AECT | ASSOCIATION FOR EDUCATIONAL COMMUNICATIONS & TECHNOLOGY

CrossMark

CULTURAL AND REGIONAL PERSPECTIVES

Applications of speech-to-text recognition and computer-aided translation for facilitating cross-cultural learning through a learning activity: issues and their solutions

Rustam Shadiev[1] · Ting-Ting Wu[2] · Ai Sun[3] ·
Yueh-Min Huang[3]

Abstract In this study, 21 university students, who represented thirteen nationalities, participated in an online cross-cultural learning activity. The participants were engaged in interactions and exchanges carried out on Facebook® and Skype® platforms, and their multilingual communications were supported by speech-to-text recognition (STR) and computer-aided translation (CAT) systems. The participants spoke in their native languages, and the STR system generated texts from their voice input. The CAT system then simultaneously translated the STR-texts into English. The aim of this study was to examine the accuracy rates of STR and CAT processes for different languages during intercultural communication. We also explored issues associated with these processes and how they were addressed in the study context. In addition, an attempt was made to determine whether or not our learning activity as supported by STR and CAT technologies facilitated cross-cultural learning. Our results showed that the lowest STR accuracy rate was for Belizean English whereas the highest STR accuracy rate was for French and Hindi. The lowest CAT accuracy rate was for Mongolian and Filipino, and the highest was for Spanish, Russian, and French. Seven issues associated with the STR process and ten issues associated with the CAT process were identified. The participants employed ten workarounds to address the STR-related issues and thirteen workarounds to address the CAT-related issues. We refer to a workaround as a method used by the participants to overcome a limitation related to either STR or CAT. Finally, our results demonstrated that cross-cultural learning took place; the participants understood and could explain foreign traditions to others and could also compare foreign traditions with their own. Based on our results, we made several suggestions and provided implications for the teaching and research community.

✉ Yueh-Min Huang
huang@mail.ncku.edu.tw

[1] School of Education Science, Nanjing Normal University, 122 Ninghai Road, Nanjing 210097, China

[2] Graduate School of Technological and Vocational Education, National Yunlin University of Science and Technology, 123, Daxue Road, Douliu 64002, Taiwan

[3] Department of Engineering Science, National Cheng Kung University, 1 University Road, Tainan 70101, Taiwan

⁂ Springer

Introduction

Technological advancements have brought the world and its people closer together. Nowadays, most people own computing devices and have access to the Internet, so intercultural interactions and exchanges take place more widely and actively (Çiftçi 2016). Thus, because cultural differences exist in our globalized society, the development of cross-cultural understanding has become an important issue (Aparicio et al. 2016; Bentley et al. 2005). People need to recognize and understand cultural differences in order to avoid problems that may arise during intercultural communication, because we all must co-exist with others peacefully in this interconnected world (Rogers et al. 2007).

Modern information and communication technologies play an important role in aiding cross-cultural learning programs (Aparicio et al. 2016; Ertmer et al. 2011; Rogers et al. 2007). Çiftçi (2016) reported that the most frequently used technologies to aid cross-cultural learning included: online discussion boards, text-based chatting, blogs, email exchanges, video recording and conferencing, and podcasting. These technologies support both asynchronous and synchronous communication and are valuable, enjoyable, and suitable for cross-cultural learning, because they enrich the nature and function of communication. Nevertheless, several critical issues related to cross-cultural education still exist that hamper intercultural interactions and exchanges among people. According to Osman and Herring (2007), Reynolds (2013, 2015), Shadiev and Huang (2016), and Wu (2014, 2015), the language barrier is the most critical factor in cross-cultural learning because the lack of a common language makes it difficult to communicate. To address this issue, we applied speech-to-text recognition (STR) and computer-aided translation (CAT) systems in the present study. During intercultural communication, the STR system generated texts from participants' speech in their native languages, and the CAT system then simultaneously translated them into foreign languages understandable to others. In this manner, it was assumed that STR and CAT systems aided the cross-cultural learning of participants from different countries who were speaking different languages. However, Barrachina et al. (2009), Hwang et al. (2012) Kuo et al. (2012), Fountain and Fountain (2009), Mellebeek et al. (2005), and Shadiev et al. (2016) warned that the accuracy rates of STR and CAT technologies should be considered when applying them to teaching and learning because content created by STR and CAT technologies can be useful and meaningful for learning only when STR and CAT texts are reasonably accurate. These scholars suggested that STR and CAT systems are still not able to deliver perfect transcriptions and translations. Since these claims were made several years ago, and technology has advanced rapidly in recent years, this issue may need to be revisited. In addition, applications of STR and CAT systems to support cross-cultural communication and learning have received very little attention from the scientific community in the past. Therefore, this study is an attempt to address these issues.

Related literature and cultural context

Cross-cultural understanding

Cross-cultural understanding is generally defined as the basic ability of individuals to correctly recognize, understand, and interpret cultural differences. According to Moran et al. (2014), diverse cultural backgrounds include differences in languages, appearance, food, traditions, values, and so on. What is allowed in one culture can be prohibited in another. For example, "Americans love beef, yet it is forbidden to Hindus, while the forbidden food in Muslim and Jewish culture is pork, eaten extensively by the Chinese and others" (Moran et al. 2014, p. 12). Thus, understanding the culture of others helps to overcome cultural differences and maintain harmonious relationships. Cross-cultural understanding can be explained by the cultural convergence theory (Gudykunst et al. 1988; Kincaid 1979). According to this theory, cross-cultural understanding takes place through communication and information exchange between/among two or more persons from different cultures when they reach a mutual understanding of each other's culture and the world in which they live. That is, experiences of and insights into other cultures that persons communicate and share among themselves enable the expansion of their cultural awareness and behavior (Gudykunst et al. 1988; Kincaid 1979). Jaakkola et al. (2010) and Talalakina (2010) argued that cross-cultural understanding can be measured based on the following four rubrics: (1) *cross-cultural knowledge* refers to familiarization of an individual with different cultural characteristics, values, beliefs, and behaviors; (2) *cross-cultural awareness* refers to an individual's internal understanding and appreciation of a culture; (3) *cross-cultural sensitivity* refers to an individual's ability to read into situations, contexts, and behaviors that are culturally rooted and to react to them appropriately, and (4) *cross-cultural competence* an individual's ability to work effectively across cultures, i. e., to respect cultural differences, adapt to changing situations, and benefit from them. Cross-cultural learning has been defined as a process involving active experience and participation that enables an individual to acquire knowledge and to absorb new attitudes and values related to different cultures (Yamazaki and Kayes 2004). Talalakina (2010) suggested fostering cross-cultural learning through multiple goals, such as (1) providing information on cross-cultural peculiarities, (2) encouraging cross-cultural communication and open-mindedness among individuals from different cultures, (3) providing tools for cross-cultural communication, and facilitating the interpretation and reactions of individuals to cross-cultural contexts, and (4) modeling situations to implement the acquisition of cross-cultural knowledge. Kohlberg (1984) proposed several techniques for implementing cross-cultural learning. One technique involves providing learners with opportunities to experience cultural diversity through readings, simulations, and watching videos (i.e., indirectly) and another is by interacting with people from other cultures (i.e., directly).

Different learning activities intended to facilitate cross-cultural understanding have been discussed in earlier studies. *Self-introduction* helps students become acquainted with one another and their cultures (Chase et al. 2002; Liu 2007; Tu 2004). In addition, more social interactions can be achieved through self-introduction (Curtis and Lawson, 2001). *Creating media content and sharing it with others* contribute to diversification of cultural expression and the empowerment of cross-cultural perceptions (Jenkins et al. 2009). *Discussing* enables students to discern and respect multiple perspectives across diverse communities (Jenkins et al. 2009). *Experiencing foreign cultures* through play, performance, and appropriation helps with adopting alternative identities. In addition, students

are able to sample and remix media content meaningfully for the purpose of improvisation and discovery (Jenkins et al. 2009). By *reflecting on foreign culture*, students are able to share their reflections and experiences with their peers (Tu 2004).

Speech-to-text recognition and computer-aided translation

Speech-to-text recognition (STR) technology synchronously transcribes text streams from an individual's speech (Ranchal et al. 2013; Shadiev et al. 2017). According to earlier research, STR has great potential to aid the learning of different target groups, i.e., students in a physical classroom, students learning online, non-native speakers, or students with cognitive or physical disabilities (Shadiev et al. 2017). For example, Hwang et al. (2012) and Kuo et al. (2012) adopted STR in online learning environments to support both teaching and learning. When students experienced problems associated with network connections and could not hear the lecturer, they could still follow the lecturer by reading STR-texts. Kheir and Way (2006) applied STR to assist the learning of hearing-impaired students. These students could use STR-texts to comprehend lecture content since they could not hear it. Ranchal et al. (2013) employed STR to generate lecture transcriptions. When STR was available during lectures, the students could focus on the lecture and did not have to take class notes. In addition, transcriptions were useful to them for studying lecture content after class. Ryba et al. (2006) adopted STR to support the learning of non-native speakers of English during lectures in English. The students listened to the lecturer and simultaneously read STR-texts to understand words with which they were unfamiliar and to verify and clarify words they had misheard or did not catch when the lecturer spoke too fast.

Computer-aided translation (CAT) technology translates texts from one language into another (Godwin-Jones 2011). According to earlier studies, CAT technology is a potentially reliable and valuable tool to aid learning, especially in the field of second/foreign language (SL/FL) learning. For example, Hermet and Désilets (2009) applied CAT during SL/FL writing activities to compose essays as well as to correct grammatical and lexical errors in essays. Omar et al. (2012) employed CAT during an online discussion. CAT translated and searched for appropriate words that helped the students express their opinions and ideas during the discussion. In addition, CAT checked grammar and spelling and helped with overcoming problems that occurred when they were constructing sentences. ElShiekh (2012) adopted CAT to research a writing course in order to explore the translation process from English into Arabic and vice versa in the field of advertising as well as Koranic texts and literary works. Shadiev and Huang (2016) introduced STR and CAT to Taiwanese and Uzbek students to facilitate their cross-cultural understanding. Students communicated and exchanged information about their local cuisine and culture. However, students in their study represented two cultures, and their communication was bilingual.

Research motivation and cultural context

In this study, we focused on facilitating cross-cultural understanding in participants from different countries who speak different languages. Informed by related theory and earlier studies, we designed a cross-cultural learning activity. The cultural context of this study was specific to each participant and represented his/her culture. That is, we had twenty-one participants from thirteen countries speaking their mother tongue in ten different languages. During the learning activity, participants introduced twenty-one traditions from

their culture. They communicated and exchanged information with one another in their mother tongue. We applied STR and CAT technologies to support multilingual interaction among the participants, help them understand each other, and facilitate their cross-cultural learning. In this study, we aimed to: (a) measure the accuracy rate of current STR and CAT technologies, (b) explore issues associated with STR and CAT processes and how they can be solved, and (c) investigate whether the use of STR and CAT applications are a feasible way by which to facilitate cross-cultural learning. The significance of this study is that we attempted to address the language barrier issue with STR and CAT technologies during cross-cultural learning. That is, this research sheds new light on the feasibility of STR and CAT systems applications to support cross-cultural multilingual communication among participants. In addition, this study provides new insights into the accuracy rate of STR and CAT systems utilized during a cross-cultural learning activity. Furthermore, this study enhances our understanding of existing issues associated with STR and CAT processes and how they can be solved. It also addresses the following research questions: (1) What is the accuracy rate of current STR and CAT technologies? (2) What are the issues associated with STR and CAT processes, and how do students resolve them? (3) Are STR and CAT applications a feasible method by which to facilitate cross-cultural learning?

Table 1 Participants, their tradition, and foreign tradition their experienced

ID	Nationality	Native language	Presented tradition	Experienced tradition
1	Taiwan	Mandarin	Atayal facial tattoo	Mongolian dumplings
2	Taiwan	Mandarin	The Night Market	Day of the Dead
3	Taiwan	Mandarin	Red envelop	Apologizing to everyone
4	Taiwan	Mandarin	Rite of passage	Apologizing to everyone
5	Taiwan	Mandarin	Praying in temple	Apologizing to everyone
6	Burkina Faso	French	Festival of masks	Eat more than three times a day
7	Burkina Faso	French	The age of maturity	Holi
8	India	Hindi	Holi	Praying in temple
9	India	Hindi	Kalbelia dance	Mongolian dumplings
10	Indonesia	Indonesian	Apologizing to everyone	Morning pilaf
11	Indonesia	Indonesian	The cracker eating contest	A Central American dish
12	Vietnam	Vietnamese	Buffalo fighting festival	Eat more than three times a day
13	Vietnam	Vietnamese	Ancestor worship	Month of the country
14	Paraguay	Spanish	Paraguayan Easter	Festival of masks
15	Panama	Spanish	Month of the country	Praying in temple
16	Honduras	Spanish	Sawdust carpets	Eat more than three times a day
17	Mexico	Spanish	Day of the Dead	Festival of masks
18	Belize	English	A Central American dish	Praying in temple
19	Mongolia	Mongolian	Mongolian dumplings	Kalbelia dance
20	Philippines	Filipino	Eat more than three times a day	Mongolian dumplings
21	Uzbekistan	Russian	Morning pilaf	Red envelop

Method

Twenty-one university students representing thirteen nationalities participated in this study (Table 1) of which twenty of the participants were between 20 and 27 years old, and one was 32 years old. This infers that there were no generational differences. The participants were enrolled in the same course, and they did not know each other prior to this study. Sixteen participants had less than 1 year of experience or no experience with the use of STR technology, and the remaining participants had more experience. Seven participants had used CAT technology for less than 3 years, and fourteen participants had used it for a longer period. Fourteen participants had less than 1 year or no cross-cultural learning experience, and seven participants had more than 1 year of experience.

We aimed to facilitate the cross-cultural understanding of the participants with a learning activity following the general recommendations of earlier studies (Chase et al. 2002; Curtis and Lawson 2001; Jenkins et al. 2009; Liu 2007; Tu 2004). Thus, the study included four one-week steps. In the first step, the participants introduced themselves, their hobbies, and their interests. In the second step, the participants introduced their local traditions and related culture. They were asked to identify one local tradition, describe it in terms of process, and explain its origin and relation to their culture. In the third step, each participant selected one tradition and experienced it; for example, student ID 21 learned about the "red envelope" culture, prepared one red envelope, and presented it to a friend as a wedding gift with an accompanying speech explaining the gift's meaning. After that, students shared their personal experiences with the selected tradition with the other participants. Before selecting a tradition to experience, each participant confirmed that he/she had no prior knowledge about it. Communication among the participants during the first three steps was carried out in a closed Facebook® group, which took place asynchronously. In the fourth step, all participants met online face to face to communicate synchronously with each other about themselves, their traditions, and their experiences with foreign cultures. The participants communicated via Skype® messenger. Student communication in the asynchronous or synchronous modes was organized based on the nature of each step of the learning activity. For example, asynchronous communication was useful during the self-introduction step because the students were not acquainted with one another, and it helped decrease anxiety and inhibitions (Yang and Chen 2007). Students were motivated to disclose personal information, which could be problematic for some during face-to-face communication (Yang and Chen 2007). Asynchronous communication allowed the students to have enough time to post their information, read the posts of others, and ask or answer questions (Shadiev et al. 2017). In addition, the students were able to communicate at convenient times (Tu 2004). Synchronous communication, on the other hand, was useful for exchanging culturally-related information, reflecting on cultural experiences, and making or receiving comments instantly (Tu 2004). Since the students were already acquainted with one another after the first three steps of the learning activity, real time communication helped increase their interaction and gave them a sense of community (Tu 2004; Shadiev et al. 2017). This is why communication among the students was synchronous in the last step.

The participants used STR technology to generate texts from their voice input in their native languages, and then they used CAT technology to simultaneously translate STR-texts into English. We employed this approach because the participants could not understand the foreign languages of the other participants, but all of them could understand English. Extracts from the participants' communication content are provided in the

Appendix as an example. Before the learning activity, the researchers trained the participants to use STR and CAT, and the participants had a 1 week period to practice the technology. The researchers met with every participant after each step to discuss technology-related issues and how these issues were addressed by the participants. If some participants had no solutions for their STR- and CAT-related problems, we offered them those proposed by other participants.

We introduced STR and CAT technologies to the participants to aid in their multilingual communication. For this purpose, we used the Google® Translate system (Fig. 1). Google® Translate is an automated machine translator application that offers free online language translation service to users with internet access. It can be used to translate words or phrases from one language into another and supports more than 100 languages. The Google® Translate algorithms are not rule-based like most CAT systems but are rather based on statistical analyses (Tobin 2015). That is, the algorithms rely on a large corpus of professionally translated texts and look for equivalences (Fountain and Fountain 2009). Google® Translate features an STR tool that is activated by clicking on a microphone button. A speaker selects a voice input language, speaks, and STR then generates text from the voice input. After that, the STR-text is simultaneously translated from the speaker's native language into English by the Google® Translate CAT tool.

The analytical data were collected from two main sources: online communication among the participants and one-on-one semi-structured interviews. Table 2 presents a number of words participants communicated in different languages during different steps of the learning activity. These numbers are presented as averages because there were several participants who were native speakers of the same language (e.g., two native French speakers and three native Spanish speakers). In addition, we included the total number of words in the table to show how many words the participants communicated in total in particular languages. This information could be useful to aid with understanding how much the STR and CAT tools can help with cross-cultural multilingual communication. We analyzed the content of the online communications to measure the accuracy rates of the STR and CAT systems. The accuracy rates of both the STR and CAT systems
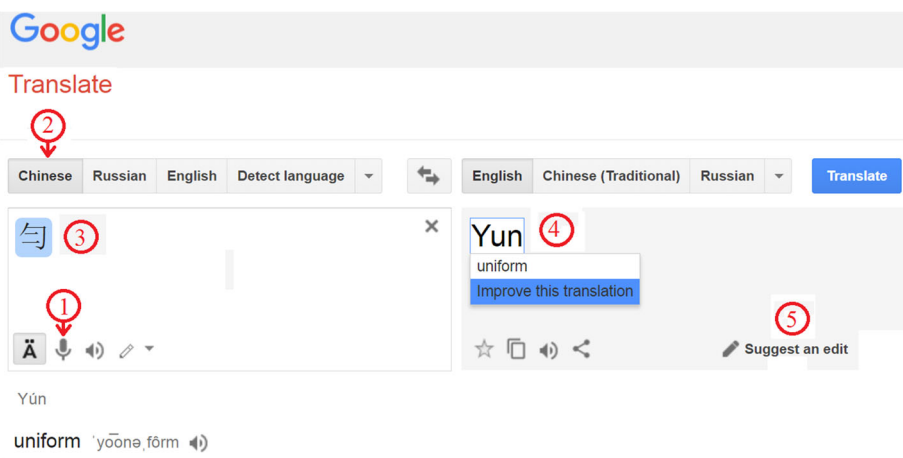

Fig. 1 Google® Translate *1* STR tool, *2* input language, *3* input area, *4* translated text, and *5* improving translation

Table 2 Number of words in posts (average and total) and accuracy rates of STR and CAT (in percentage)

Language	Number of words in posts (average)				Number of words (total)	STR					CAT				
	Step 1	Step 2	Step 3	Step 4		Step 1	Step 2	Step 3	Step 4	Average	Step 1	Step 2	Step 3	Step 4	Average
Mandarin (n = 5)	130	264	274	385	5262	93.59	96.23	98.10	99.89	96.95	94.59	93.88	95.49	98.73	95.67
French (n = 2)	141	205	177	324	1690	97.56	98.74	98.44	99.30	98.51	97.83	97.09	97.10	99.79	97.95
Hindi (n = 2)	179	248	217	305	1896	98.47	97.86	98.32	99.40	98.51	94.24	95.32	94.82	99.32	95.93
Indonesian (n = 2)	163	172	169	519	2043	97.78	98.23	98.21	99.70	98.48	96.29	96.91	97.12	99.20	97.38
Vietnamese (n = 2)	113	194	203	280	1578	98.40	94.48	98.43	99.29	97.65	94.10	93.05	95.63	98.68	95.37
Spanish (n = 3)	176	232	242	260	3637	97.13	97.57	97.21	99.25	97.79	98.16	96.81	97.81	99.80	98.15
English (n = 1)	341	243	287	550	1421	93.75	90.95	92.15	98.91	93.94	NA	NA	NA	NA	NA
Mongolian (n = 1)	112	160	158	228	658	NA	NA	NA	NA	NA	92.68	92.61	94.38	97.79	94.37
Filipino (n = 1)	115	119	170	294	698	99.26	95.89	98.45	99.64	98.31	90.22	92.59	97.35	98.23	94.60
Russian (n = 1)	136	244	291	359	1030	99.06	96.17	98.79	99.67	98.42	97.50	96.69	98.60	99.30	98.02

AECT

were calculated by dividing the total of all correct words in the text by the total words and multiplying by 100.

One experienced researcher, a co-author of this paper, carried out in-depth, one-on-one semi-structured interviews with all participants. Interviews were carried out face to face in the researcher's office and online via Skype at the end of each step of the learning activity. All interviews were conducted in English. In the interviews, the researcher asked the participants about their experiences with the technology during the learning activity. The interview content was then analyzed by the researcher and his two assistants to derive issues related to STR and CAT processes and the workarounds that the participants used to address such issues. Each interview took approximately 30 min. The researcher recorded all interviews using a digital recorder and transcribed the content for analysis. Subsequently, the researcher and two assistants reviewed the interview transcripts and separately coded the text segments that contained information related to the research focus. They aggregated codes with similar meanings and formed categories to produce a framework in order to report the findings. Notable differences in the coding and categorization processes were resolved through rater discussion until a consensus was achieved. The inter-rater reliability of the interview data was evaluated by using Cohen's kappa, and the result exceeded 0.90, thus demonstrating high inter-rater reliability.

We analyzed the communication content of the participants generated during the learning activity to measure their cross-cultural understanding. Since the participants communicated in ten different languages, we analyzed the content already translated into English by the participants. We adopted a coding unit concept for this analysis. Text segments that represented cross-cultural understanding of foreign traditions were highlighted and coded. Codes with related meanings were then collected and grouped. Established groups of codes produced a framework for reporting the research findings. We evaluated cross-cultural understanding with respect to *cross-cultural knowledge* (i.e. student familiarization with different cultural characteristics, values, beliefs, and behaviors), *awareness* (i.e. internal understanding and appreciation of a culture), and *sensitivity* (i.e. ability to read into situations, contexts, and behaviors that are culturally rooted and to react to them appropriately) (Jaakkola et al. 2010; Talalakina 2010). We did not focus on *competence* (i.e. ability to work effectively across cultures) because our learning activity took place online, so the participants had no chance to experience foreign culture and traditions in authentic contexts (e.g., to attend the morning pilaf ceremony in Uzbekistan), and thus, could not demonstrate their competence. A score of "1" was given if a participant's communication content represented *cross-cultural knowledge*, "2" for awareness, and "3" for *sensitivity*. If communication content did not represent any of these, a participant got a score of "0." Three raters were involved in the evaluation process. The inter-rater reliability coefficients among them were calculated using Cohen's kappa. The mean inter-rater reliability among the three raters exceeded 0.90, which indicates excellent agreement.

Results and discussion

Accuracy rate of STR and CAT

Results related to the accuracy of the STR and CAT systems with respect to different languages are presented in Table 2 and Fig. 2 (as a percentage). The accuracy rate of STR
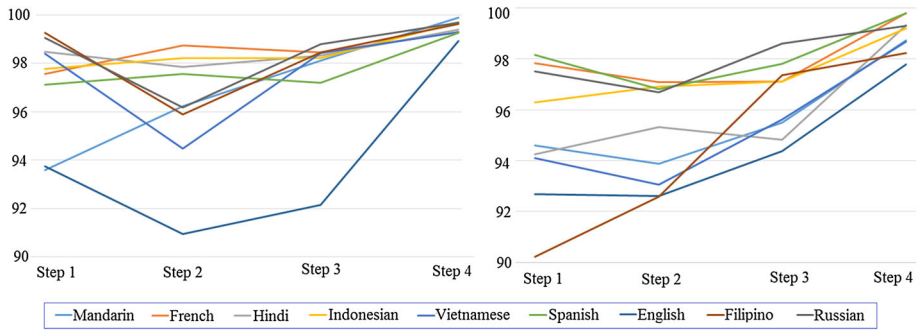
Fig. 2 Accuracy rates of STR (left) and CAT (right)

for Mongolian is absent because Google® Translate does not feature an STR function for this language. One reason for this is because Mongolian is not a widespread language as compared to the other languages under consideration here. We therefore asked the Mongolian participant to type content into CAT using a keyboard. In addition, since we asked the participants to translate their content from their native language into English, and the Belizean participant was a native speaker of English, the CAT accuracy rate for Belizean English is absent.

According to our results, the lowest STR accuracy rate was for English (the average was 93.94%), and the highest STR accuracy rate was for French and Hindi (the average was 98.51%). This result was surprising given that participant ID18 is a native speaker of English and since this is the most commonly used language, so the STR tool for English should be quite mature, and thus the STR accuracy rate should be high. We interviewed participant ID18 to discover the reason behind this low accuracy rate. The participant admitted that since he is from Belize, he has a very strong accent; therefore, STR was unable to recognize some spoken words correctly. Holm (1982) suggested that Belizean English differs from Standard English (i.e., British or the United States) in terms of both dialect and accent. This difference reflects the varied history of the speech community and the degree of contact that the community had with Spanish and Creole. STR is based on Standard English acoustic and language models. Thus, when participant ID18 pronounced words with an accent, STR was unable to recognize them correctly and generated text with errors. Since French and Hindi are widely spoken languages, the accuracy rate of STR for these languages was high. In addition, the French and Hindi participants mentioned that they had rehearsed using STR beforehand and knew several strategies to apply to the use of this technology. This is why the highest accuracy rate was for French and Hindi.

According to our results, the lowest CAT accuracy rate was for Mongolian (94.37%) and Filipino (94.60%), and the highest rate was for Spanish (98.15%), Russian (98.02%), and French (97.95%). Perhaps, this is due to the wide use of Spanish, Russian, and French in contrast to Mongolian and Filipino. The CAT database is bigger for Spanish, Russian, and French and smaller for Mongolian and Filipino. When there is a small language database, CAT translates texts with lower accuracy rates as compared to when there is a large database (Aiken and Balan 2011). Another reason for differences in accuracy rates is the similarities/differences between English and these languages, where a larger difference results in lower accuracy (Tobin 2015). Mongolian and Filipino are Asian languages and thus have greater differences with English compared to European languages such as Spanish, Russian, and French.

Scholars have suggested that texts produced by STR and CAT systems may contain mistakes and ambiguities (ElShiekh 2012; Fountain and Fountain 2009; Tobin 2015). Therefore, STR/CAT-texts are acceptable and useful for learning only when they are generated with reasonable accuracy (Barrachina et al. 2009; Mellebeek et al. 2005). Hwang et al. (2012) and Kuo et al. (2012) argued that accuracy rates of texts of more than 85% can be considered to be reasonable and that texts with these accuracy rates will indeed enable teaching and learning. Following this argument and the fact that STR and CAT have advanced greatly after nearly 4 years of development, an accuracy rate of 85% can now be easily achieved. Therefore, the participants in this study agreed that the STR- and CAT-generated texts were acceptable and useful. These findings are in line with those of previous studies (Shadiev and Huang 2016).

Our results show that the STR and CAT accuracy rates for some languages were not consistent and changed throughout the steps of the learning activity. For example, in the case of Hindi, Vietnamese, English, Filipino, and Russian, the STR accuracy rate was higher in the first step as well as in the second one. For Mandarin, French, Vietnamese, Spanish, Mongolian, and Russian, the CAT accuracy rate was higher in the first step as compared to the second one. Theoretically, the accuracy rate of the STR and CAT system should increase throughout the steps as the participants gain experience with using the systems and subsequently utilize them during communication more efficiently. However, our results are contradictory, so we interviewed the participants who speak these languages in order to explore any possible explanations for the results. In the interviews, the participants mentioned that they used words that are very common, easy, and unambiguous for the first step (i.e., self-introduction). However, they had to use some specific words or terminology related to local traditions in the second step, so STR and CAT could not always recognize and translate them correctly. This is perhaps because the STR and CAT systems do not have such words in their database. Labov (2011) suggested that languages change over time because some words are borrowed from other languages or invented due to different cultural environments. For example, "panades" and "garnaches" are words related to popular food items in Belize that were borrowed from Spanish. These words are currently in wide use in Belizean English but do not exist in the CAT database for English. They are likely to be accurately recognized/translated by the STR/CAT for Spanish rather than by the STR/CAT for English.

In terms of perceptions of STR and CAT systems, in the interviews, the participants mentioned that it was easy to use both. In addition, the participants claimed that the accuracy rates of the translated texts were very good and acceptable. They could fully understand the translated content posted by other participants. The participants said that their experience with STR and CAT enabled them to find the strengths and limitations of the technologies, and they consequently could fully utilize these technologies during the learning activity. With each step, they obtained content with fewer errors due to their experience with the technologies except in some cases when they had to use specific words and terminology, e.g., to describe traditions and related culture. However, one important issue should be considered. Some participants said that Internet Explorer or Mozilla Firefox were their favorite browsers, but it turned out that these browsers do not feature the STR function in Google® Translate. Some other browsers (i.e., Opera®) feature the STR tool, but it may stop recording speech input before a speaker stops speaking. Therefore, the participants had to use Google® Chrome for their experience with using Google® Translate to be efficient.

Although earlier studies have suggested that STR (Hwang et al. 2012; Kheir and Way 2006; Kuo et al. 2012; Ranchal et al. 2013; Ryba et al. 2006) and CAT (Hermet and

Désilets 2009; Godwin-Jones 2011; Omar et al. 2012) are valuable tools to support learning, applications of these two technologies to cross-cultural learning have received little attention. In this study, we attempted to bridge this gap. Our results demonstrated the accuracy rates for different languages when using STR and CAT systems during multi-lingual cross-cultural communication.

Improving the accuracy rates of STR and CAT through various workarounds

In this section, we discuss several issues reported by the participants that are associated with STR/CAT-text accuracy rates and several workarounds the participants applied to address them. With regard to the STR process, the participants experienced seven issues and employed ten workarounds to address them (Table 3). *STR issue 1* The STR technology did not add punctuation marks in the STR-texts. As a result, the participants got long strings of words and could not distinguish one sentence from another. The participants then dictated punctuation marks (*STR workaround 1*) or inserted them manually (*STR workaround 2*). This approach helped the participants distinguish when one sentence ended and another started, showed how the sentence should be read, and made the meaning clear. However, it should be noted that the STR function of Google® Translate does not add punctuation marks for Mandarin, Hindi, Indonesian, Vietnamese, and Filipino but does for English, Spanish, French, and Russian.

STR issue 2 The STR system for Traditional Chinese generated some characters in Simplified Chinese (characters that are currently used in Mainland China) instead of in Traditional Chinese (characters that are currently used in Hong Kong, Macau, and Taiwan). The participants tried to provide the content verbally several times, but the STR system kept generating these characters in Simplified Chinese. Since the native Chinese speaking participants were from Taiwan and were unfamiliar with Simplified Chinese characters, they did not know if the STR-text in Simplified Chinese was correct or not, so they changed the simplified Chinese characters into traditional ones manually (*STR workaround 3*).

STR issue 3 The STR system misrecognized some specific names and special terminology and generated them into wrong words. For example, according to participant ID7 (a native speaker of French), the name of the culture in Burkina Faso is "Dogon," but the STR system recognized it as "de gang" (i.e., "gang" in English) because the word "Dogon" is a specific name of a culture and does not exist in French whereas "de gang" does. Participant ID7 tried to pronounce this word differently, but the STR system still generated an incorrect word. Then, the participant changed the misrecognized word into the correct one manually (*STR workaround 4*).

STR issue 4 The STR system generates errors when the spoken content is multilingual. For example, using Western names is popular in Asia nowadays, and if Asians say "my name is" in their native language and then add a name in English, the first part will be recognized correctly by the STR, but the second one will be incorrect. As shown in Table 3, STR-generated characters in Mandarin for the name "Jack," written as "這個" (i.e., "zhege"), are pronounced similarly. The participants used two workarounds to deal with this issue. They changed the misrecognized Chinese characters into the English word manually (*STR workaround 5*) or spoke one part of the sentence in Mandarin first and then switched the STR input language to a specific other language (i.e., English in this case), and spoke the second part in that language (*STR workaround 6*).

STR issue 5 If a speaker paused for a few seconds (e.g., to think about what to say next) and then resumed speaking again, the STR-text for the second part of the speech could

Table 3 Mistakes in texts generated by STR and strategies to address them

STR related issue	Example of an issue	STR strategy	Example of addressing an issue
1. STR does not add punctuation marks	Hello my friends I am ID14* I study Energy Engineering	1. Dictate punctuation marks 2. Insert punctuation marks manually	Speech: Hello my friends comma I am ID14 comma I study Energy Engineering full stop STR-text: Hello my friends, I am ID14, I study Energy Engineering.
2. STR generates character in Simplified Chinese form instead of Traditional Chinese form	Simplified Chinese characters: 中国 电影	3. Change Simplified Chinese characters into Traditional Chinese characters manually	Traditional Chinese characters: 中國 電影
3. STR misrecognizes names/terminologies	de gang [in French] quiere [in Spanish]	4. Change misrecognized word into correct one manually	Dogon [in French] tereré [in Spanish]
4. STR misrecognizes speech with multilingual content	Speech: 我是這個 [in Chinese] Read phonetically: Wǒ shì zhège [Translation: I was the]	5. Change Chinese characters "這個" to "Jack" manually 6. Speak whole sentence in Mandarin first (e.g. 我是) then switch STR input language to a specific one and speak in that language (e.g. Jack)	Speech: 我是Jack Read phonetically: Wǒ shì Jack [Translation: I'm Jack]
5. STR overrides text	Cantar y conversar con amigos [in Spanish]	7. Speak without too long pauses so that previously STR-generated text will not be overridden with newly generated one 8. Prepare script for a speech	Me gusta correr, cantar y conversar con amigos [in Spanish]
6. STR is unable to recognize some words	न*****	9. Change न***** to नमस्ते! manually	नमस्ते!
7. Similar pronunciation	雲 → Yun [one character in the name is misrecognized] il versus ils [in French he vs they]	10. Change 雲 to 勻/il or ils manually	勻 → Yun [this is correct character and have the same pronunciation and tone as misrecognized one]

* name was replaced by ID

override the STR-text for the first part. The participants found this experience unpleasant because it took some time for them to regenerate the overridden STR-texts. To address this issue, the participants spoke without long pauses (*STR workaround 7*), and they also prepared a script for their speech, so they could speak fluently without any pauses (*STR workaround 8*).

STR issue 6 The STR system did not recognize some words at all. For example, participant ID8 (a native speaker of Hindi) found that the STR system recognized only the

first syllable of the word "नमस्ते!" (i.e., "hello"), and the other part of the word was generated as asterisks, i.e., "न*****". This word is very common in Hindi, and the participant found it strange that the STR system could not recognize it correctly even after several attempts. Then, the participant changed the word "न*****" to "नमस्ते!" manually (*STR workaround 9*).

STR issue 7 The STR system was unable to recognize words with similar pronunciations (e.g., "il" instead of "ils" in French or "雲" instead of "匀" in Mandarin because they have similar pronunciations but different meanings). When the STR system generated wrong words instead of the correct ones, the STR-texts did not make any sense. Therefore, the participants had to change "雲" to "匀" or "il" to "ils" manually (*STR workaround 10*).

With regard to the CAT process, the participants experienced ten issues and employed thirteen workarounds to address them (Table 4). *CAT issue 1* Word order in the translated text was not the same as in the original sentence, which made the content of the CAT text incorrect and even incomprehensible in some cases. For example, participant ID21 made a self-introduction in Russian, and one word was relocated from the beginning of the translated text to the end, so the sentence became meaningless (Table 4). On the other hand, the CAT system translated the sentence and kept words in the same order as in a spoken sentence for Mandarin and Vietnamese. Although the words were translated correctly, their order in the translated sentence made it incorrect and meaningless. This is because word order in some languages is not the same as in English. It is suggested that if a sentence is similar in structure to a target language, the CAT system can translate more easily. Sentences in some languages, like Mandarin and Vietnamese, do not have a structure similar to that of English. In these languages, sentences do not have spaces between words, which adds complexity since the CAT system may not even know what constitutes a word. In order to address this issue, most participants tried to make their sentences shorter by breaking one sentence into small parts, and they separated parts using paragraphs (*CAT workaround 1*). Some students made their sentences shorter and used punctuation marks after each part, i.e., commas, semicolons, hyphens, and full stops (*CAT workaround 2*). In addition, students came up with the idea of making sentences in their native language in the active form, i.e., subject + verb + the remaining words (*CAT workaround 3*). They said that this workaround helped them obtain a better accuracy rate.

CAT issue 2 Some words in a sentence were missing after translation (mostly the verb "to be"). To address this issue, the participants added missing words to the translated text manually (*CAT workaround 4*), or they changed the spoken sentence by using different words to convey the same meaning (*CAT workaround 5*).

CAT issue 3 CAT added some extra or repeated words in the translated text. For example, when one participant greeted other participants in his/her native language, CAT doubled the word "Hello" in a translated greeting (Table 4). The participants deleted the extra or repeated word manually (*CAT workaround 6*) to address this issue.

CAT issue 4 The CAT system translated long sentences with more errors. For example, if there were several sentences, and no punctuation marks were inserted, CAT considered them to be one sentence and translated the words using its algorithms. As a result, the outcome was one translated sentence with words from different sentences mixed together and in the wrong order. To address this issue, the participants broke a long sentence into small parts (or shorter sentences) and separated the sentences into paragraphs (*CAT workaround 1*), or they inserted punctuation marks, i.e., commas, semicolons, hyphens, and full stops, into the sentences (*CAT workaround 2*).

CAT issue 5 Some participants do not use the verb "to be" in their native language for daily communication, e.g., when they say who they are. Therefore, when they

Table 4 Mistakes in texts translated by CAT and strategies to address them

CAT related issue	Example of an issue	CAT strategy	Example of addressing an issue
1. Wrong word order in translated text	In native language: *Привет всем меня зовут ID21 я из Узбекистана* [in Russian] Translation: *Hello everyone my name is I'm from Uzbekistan ID21*	1. Break a sentence into small parts and separate them using paragraph/s 2. Add punctuation marks, i. e. coma, semicolon, hyphen, full stop 3. Make sentence in active form, i.e. subject + verb + the rest words	In native language: *Привет всем. Меня зовут ID21. Я из Узбекистана.* Translation: *Hello everybody. My name is ID21. I am from Uzbekistan.*
2. Missing words in translated text	In native language: *Halo semuanya, nama saya ID11.* [in Indonesian] Translation: *Hello everyone, my name ID11.*	4. Add missing words ("*adalah*" here which means "*is*") to translated text manually 5. Change spoken sentence, i.e. use different words	In native language: *Halo semuanya, nama saya adalah ID11.* Translation: *Hello everyone my name is ID11.*
3. Extra/ repeated word/s in translated text	Hello Hello my name is	6. Delete extra/repeated word manually	Hello my name is
Wrong translation if… 4. A sentence is too long	In native language: *Hola que tal mi nombre es ID15 soy panameña actualmente estoy cursando mi primer año de diseño Industrial aparte del arte también me gusta jugar futbol y cocinar.* [in Spanish] Translation: *Hello my name is ID15 such'm Panamanian am currently studying my first year of Industrial design other art I like playing football and cooking.*	1. Break a sentence into small parts and separate them using paragraph/s 2. Add punctuation marks, i. e. coma, semicolon, hyphen, full stop	In native language: *Hola, mi nombre es Stefani. Soy panameña. actualmente estoy cursando mi primer año de diseño Industrial. Aparte del arte, también me gusta jugar futbol y cocinar.* Translation: *Hello, my name is Stefani. I am Panamanian. I am currently studying my first year of Industrial Design. Aside from the art, I also like to play football and cooking.*
5. Some words are missing in spoken sentence	In native language:我的名字ID5 [in Mandarin] Translation: My name ID5	7. Add missing words ["是" translated as "is"] in spoken sentence	In native language: 我的名字是ID5 Translation: My name is ID5
6. Use specific term/name in spoken sentence	In native language: buuz [in Mongolian] Translation: buuz In native language: 匀 [in Mandarin] Translation: uniform	8. Edit term/name in translated text manually 9. Improving translation by suggesting an edit: Click on "Suggest an edit" and type the correct word in the open box.	In native language: buuz Translation: dumplings In native language: 匀 Translation: Yun

Table 4 continued

CAT related issue	Example of an issue	CAT strategy	Example of addressing an issue
7. Use informal (casual) language in spoken sentence	In native language: 早! [in Mandarin] Translation: Early!	9. Improving translation by suggesting an edit: Click on "Suggest an edit" and type the correct word in the open box. 10. Use formal language in spoken sentence	In native language: 早上好! Translation: Good morning!
8. Use complex words in spoken sentence	In native language: *Би бол хоёр хүүхэдтэй айлын бага нь бөгөөд дээрээ нэг эгчтэй. Манай эгч бөөрний нарын мэргэжлийн их эмч.* [in Mongolian] Translation: *I have a two-child family is his sister. medical practitioners of our sister kidney.*	11. Use simple words in spoken sentence	In native language: *Би бол айлын отгон хүүхэд. Миний эгч бол бөөрний их эмч.* Translation: *I am the youngest child of a family. My sister is a kidney doctor.*
9. Use ambiguous words in spoken sentence	In native language: *Mi muñeca* [in Spanish] Translation: My doll	12. Revised "My doll" into "My wrist" in translated text manually	My wrist
10. Use cultural idioms in spoken sentence	In native language: Кто рано встает тому бог даёт [in Russian] Translation: Early risers that God gives	13. Revise the idiom in translated text manually	The early bird catches the worm

communicated this type of content, CAT did not include a form of "to be" in the translated text and thus made it incorrect. To address this issue, the participants had to add form of "to be" to the spoken sentences in their native language (*CAT workaround 7*). They said that it was correct when they used "to be," but the sentence became more formal.

CAT issue 6 CAT was not able to translate some specific terms/names correctly. For example, one participant tried to translate the Mongolian word "buuz" into English, and the CAT system translated it as the original word "buuz" instead of as "dumplings." In such cases, the participants edited the incorrect translation to the correct word manually (*CAT workaround 8*). Some participants suggested improving translation by adding a "suggesting an edit" function to the system (*CAT workaround 9*). That is, they suggested that the system could allow clicking on a "Suggest an edit" prompt that would then allow them to type the correct word in the box that opened (Fig. 1).

CAT issue 7 CAT was also unable to translate some words into informal or casual language. For example, it is normal to say, "Good morning!" in Mandarin either as "早" or "早上好." However, the former is more casual, and it was translated into English by the CAT system as "Early." The participants improved the translation by suggesting an edit (*CAT workaround 9*), or they had to use formal language in the spoken sentence in order to get the correct translation, i.e., "Good morning!" (*CAT workaround 10*).

CAT issue 8 The CAT system translated a sentence with more errors if some complex words were used in a spoken sentence (Table 4). When the participants used simple words in a spoken sentence (*CAT workaround 11*), the CAT translation accuracy rate was higher compared to when they used more complex words.

CAT issue 9 The CAT was not able to translate ambiguous words (i.e., words with two or more translations) based on context. For example, the Spanish word "Mi muñeca" can be translated either as "My doll" or "My wrist." CAT translated it as "My doll" whereas the participants needed it to be translated as "My wrist." The participant tried to translate it several times, but every time got the wrong translation and had to revise "My doll" into "My wrist" manually in the translated text (*CAT workaround 12*).

CAT issue 10 CAT was unable to translate some cultural idioms correctly. For example, a participant tried to use a Russian idiom to talk about local culture, and the CAT system translated it word by word, so the meaning of idiom was lost (Table 4). The participant had to revise the idiom in the translated text manually (*CAT workaround 13*).

Some workarounds, e.g., manually typing, changing, or correcting some words/characters in the texts produced by the STR or CAT systems can prove to be far from novel as actual workarounds for some experienced users. However, there are some categories of users, e.g., novice users, for whom these workarounds can be new and useful. Because our participants were from thirteen different countries, not all of them had easy or full access to information technology. In fact, thirty percent of our participants did not have much experience with the use of CAT technology prior to this study; therefore, some of these participants did not think of using these workarounds at all. When we discussed these workarounds with these participants, they found them very useful.

Facilitating cross-cultural understanding

We asked the participants at the beginning of the study to select and experience a culture they were not familiar with. The participants admitted that before our cross-cultural learning activity, they had no prior knowledge regarding the culture and traditions they experienced. After the cross-cultural learning activity, we assessed the cross-cultural knowledge, awareness, and sensitivity of the participants, and the results showed that all participants reached all levels of cross-cultural understanding. That is, the participants learned and understood the foreign traditions they experienced in the learning activity. According to the results, the participants had become familiarized with different cultural characteristics, values, beliefs, and behaviors, and they could recall, interpret, summarize, compare, and explain the traditions they learned about and experienced. Furthermore, the participants were able to appreciate a foreign culture and to reflect on their own experiences of a foreign culture. We provide extracts from the participants' communication representing key concepts related to foreign traditions (Appendix: Steps 3 and 4). For example, participant ID2 explained the Day of the Dead tradition and pointed out the differences between this tradition and that of Chinese culture; participant ID10 summarized and interpreted the "morning pilaf" tradition; participant ID20 understood what "buuz" was and was able to explain how to prepare it; participant ID21 talked about the red envelope tradition and its origins.

We interviewed the participants to explore their perceptions regarding the cross-cultural learning activity supported by the STR and CAT systems. In the interviews, the participants mentioned that applications of STR and CAT during the learning activity were interesting and useful for their cross-cultural learning. First, the participants were able to introduce their traditions and also reflect on their experiences with foreign traditions in

their native language. Second, the participants were able to read what others said in foreign languages without any knowledge of these languages. Therefore, applications of STR and CAT helped the participants communicate with each other in their own native language because the technologies translated the communication content. As a result, the participants understood foreign traditions through the content communicated by other participants in foreign languages during the learning activity. More importantly, they learned traditions from those who represented their own traditions. The participants mentioned that they could communicate with representatives of thirteen nationalities and select any of twenty-one traditions. Information posted by the participants was special and unique. In addition, the participants could ask some specific questions of those who introduced traditions they were interested in and could obtain instant answers. Without STR and CAT, the participants would not have been able to communicate with the other participants due to language barriers. Our findings are in line with the cultural convergence theory (Gudykunst et al. 1988; Kincaid 1979) and the findings of earlier studies. According to this theory, cross-cultural learning takes place if learners from different cultures communicate and exchange learning information with each other. We designed a learning activity to implement cross-cultural knowledge, and the participants exchanged information about their culture, traditions, and cross-cultural peculiarities using online tools (Talalakina 2010). The participants discussed their traditions, experienced foreign traditions, and shared their experiences; such communication and experiences thus enabled cross-cultural learning to take place (Kohlberg 1984).

One important point worth mentioning here is our learning activities design and the technological support obtained by combining STR and CAT technologies to facilitate cross-cultural understanding occurred in a way that was impossible or ineffective to achieve using other methods of instruction. In contrast to other designs used for cross-cultural learning activities, e.g. reading or watching videos (Kohlberg 1984), our learning activities were focused on communication and exchange of culture-related information among students. Participation in our learning activities led students not only to receive information but also to experience it and reflect on their experiences. Furthermore, students were able to discuss culture-related information with a host from a given culture. Therefore, our learning activities helped promote peer learning practice and increased a sense of authenticity. On the other hand, some earlier studies also focused on interaction among participants to enhance their cross-cultural understanding (Aparicio et al. 2016; Çiftçi 2016; Ertmer et al. 2011; Rogers et al. 2007); however, the language barrier was reported as the critical issue that hindered communication and exchange of culture-related information among students speaking different languages (Osman and Herring 2007; Shadiev and Huang 2016). Employing STR and CAT technologies during our learning activities enabled students from different cultures who spoke different native languages to communicate with each other without language barriers. As a result, their sense of connectivity increased. When we applied STR and CAT technologies together, they complemented each other. STR technology is a valuable tool for education because it synchronously generates text streams from speech input, and students can use these texts to confirm what is being said, to attain a better understanding of learning material, to take notes, and to complete homework (Hwang et al. 2012; Kuo et al. 2012; Kheir and Way 2006; Ranchal et al. 2013; Ryba et al. 2006). CAT extends applications of STR by translating communication content into many different languages, so translated STR-texts can also be helpful to those who cannot speak the language in which communication content is delivered. STR extends applications of CAT by providing a hands-free input method. Shadiev and Huang (2016) argued that STR input is more fun, convenient, and faster than typing. When considering

mobile technology, which is an important everyday tool not just in our lives but also in the educational milieu, STR technology is a very helpful method by which to input cross-cultural communication content since typing on mobile devices is not as easy to accomplish.

The second point is how this present study differs from our earlier research, e.g. Shadiev and Huang (2016), Shadiev et al. (2017a), and Shadiev et al. (2017b). Shadiev et al. (2017) explored the effectiveness of STR applications on learning performance, which is a different research direction from that of this study. Shadiev et al. (2016) focused on facilitating cross-cultural understanding with project-based collaborative learning in an online environment but without applications of STR and CAT. Both the study of Shadiev and Huang (2016) and the present study focused on applications of STR and CAT to facilitate cross-cultural understanding. However, the participants in this study communicated in ten different languages, so their communication was multi-lingual, whereas communication among the participants in the previous research was bi-lingual (Shadiev and Huang 2016). In addition, our participants were from thirteen countries and therefore, represented multiple cultures, whereas the participants in Shadiev and Huang (2016) were only from two countries, thus representing only two cultures. Therefore, the current exploration focuses on whether applications of STR and CAT are effective in terms of supporting and facilitating the cross-cultural understanding of participants from several countries interacting about their culture and traditions in multiple languages. In addition, previous research (Shadiev and Huang 2016) explored the accuracy rates of STR and CAT in two languages only, but we focused on ten. Furthermore, we explored issues associated with STR and CAT processes and how they can be solved, which was not the focus in Shadiev and Huang (2016).

Conclusions

Our results suggest that using STR and CAT systems are useful for simple, daily life communication in most languages. However, when considering communication on complex and advanced topics, STR and CAT produced more accurate content only for widely used languages that are similar to English (e.g., Russian, French, and Spanish). At this time, STR and CAT should not be considered a well-rounded professional translation mechanism from voice input since they have limitations that need to be considered. As stated by Scigliano (2010), the translations might still lack correct grammar and punctuation in some cases. Following this notion, Google® mentioned: "Even today's most sophisticated software, however, does not approach the fluency of a native speaker or possess the skill of a professional translator. Automatic translation is very difficult, as the meaning of words depends on the context in which they are used. While we are working on the problem, it may be some time before anyone can offer human quality translations." Recent evidence suggests that the accuracy rate of STR and CAT is improving as time goes on (Simonite 2016). In the case of STR, it was about 88% in 2006 (Kheir and Way 2006), about 90% in 2012 (Hwang et al. 2012; Kuo et al. 2012), and more than 90% in 2015 (Shadiev and Huang 2016) and 2016 (i.e., the present study). In the case of CAT, it was about 74–89% in 2015 (Shadiev and Huang 2016) and was more than 90% in 2016 (i.e., the present study). The results of this study showed that translated communication content was accurate enough across the languages so that the participants found it comprehensible and useful for communication and understanding foreign traditions. Meanwhile, there is still

room for improvement since the accuracy rate is not yet ideal (Barrachina et al. 2009; Fountain and Fountain 2009; Mellebeek et al. 2005). Therefore, our results can be useful for educators and researchers who plan to apply STR and CAT to support multilingual communication during a learning process in the near future. Based on our results, they will be able to make the design of learning activities with STR and CAT support more efficient. Notably, the workarounds reported by our participants can be useful to improve the accuracy rate of STR and CAT. It is also important to note that while the experiences of our participants relate to the usage of STR and CAT systems for ten languages only, many of the issues and possible solutions discussed in this paper are relevant to other languages as well. In addition, we suggest that applications of STR and CAT systems can be useful not only for cross-cultural learning programs but for other programs as well. For example, students may learn such subjects as geography by exchanging related learning information with students in different countries using multilingual communication.

Some limitations of this study must be acknowledged. One limitation relates to the small sample size, short term exposure to applications of STR and CAT during the cross-cultural learning activity, and the limited number of languages that participants spoke as their mother tongue. This issue may limit generalization of our results to the wider population. Therefore, researchers need to consider these limitations and address them in future studies. Another limitation relates to potential biases in the validity of the accuracy rate. Participants' ability to understand English enabled them to compare translations with original texts in their mother tongue. This, in turn, could affect implementation of the workarounds discussed earlier related to using these technologies more effectively. Therefore, continuous usage of the workarounds throughout this study could lead to increased accuracy rate of texts generated by the technology. In the future, we will focus on scenarios where when students speak, the technology generates texts in the native languages of their partners, which the students do not understand. We will investigate whether the accuracy rate will be the same or different. In future studies, researchers may also wish to explore what language characteristics influence the accuracy rates of STR and CAT with regard to recognizing some culture-related words or terms. It currently is not clear whether the accuracy rate of STR and CAT was influenced by the fact that words or terms for some traditions are more easily recognized by STR and CAT than those from others.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

Extracts from posts of the participants in their mother tongue and English translation.
    Step 1
    ID9 हेलो एवरीवन मेरा नाम XXX है मैं इंडिया से हूं में अजमेर सिटी से हूं मुझे/Hello Everyone. My name is XXX. I am from India. I am from Ajmer City.
    ID 12 Xin chào các bạn. Tôi tên là XXX. Tôi đến từ Việt Nam. Hiện tại Tôi đang học tại trường đại học./Hello friends. My name is XXX. I'm from Vietnam. Currently I am studying at the university.
    ID19 Сайн байна уу. Миний нэрийг XXX гэдэг. Би бол Монгол хүн./Hello. My name is XXX. I am a Mongolian.
    Step 2

ID5 台灣是一個充滿許多廟宇的國家因為許多人會去廟宇祈福,目的也有所不同, 像是求平安、考試順利、交女朋友、生小孩,因此各個廟宇所供奉的神明也有所 不同,像是海之神、土地神、愛神、地獄神。/Taiwan is a country filled with many temples because many people will go to the temple to pray, the purpose is also different, like to pray for peace, the examination successfully, a girlfriend, a baby, and therefore each of the temples dedicated to the gods are different, like the god of the sea, land God, love God, God of Hell.

ID16 La tradicion de las alfombras de aserrin en Honduras ha sido una tradicion de mas de 50 anos en el pais. Ningún versículo de La Biblia las describe; sin embargo, las coloridas alfombras de aserrín se han convertido en elementos indispensables para conmemorar la pasión, muerte y resurrección de Jesucristo./The tradition of sawdust carpets in Honduras has been a tradition of over 50 years in the country. No verse of the Bible describes; however, colorful sawdust carpets have become indispensable elements to commemorate the passion, death and resurrection of Jesus Christ.

Step 3

ID7 Bonjour tout le monde. Nous voici déjà à l'étape 3 du projet et pour dire lequel culture nous a beaucoup marqué. Pour moi c'est la culture de l'Inde qui est en rapport avec le festival HOLI. Mon choix s'est porté sur celle-là à cause des couleurs de l'événement./ Hello everybody. We are already in stage 3 of the project and to say which culture has greatly influenced us. For me it is the culture of India which is related to the HOLI festival. My choice fell on that one because of the color of the event.

ID20 Ang aking piniling maranasan ay ang kultura ng mga Mongolians na tinawag na "tea dumplings". Ito ay isa sa kanilang pangaraw-araw na pagkain na kanilang tinatawag na "buuz". Ang sabaw ay binubuo ng tsaa./My choice is to experience the culture of the Mongolians called "tea dumplings". It is one of their daily meals, which they called "buuz". The broth is made up of tea.

Step 4

ID2 這個節日非常有趣。在亡靈節,墨西哥人不哭泣,他們開心慶祝。在墨西哥的 文化裡,死亡並不是一件悲傷的事,而是一種生命的循環 。在這天,墨西哥人慶祝能和 過世的親人相聚,並且好好得活著 。這個觀念和華人非常不一樣 。/This festival is very interesting. In the Day of the Dead, Mexicans do not cry, they are happy to celebrate. In Mexican culture, death is not a sad thing, but a life cycle. On this day, Mexicans celebrate the death of loved ones and can get together and have a good living. This concept and the Chinese are very different.

ID10 Saya memilih tradisi dari Usbekistan – "morning pilaf". Tradisi ini sangat menarik karena kita harus berkumpul bersama dan makan bersama di pagi hari. Tradisi ini diadakan ketika ada acara khusus pada hari tersebut. Seperti acara pernikahan, dan kelahiran./I chose the tradition of Uzbekistan - "morning pilaf". This tradition is very interesting because we have to get together and eat together in the morning. This tradition is held when there are special events on that day. Such as weddings and births.

ID21 Я попробовал традицию вручения красного конверта. Красный цвет конверта символизирует удачу и предназначается для отпугивания злых духов. Обычно люди получают красные конверты во время праздников или специальных мероприятий. Мамы получают конверты на день мам, молодожены получают конверты на свадьбе, а дети или старики на китайский новый год и так далее. Обычно в красный конверт кладутся деньги. Эта традиция очень давняя. Много лет назад когда деньги были в виде монет, их завязывали вместе красной ниткой. Согласно легенде, злые духи дотрагивались до голов детей. После этого дети становились серьезно больными. Монеты клали над головой детей, что

отпугивало злых духов. Позже, появились банкноты. Они заменили монеты. Люди стали использовать красные конверты вместо нити./I tried the tradition of presenting a red envelope. Red color of the envelope symbolizes good luck and is meant to scare away evil spirits. Usually, people receive red envelopes during holidays or special events. Moms get envelopes on the day of mothers, newlyweds receive envelopes at a wedding, and children or old men on the Chinese New Year and so on. Typically, in a red envelope money placed. This tradition is very old. Many years ago, when money was in the form of coins, they are tied with red thread. According to legend, evil spirits touched the heads of children. After that, the children became seriously ill. Coins placed over the head of children that scare away evil spirits. Later, there were banknotes. They replaced the coin. People began to use red envelopes instead of thread.

References

Aiken, M., & Balan, S. (2011). An analysis of Google Translate accuracy. *Translation Journal, 16*(2). Retrieved from http://translationjournal.net/journal/56google.htm.

Aparicio, M., Bacao, F., & Oliveira, T. (2016). Cultural impacts on e-learning systems' success. *The Internet and Higher Education, 31,* 58–70.

Barrachina, S., Bender, O., Casacuberta, F., Civera, J., Cubel, E., Khadivi, S., et al. (2009). Statistical approaches to computer-assisted translation. *Computational Linguistics, 35*(1), 3–28.

Bentley, J. P. H., Tinney, M. V., & Chia, B. H. (2005). Intercultural Internet-based learning: Know your audience and what it values. *Educational Technology Research and Development, 53,* 117–127. https://doi.org/10.1007/BF02504870.

Chase, M., Macfadyen, L., Reeder, K., & Roche, J. (2002). Intercultural challenges in networked learning: Hard technologies meet soft skills. *First Monday*, 7(8). Retrieved from https://firstmonday.org/ojs/index.php/fm.

Çiftçi, E. Y. (2016). A review of research on intercultural learning through computer-based digital technologies. *Journal of Educational Technology & Society, 19*(2), 313–327.

Curtis, D. D., & Lawson, M. J. (2001). Exploring collaborative online learning. *Journal of Asynchronous Learning Networks*, 5(1), 21–34.

ElShiekh, A. A. A. (2012). Google translate service: transfer of meaning, distortion or simply a new creation? An investigation into the translation process & problems at google. *English Language and Literature Studies, 2*(1), 56–68.

Ertmer, P. A., Newby, T. J., Liu, W., et al. (2011). Students' confidence and perceived value for participating in cross-cultural wiki-based collaborations. *Educational Technology Research and Development, 59,* 213–228. https://doi.org/10.1007/s11423-011-9187-4.

Fountain, A., & Fountain, C. (2009). A New Look at Translation: Teaching tools for language and literature. *Empowerment through Collaboration: Dimension, 2009,* 1–15.

Godwin-Jones, R. (2011). Emerging technologies: Mobile apps for language learning. *Language Learning & Technology*, 15(2), 2–11.

Gudykunst, W. B., Ting-Toomey, S., & Chua, E. (1988). *Culture and interpersonal communication.* Thousand Oaks: Sage Publications.

Hermet, M., & Désilets, A. (2009). Using first and second language models to correct preposition errors in second language authoring. In *Proceedings of the Fourth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 64–72). Association for Computational Linguistics.

Holm, J. (1982). *Central American English* (Vol. 2). Amsterdam: John Benjamins Publishing.

Hwang, W. Y., Shadiev, R., Kuo, T. C. T., & Chen, N. S. (2012). A study of speech to text recognition and its effect to synchronous learning. *Journal of Educational Technology & Society, 15*(1), 367–380.

Jaakkola, H., Heimbürger, A., & Linna, P. (2010). Knowledge-oriented software engineering process in a multi-cultural context. *Software Quality Journal, 18,* 299–319.

Jenkins, H., Purushotma, R., Weigel, M., Clinton, K., & Robison, A. J. (2009). *Confronting the challenges of participatory culture: Media education for the 21st century.* Cambridge, MA: MIT Press.

Kheir, R., & Way, T. (2006). Improving speech recognition to assist real-time classroom note taking. In Proceeding of the rehabilitation engineering and assistive technology society of North America conference pp. 1–4.

Kincaid, D. L. (1979). *The convergence model of communication.* Honolulu: East-West Center.

Kohlberg, L. (1984). *The psychology of moral development. The nature and validity of moral stages.* San Francisco: Harper & Row.

Kuo, T. C. T., Shadiev, R., Hwang, W. Y., & Chen, N. S. (2012). Effects of applying STR for group learning activities on learning performance in a synchronous cyber classroom. *Computers & Education, 58*(1), 600–608.

Labov, W. (2011). *Principles of linguistic change, cognitive and cultural factors*. New Jersey: Wiley.

Liu, Y. (2007). Designing quality online education to promote cross-cultural understanding. In A. Edmundson (Ed.), *Globalized e-learning cultural challenges* (pp. 35–59). Hershey, PA: IGI Global. https://doi.org/10.4018/978-1-59904-301-2.ch003.

Mellebeek, B., Khasin, A., Owczarzak, K., Van Genabith, J., & Way, A. (2005). Improving online machine translation systems. *In Proceedings of the X MT Summit* (pp. 290–297). Phuket, Thailand.

Moran, R. T., Abramson, N. R., & Moran, S. V. (2014). *Managing cultural differences*. Routledge: CRC Press.

Omar, H., Embi, M. A., & Yunus, M. M. (2012). ESL learners' interaction in an online discussion via Facebook. *Asian Social Science*, *8*(11), 67–74.

Osman, G., & Herring, S. C. (2007). Interaction, facilitation, and deep learning in cross-cultural chat: A case study. *The Internet and Higher Education, 10*(2), 125–141.

Ranchal, R., Taber-Doughty, T., Guo, Y., Bain, K., Martin, H., Robinson, J., et al. (2013). Using speech recognition for real-time captioning and lecture transcription in the classroom. *IEEE Transactions on Learning Technologies, 6*(4), 299–311.

Reynolds, B. L. (2013). A web-based ELF writing environment as a bridge between academic advisers and junior researchers: A pilot study. *British Journal of Educational Technology, 44*(3), E77–E80.

Reynolds, B. L. (2015). Helping Taiwanese graduate students help themselves: Applying corpora to industrial management English as a foreign language academic reading and writing. *Computers in the Schools, 32*(3–4), 300–317.

Rogers, P. C., Graham, C. R., & Mayes, C. T. (2007). Cultural competence and instructional design: Exploration research into the delivery of online instruction cross-culturally. *Educational Technology Research and Development, 55,* 197–217. https://doi.org/10.1007/s11423-007-9033-x.

Ryba, K., McIvor, T., Shakir, M., & Paez, D. (2006). Liberated learning: analysis of university students' perceptions and experiences with continuous automated speech recognition. *Journal of Instructional Science and Technology, 9*(1), 1–19.

Scigliano, D. A. (Ed.). (2010). *Telementoring in the K-12 classroom: online communication technologies for learning: online communication technologies for learning*. Hershey: IGI Global.

Shadiev, R., & Huang, Y. M. (2016). Facilitating cross-cultural understanding with learning activities supported by speech-to-text recognition and computer-aided translation. *Computers & Education, 98,* 130–141.

Shadiev, R., Huang, Y. M., & Hwang, J. P. (2017a). Investigating the effectiveness of speech-to-text recognition application on learning performance, attention, and meditation. *Educational Technology Research and Development*. https://doi.org/10.1007/s11423-017-9516-3.

Shadiev, R., Hwang, W. Y., Huang, Y. M., & Liu, C. J. (2016). Investigating applications of speech to text recognition for face to face seminar to assist learning of non-native English participants. *Technology, Pedagogy and Education, 25*(1), 119–134.

Shadiev, R., Wu, T. T., & Huang, Y. M. (2017b). Enhancing learning performance, attention, and meditation using a speech-to-text recognition application: Evidence from multiple data sources. *Interactive Learning Environments, 25*(2), 249–261.

Simonite, T. (2016). Google's new service translates languages almost as well as humans can. MIT Technology Review. Retrieved from: https://www.technologyreview.com/s/602480/googles-new-service-translates-languages-almost-as-well-as-humans-can/?utm_campaign=socialflow&utm_source=facebook&utm_medium=post.

Talalakina, E. V. (2010). Fostering cross-cultural understanding through e-learning: Russian–American forum case-study. *International Journal of Emerging Technologies in Learning, 5*(3), 42–46.

Tobin, A. (2015). Is Google Translate good enough for commercial websites?: A machine translation evaluation of text from English websites into four different languages. *Reitaku Review, 21,* 94–116.

Tu, C. H. (2004). *Online collaborative learning communities: Twenty-one designs to building an online collaborative learning community*. Westport, CT: Greenwood.

Wu, T. T. (2014). The use of a mobile assistant learning system for health education based on project-based learning. *Computers, Informatics, Nursing, 32*(10), 497–503.

Wu, T. T. (2015). A learning log analysis of an English-reading e-book system combined with a guidance mechanism. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2015.1070272.

Yamazaki, Y., & Kayes, D. C. (2004). An experiential approach to cross-cultural learning: A review and integration of competencies for successful expatriate adaptation. *Academy of Management Learning & Education*, *3*(4), 362–379.

Yang, S. C., & Chen, Y. J. (2007). Technology-enhanced language learning: A case study. *Computers in Human Behavior*, *23*(1), 860–879.

Rustam Shadiev is a professor at the School of Education Science, Nanjing Normal University, China. His research interests include online learning and its technology, educational applications of multimedia technology, and computer supported EFL/ESL learning.

Ting-Ting Wu is an assistant professor at the Graduate School of Technological and Vocational Education, National Yunlin University of Science and Technology, Taiwan. Her research interests are mobile and ubiquitous learning, information technology-applied instructions and intelligent learning systems.

Ai Sun is currently a postdoctoral researcher at the Department of Engineering Science, National Cheng Kung University, Taiwan. Her research interests include e-learning and computer assisted language learning.

Yueh-Min Huang is a Chair Professor at the Department of Engineering Science, National Cheng Kung University, Taiwan. His research interests focus on e-Learning, multimedia communications, artificial intelligence, and embedded systems.