but in the end, it was not translated appropriately. Therefore, his partner, Participant 13, gave up trying to understand and proceeded with the conversation by starting another question.

4.3. Turn-Taking

We also observed the turn-taking characteristic of machine translation users. In consecutive translation, the original language is not conveyed to the other party at the time of speech, and can only be understood through translation, resulting in a time lag. This inevitably lengthens the turn time and increases the time that neither of the dialogue partners speaks (latency). However, in this observation, even if the latency was long, participants did not take their partner's turn to speak until the translation results were available or even after a long silence. As a striking example, in the conversation between Participant 7 and Participant 8, after Participant 7's utterance and its translation, there was a 17-second silence in which Participant 8 seemed to be thinking about something. During that time, Participant 7 did not steal a turn but simply waited.

Another important finding was that the turns were alternated in an orderly fashion without any instructions from the investigator. In other words, one person did not talk too much or remain silent, and the turns were done mechanically; when one person spoke, the other person also spoke for about the same length of time, and when that was over, the other person spoke again. This kind of behavior was observed in 18 of the 20 participants. The above exchange is quite unnatural turn-taking when compared to normal conversation. In addition, two of the 20 participants (in the same pair) did not pass

the turn to the partner in a single translation but rather made several short statements and then gave the partner a turn. Interestingly, even in this case, they took turns the same way as each other, even though they had not discussed and decided beforehand how to take turns. This indicates that in face-to-face communication using consecutive speech machine translation, speakers tend to follow the same turn-taking pattern as their interlocutor.

The participants seemed to understand when to start their own turn by the silence after the translation was done and by the action of their partner pointing his/her hand at them to encourage them to turn. They were also observed placing their hand on their own chests to see if it was their turn (10 out of 20 participants). A further interesting observation was that when making joking remarks, the participants did not laugh at the time of speech but laughed when the other person understood the content via the translator. This suggests the possibility that in the case of consecutive translation, the timing for expressing emotion tends to occur when the other person understands the translation rather than at the time of utterance. The speaker seemed to subconsciously understand that sharing the emotion "This is funny, isn't it?" is meaningless unless the other party understands the translation.

4.4. The Use of Common Language and Gestures

Another finding from the observations of this study is that common ground, such as common language and gestures, is not the primary means of removing communication barriers caused by mistranslations. For the four pairs who were instructed not to use English or Japanese, the use of common language was limited to cases such as when Participant 8 (a Japanese language learner) noticed a mistranslation of her own statement and unintentionally said in Japanese, "Sorry" and not with the intention of removing communication barriers. For the other pairs who were not instructed to use the common language, there was no use of English or Japanese as a means to eliminate communication barriers. For example, Participant 4 said, "I'll finish the sentence" in English because the translation started before he finished what he wanted to say. His use of English was not intended to remove a barrier to communication; it was only an inadvertent remark about a procedural matter.

When there was a mistranslation, participants did not compensate for it with a gesture but rather tried to remove the communication barrier by rephrasing it into a more easily translatable word. For example, Participant 14 was trying to explain that people in Singapore celebrate two New Years (January 1 and Chinese New Year), but she quickly realized that it was mistranslated; she immediately rephrased the sentence to make it work. In most other cases, participants attempted to remove communication barriers primarily by rephrasing and rarely by using gestures. In fact, there was only one instance (by Participant 2) in which a gesture was used for such a purpose.

In addition, the use of gestures was not varied but was mostly an act of expressing understanding /agreement or incomprehension/denial/questioning. Regardless of the concreteness or abstraction of the conversation topic, most gestures were used to confirm that the translation was understandable or to communicate that the translation was not understood. Nodding, which communicates understanding, was a behavior seen by all participants, except that it was extremely infrequent in the conversation between Participant 17 and Participant 18. In cases of incomprehension, shaking the head (3 of the 20 participants), raising the eyebrows, and tilting the head (2 of 20 participants) were observed (4 of the 20 participants said "I don't understand" via machine translation). Conversely, when participants realized that the original statement was incorrectly recognized by the device, they would often indicate, "This translation is wrong" by shaking their heads or waving a hand from side to side (this was noticeable in 12 of the 20 participants).

5. Discussion

This study aimed to explore the influence of technology by identifying communication characteristics using consecutive speech machine translation in an exploratory manner. The results from the observations not only confirmed some of the findings of previous studies in other contexts, such as telecommunication via text-based and speech machine translation, but also provided new insights that may shed light on future research on face-to-face communication using speech machine translation.

First of all, the current study confirmed some of the identified factors that hinder communication in terms of translation accuracy. The most prominent was language pair, which significantly affected the quality of communication (Ogura et al., 2005; Pituxcoosuvarn et al., 2018; Taira et al., 2021). In addition, similar to the study by Hara and Iqbal (2015), the use of proper nouns also affected communication. This is because they are difficult to translate via neural machine translation (NMT), on which the translator used in this study is based. Since NMT needs a large corpus for better translation, infrequent proper names are particularly difficult to translate (Sharma et al., 2023). Furthermore, the influence of foreign words used as loanwords in conversation is noteworthy. Although models have been proposed (e.g., Mi et al., 2020) to improve the identification of loanwords, this potential difficulty should be recognized as a caveat in communication using speech machine translation.

The study also found that conversations continued even in the presence of mistranslations, which was observed in other studies (e.g., Calefato et al., 2015; Yasuoka & Bjørn, 2011). An interesting observation, however, is that some conversations continued not because the mistranslation did not interfere with the understanding of the overall message but because it was completely incomprehensible, and therefore the participants gave up on understanding. These observations indicate that the act of continuing the conversation involves either not correcting minor mistranslations because the majority of the message can be understood or giving up because the mistranslations cannot be understood at all. In both cases, the behavior could be seen as labor-saving in terms of the effort required to correct mistranslations and confirm their meaning. In the present study, since the conversation was topic-based rather than having a goal to be achieved, such as the completion of a task, it is possible that the tendency to give up when the translation was not understood at all was strongly expressed.

Although Hara and Iqbal (2015) identified the difficulty of turn-taking in communication using machine translation, there are not many studies that extensively discuss the characteristics of turn-taking in that context; therefore, it has potential significance for future research. One of the most distinctive characteristics found in this study is that participants did not steal a turn from their partner and took turns alternatively in a mechanical fashion. It is conceivable that they could not intervene because they did not understand the meaning until they heard the translation results, but that would not explain the above case in which there was a long silence (17 seconds) after the translation. One potential reason, instead, is that both parties understood that they were using machine translation and that it takes time to get used to unusual ways of conversation, so they seemed to respect the other person's time needed to do so; therefore, they waited until their partner responded back no matter how long it takes. In fact, another finding of Hara and Iqbal (2015) was that participants tended to “forgive” the difficulty of communication because they understood that it was via machine translation. This can be thought of as a consideration for the interlocutor, based on the meta-awareness that communication is taking place via machine translation.

Another important finding about turn-taking was that we observed behaviors that attempted to match turn-taking patterns and speech volume with each other. Since mechanical and unfamiliar turn-taking was, in a sense, "new" communication for all research participants, they may have unconsciously tried to determine their own turns using their partner's turn as a reference axis. The finding that turn-taking patterns and speech volume tended to be the same suggests the possibility that, in communication via machine translation, both speakers can communicate equally in terms of power, regardless of their knowledge or proficiency in the second language. This is because, in second language communication, native speakers of the language used tend to take linguistic initiative over the non-native speakers with whom they interact (Liddicoat, 2016). However, with machine translation, both speakers are speaking in their native or most fluent language, so it may be less likely that one of them will take the initiative. It is important that opportunities to speak are not deprived by unbalanced linguistic power relations, especially in education, since motivation to speak is crucial in language learning (Hashimoto, 2002). Kelly and Hou (2021), in fact, advocate the use of machine translation as a tool to empower students in the classroom because the study found that the learners saw machine translation as a legitimate tool for multilingual communication that they can use and integrate with their language resources.

As mentioned earlier, previous studies such as Yasuoka and Bjørn (2011) and Pituxcoosuvarn et al. (2018) discuss the importance of common ground as a tool to address communicative obstacles caused

by machine translation and to ensure successful communication. The current study, on the other hand, identified that common language and gestures were rarely used to remove communication barriers caused by mistranslations; instead, rephrasing was preferred. The use of gestures was effective in situations where verbal communication was not immediately possible (due to machine translation with long latency) as a means of telling the other party whether the translation was correct.

Some participants also used the thumbs-up sign or OK sign without speech, both of which were used to express agreement (4 out of the 20 participants). Agreement can be expressed by showing only a gesture, so the participants probably felt that there was no need to have trouble communicating with the other person via machine translation. In a normal conversation, simply saying "yes" is sufficient, but when using machine translation, there is a greater physical, cognitive, and temporal burden on the speaker to press a button, say "yes" to the device, check whether the speech is correctly recognized, and wait for the translation to be done. Therefore, speaking through machine translation requires a communicative burden, compared to the simplicity of using a gesture to indicate agreement. In other words, nonverbal behavior helped the participants avoid the procedural burden of using machine translation. This explains why most gestures used in the observed conversations were those that express understanding /agreement or incomprehension/denial/questioning.

The use of gestures also seemed to be associated with the facilitation of turn-taking; gestures can play a role in facilitating communication, not in the sense of removing obstacles due to mistranslations, but in the sense of facilitating turns. For example, speakers can signal the turn switch by pointing with their hand to their partner, or by nodding to indicate that they understand their partner's translation, thereby facilitating the next turn. Without these nonverbal behaviors, they may not know if they need to rephrase due to a mistranslation or if they can proceed with the conversation (i.e., go to the next turn). Nodding is seen in normal conversation as a nonverbal behavior that indicates that one is listening to the other person (Duncan & Fiske, 1977), but in this study, some other gestures were also observed that would appear unnatural in normal conversation, such as indicating a turn switch by pointing at the other person with the hand. The use of such unnatural gestures suggests that in conversations via machine translation, it is important to use nonverbal behaviors such as pointing gestures for smooth turn-taking.

The findings on the use of gestures discussed above reveal some important implications for understanding the nature of face-to-face communication using consecutive speech machine translation technology. As noted above, gestures may be primarily used to express simple answers to convey intentions such as agreement or incomprehension, especially when speakers find it cumbersome to use the translator. In other words, the use of gestures is expected to increase as the communicative burden increases. Therefore, this study assumes that the speaker's use of gestures changes with the degree of communicative burden caused by the operation of the translator and the associated waiting time. The study proposes to call this burden “speech production load," and argues that it serves as one of the factors that determines the relationship between a speech machine translation system and the communicative behavior of speakers using it.

The implications above are important, especially when we use a consecutive machine translation device. As speech production load is reduced, resistance to speech is also reduced, which may result in a switch from gesture to speech; for example, instead of relying solely on the use of gestures to express agreement, speakers may increasingly express their intentions through speech (while still accompanied by gestures). In this sense, the introduction of automatic simultaneous machine translation systems may bring about a change in the way speakers use gestures, as the speech production load is expected to decrease by reducing device operation and waiting time. However, Hara and Iqbal (2015), who used a simultaneous machine translation system, found that participants experienced communication difficulties due to speech overlap, in which the device played translations while participants were speaking; they, in fact, preferred text translations in this regard. Therefore, when considering the high speech production load of consecutive translation and the possibility of speech overlap of simultaneous translation, one way to achieve better and smoother communication may be to use simultaneous text translation in which speakers can audibly speak to the system while the translation results are displayed on a screen. This could lead to the lowest speech production load.

The concept of speech production load could also provide a general conceptual description of the process by which companies producing translation devices are developing more user-friendly devices. For example, a wearable translator worn like a pair of glasses can automatically display the translation results on a board like the lens of the glasses while you are working (e.g., on a construction site). When using such a translation device, the speech production load is considered low because translation is performed automatically by simply speaking. However, hand-held translation devices require the user to stop working, lift up the translation device, press a button, speak, put down the translation device when finished, and start working again, a cumbersome process that places a very heavy burden on the user's speech. In this sense, the development of easy-to-use translation devices, such as wearable translators, can be viewed as a process of reducing speech production load.

The purpose of this study was to discover, in an exploratory manner, the characteristics of communication through consecutive speech machine translation technology, which have not been clarified to date. Therefore, the current discussion may include hypothetic inferences that require multiple studies before they can be fully elucidated. For example, whether the concept of speech production load is really one factor that determines the relationship between machine translation systems and human communication behavior requires further research. One important aspect of such further research is to observe communication using automatic simultaneous speech translation in which participants audibly speak to the device while the translation results are displayed in text, which is expected to reduce the speech production load, as discussed above. Another important research topic for turn-taking would be to compare turn-taking patterns in normal second language communication with those using machine translation in order to see the possibility that linguistic power relations may be balanced by the use of machine translation. In addition, the design of the study needs to be reconsidered. As we discussed, conversations observed in the present study were topic-based, and it may have led participants to give up on understanding difficult mistranslations compared to conversations aimed at task completion. Future research should take this possibility into account in the study design.

Disclosure Statement

The author claims no conflict of interest.

Funding

The research did not receive any specific grants from funding agencies.

References

Calefato, F., Lanubile, F., Conte, T., & Prikladnicki, R. (2015). Assessing the impact of real-time machine translation on multilingual meetings in global software projects. In P. Runeson (Ed.), *Proceedings of the ACM-IEEE International Symposium on Empirical Software Engineering and Measurement* (Vol. 21, pp. 1002-1034). Association for Computing Machinery. https://doi.org/10.1145/2372251.2372299

Dew, K. N., Turner, A. M., Choi, Y. K., Bosold, A., & Kirchhoff, K. (2018). Development of machine translation technology for assisting health communication: A systematic review. *Journal of Biomedical Informatics, 85*, 56-67. https://doi.org/10.1016/j.jbi.2018.07.018

Duncan, S. & Fiske, D. (1977). *Face-to-face interaction: Research, methods, and theory*. Lawrence Erlbaum Associates. https://doi.org/10.4324/9781315660998

Gaudio, R., Burchardt, A., & Branco, A. (2016). Evaluating machine translation in a usage scenario. In N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation* (pp. 1-8). European Language Resources Association.

Hara, K., & Iqbal, S. T. (2015). Effect of machine translation in interlingual conversation: Lessons from a formative study. In B. Befole & J. Kim (Eds.), *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 3473-3482). ACM. https://doi.org/10.1145/2702123.2702407

Hashimoto, Y. (2002). Motivation and willingness to communicate as predictors of reported L2 use: The Japanese ESL context. *Second Language Studies, 20(*2), 29-70.

Hill, D. C., Gombay, C., Sanchez, O., Woappi, B., Romero Vélez, A. S., Davidson, S., & Richardson, E. Z. L. (2022). Lost in machine translation: The promises and pitfalls of machine translation for multilingual group work in global health education. *Discover Education, 1*, Article 3. https://doi.org/10.1007/s44217-022-00004-z

Ikuta, G. (2021, December 17). *"Pocketalk" translation device goes public*. The Nikkei Business. https://business.nikkei.com/atcl/seminar/19/00030/121600297/?n_cid=nbpnb_mled_epu

Inaba, R. (2007). Usability of multilingual communication tools. In N. Aykin (Ed.), *Usability and internationalization. Global and local user interfaces*. *Lecture notes in computer science* (Vol. 4560, pp. 91–97). Springer. https://doi.org/10.1007/978-3-540-73289-1_11

Ishida, T. (2006). Language grid: An infrastructure for intercultural collaboration. In S. Shimojo & J. Urban (Eds.), *Symposium on Applications and the Internet* (pp. 96-100). IEEE. https://doi.org/10.1109/SAINT.2006.40

Jain, A., Mhaskar, S., & Bhattacharyya, P. (2021). Evaluating the performance of back-translation for low resource English-Marathi language pair: CFILT-IITBombay @ LoResMT 2021. In J. Ortega, A. K. Ojha, K. Kann, & C. Liu (Eds.), *Proceedings of the 4th Workshop on Technologies for MT of Low Resource Languages* (pp. 158-162). Association for Machine Translation in the Americas.

Kasper, G., & Kellerman, E. (1997). *Communication strategies: Psycholinguistic and sociolinguistic perspectives*. Addison Wesley Longman.

Kelly, R., & Hou, H. (2021). Empowering learners of English as an additional language: Translanguaging with machine translation. *Language and Education, 36(*6), 544-559. https://doi.org/10.1080/09500782.2021.1958834

Liddicoat, A. (2016). Native and non-native speaker identities in interaction: Trajectories of power. *Applied Linguistics Review, 7*(4), 409-429. https://doi.org/10.1515/applirev-2016-0018

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a second language: A situational model of second language confidence and affiliation. *The Modern Language Journal, 82*(4), 545-562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

Mi, C., Xie, L., & Zhang, Y. (2020). Loanword identification in low-resource languages with minimal supervision. *ACM Transactions on Asian and Low-Resource Language Information Processing, 19*(3), 1-22. https://doi.org/10.1145/3374212

Ministry of Internal Affairs and Communications. (2020). *Global communication plan 2025*. https://www.soumu.go.jp/main_content/000678485.pdf

Ministry of Justice. (2023). *Comprehensive measures for acceptance and coexistence of foreign nationals (FY2023 revised)*. https://www.moj.go.jp/isa/content/001404440.pdf

Nose, T., & Hishiyama, R. (2014). An analysis of interaction protocol within an MT-mediated gaming simulation. *The transactions of Human Interface Society, 16*(1-4), 235-248.

Ogura, K., Hayashi, Y. Nomura, S., & Ishida, T. (2005). Language-dependency in user's adaptation for MT systems in MT-mediated communication. *Journal of Natural Language Processing, 12*(3), 183-201. https://doi.org/10.5715/jnlp.12.3_183

Pituxcoosuvarn, M., Ishida, T., Yamashita, N., Takasaki, T., & Mori, Y. (2018). Machine translation usage in a children's workshop. In H. Egi, T. Yuizono, N. Baloian, T. Yoshino, S. Ichimura, & A. Rodrigues. (Eds.), *Collaboration technologies and social computing. Lecture notes in computer science* (Vol. 11000, pp. 59–73). Springer. https://doi.org/10.1007/978-3-319-98743-9_5

Seligman, M., & Dillinger, M. (2015). Evaluation and revision of a speech translation system for healthcare. In M. Federico, S. Stüker, & J. Niehues (Eds.), *Proceedings of the 12th International Workshop on Spoken Language Translation* (pp. 209–216). ACL.

Sharma, R., Katyayan, P., & Joshi, N. (2023). Improving the quality of neural machine translation through proper translation of name entities. In D. K. Sharma & R. Agarwal (Eds.), *6th International Conference on Information Systems and Computer Networks* (pp. 1-5). IEEE. https://doi.org/10.48550/arXiv.2305.07360

Shi, C., Ishida, T., & Lin, D. (2014). Translation agent: A new metaphor for machine translation. *New Generation Computing, 32*(2), 163-186. https://doi.org/10.1007/s00354-014-0204-0

Shigenobu, T. (2007). Evaluation and usability of back translation for intercultural communication. In Aykin, N. (Ed.), *Usability and internationalization. Global and local user interfaces. Lecture notes in computer science* (Vol. 4560, pp. 259-265). Springer. https://doi.org/10.1007/978-3-540-73289-1_31

Sidnell, J. (2007). Comparative studies in conversation analysis. *Annual Review of Anthropology, 36*, 229-244. https://doi.org/10.1146/annurev.anthro.36.081406.094313

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoon, K., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences, 106*(26), 10587-10592. https://doi.org/10.1073/pnas.0903616106

Taira, B. R., Kreger, V., Orue A., & Diamond, L. C. A. (2021). Pragmatic assessment of Google Translate for emergency department instructions. *Journal of General Internal Medicine, 36*(11), 3361-3365. https://doi.org/10.1007/s11606-021-06666-z

Tarone, E. (1980). Communication strategies, foreigner talks, and repair in interlanguage. *Language Learning, 30*(2), 417-431. https://doi.org/10.1111/j.1467-1770.1980.tb00326.x

Wu, C., Xia, F., Deleger, L., & Solti, I. (2011). Statistical machine translation for biomedical text: Are we there yet. *AMIA Annual Symposium Proceedings*, *2011*, 1290–1299.

Yamashita, N., & Ishida, T. (2006). Effects of machine translation on collaborative work. In P. Hinds & D. Martin (Eds.), *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work* (pp. 515-524). Association for Computing Machinery. https://doi.org/10.1145/1180875.1180955

Yamashita, N., Inaba, R., Kuzuoka, H., & Ishida, T. (2009). Difficulties in establishing common ground in multiparty groups using machine translation. In D. R. Olsen & R. B. Arthur (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 679-688). Association for Computing Machinery. https://doi.org/10.1145/1518701.1518807

Yasuoka, M., & Bjørn, P. (2011). Machine translation effects on communication: What makes it difficult to communicate through machine translation? In K. Hachimura, T. Ishida, & N. Tosa (Eds.), *International Conference on Culture and Computing* (pp. 110-115). IEEE. https://doi.org/10.1109/Culture-Computing.2011.28

Appendix

Conversation Topics

Topic 1: Introduce yourselves to your partner. You can talk about anything, including your country of origin, the language you speak, your recent interests, etc.

Topic 2: Tell your partner how you spent your spring break. Listen to his/her stories, give your opinions, and ask questions.

Topic 3: Please tell your partner about how you spend New Year's in your country. Listen to what he/she has to say, give your opinions, and ask questions.

Topic 4: Talk about a restaurant that originated in your country and is popular among people. Listen to what he/she has to say, give your opinions, and ask questions.

Topic 5: What is the most impressive concept you have learned in your field of study? Listen to what he/she to say, give your opinions, and ask questions.

Topic 6: What does "happiness" mean to you? Listen to what he/she has to say, give your opinions, and ask questions.

Topic 7: Assuming that a meeting will be held next week, determine the best date, time, and location for the meeting based on your actual schedule.