PLACE-BASED FOREIGN LANGUAGE LEARNING
VIA TELEPRESENCE ROBOTS

A Dissertation in
Learning, Design, and Technology

by

Jian Liao

© 2020 Jian Liao

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2020
The dissertation of Jian Liao was reviewed and approved by the following:

Simon Hooper  
Professor of Learning, Design, and Technology  
Dissertation Advisor  
Chair of Committee

Heather Zimmerman  
Associate Professor of Learning, Design, and Technology

Tyler Hollett  
Assistant Professor of Learning, Design, and Technology

Xiaofei Lu  
Professor of Applied Linguistics

Susan Land  
Professor of Learning, Design, and Technology  
Director of Graduate Studies, Department of Learning Performance and Systems
ABSTRACT

It has been found that FL learners, who only focus on language forms such as vocabulary, pronunciation, and grammar, may not be able to use the learned linguistic knowledge appropriately in real conversational situations. From the perspective of communicative language teaching, FL learners need to be exposed to and engaged in authentic language use in the target countries to improve their communicative competence, i.e., the ability to communicate with others in spontaneous interaction in given social contexts. However, only a few FL learners are able to travel to the target language countries for language learning purposes due to the expensive costs of travelling and boarding. This three-article dissertation proposes to use telepresence robots to engage FL learners in authentic conversation in the target countries. The first article investigates the perceived benefits and challenges of using Romo, a small telepresence robot, to mediate the conversations between FL learners in China and a native English speaker on an American university campus. The second article frames this learning form as telepresence-place-based foreign language learning (TPFLL) and proposes three design principles based on the findings of the first article. A case study using KeeBot, a bigger telepresence robot, was conducted to examine the proposed design principles at an arboretum at the American university campus. The third article further investigates how different places as well as the people and social activities at these places influence the use of target language from a perspective of geosemiotics. The findings of the three articles are expected to help demonstrate the learning process of TPFLL and provide insights to guide the design of TPFLL learning activities and the future design of telepresence robots.

Keywords: telepresence robots, robot-assisted language learning, communicative competence, place-based learning, communicative-language teaching
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>PREFACE</td>
<td>x</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>xiii</td>
</tr>
<tr>
<td>CHAPTER 1 : INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Why Place Matters</td>
<td>1</td>
</tr>
<tr>
<td>Teaching Methods to Improve Communicative Competence</td>
<td>7</td>
</tr>
<tr>
<td>Technologies Applied in Language Learning and the Pushes for Authenticity</td>
<td>12</td>
</tr>
<tr>
<td>Telepresence Technologies and Telepresence Robots</td>
<td>14</td>
</tr>
<tr>
<td>Research Purpose and Structure of the Dissertation</td>
<td>18</td>
</tr>
<tr>
<td>Conclusion</td>
<td>19</td>
</tr>
<tr>
<td>References</td>
<td>20</td>
</tr>
<tr>
<td>CHAPTER 2 : EXPLORING THE AFFORDANCES OF TELEPRESENCE ROBOTS IN FOREIGN LANGUAGE LEARNING</td>
<td>27</td>
</tr>
<tr>
<td>Abstract</td>
<td>28</td>
</tr>
<tr>
<td>Introduction</td>
<td>29</td>
</tr>
<tr>
<td>Theoretical Background</td>
<td>30</td>
</tr>
<tr>
<td>Research Questions</td>
<td>32</td>
</tr>
<tr>
<td>Methodology</td>
<td>33</td>
</tr>
<tr>
<td>Findings</td>
<td>36</td>
</tr>
<tr>
<td>Discussion</td>
<td>44</td>
</tr>
<tr>
<td>Conclusion</td>
<td>47</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>47</td>
</tr>
<tr>
<td>Notes</td>
<td>47</td>
</tr>
</tbody>
</table>
# References

- **CHAPTER 3: TELEPRESENCE-PLACE-BASED FOREIGN LANGUAGE LEARNING AND ITS DESIGN PRINCIPLES**  ......................................................... 52  
  - Abstract ............................................................................................................................ 53  
  - Introduction ....................................................................................................................... 54  
  - Literature review .............................................................................................................. 55  
  - Framing the study as TPFLL ............................................................................................. 57  
  - Methodology .................................................................................................................... 61  
  - Findings ............................................................................................................................. 67  
  - Discussion .......................................................................................................................... 81  
  - Conclusion ......................................................................................................................... 83  
  - References ......................................................................................................................... 84

- **CHAPTER 4: GEOSEMIOTIC ANALYSIS OF TELEPRESENCE-PLACE-BASED FOREIGN LANGUAGE LEARNING** ................................................................. 90  
  - Abstract ............................................................................................................................. 90  
  - Introduction ....................................................................................................................... 91  
  - Literature Review ............................................................................................................ 92  
  - Methodology ..................................................................................................................... 99  
  - Findings ............................................................................................................................. 104  
  - Discussion ......................................................................................................................... 118  
  - Conclusion ......................................................................................................................... 125  
  - References ......................................................................................................................... 125

- **CHAPTER 5: DISCUSSION AND CONCLUSION** ................................................................... 130  
  - Summary of the Three Articles ....................................................................................... 130  
  - Discussions in Theoretical Topics ................................................................................... 132
LIST OF TABLES

Table 2.1. Demographic Information of Remote Participants .................................................. 33
Table 2.2. Themes and codes .................................................................................................... 37
Table 3.1. Demographic information of the learners ................................................................. 62
Table 3.2. Design principles and specific design of the study .................................................. 64
Table 3.3. Survey results on perceived learning outcomes ....................................................... 71
Table 4.1. Demographic Information of the Learners ............................................................... 99
Table 4.2. Overarching findings of the geosemiotics analysis ................................................ 105
Table 4.3. Presenting pumpkins at an exhibit/display space near the arboretum gate .......... 107
Table 4.4. Micro-role-playing for expanding discourse to sociocultural aspects of an object ... 110
Table 4.5. Etic engagement with baseball game at a special use space .................................. 112
Table 4.6. Emic engagement into a picnic at a special use space ........................................... 113
Table 4.7. Leading the learner at the passage space between the witness tree and the pavilion 115
Table 4.8. Talking while walking from the witness tree to the pavilion ................................ 116
Table 5.1. Three overlapping research areas relevant to PBL, FL, and telepresence technologies .................................................................................................................................. 142
LIST OF FIGURES

Figure 1.1. 2D Telepresence approach using a fixed camera and monitor ........................................... 15
Figure 1.2. 2.5D Telepresence approach using the “TriPad” ............................................................... 16
Figure 1.3. Telepresence robots .......................................................................................................... 17
Figure 2.1. Communication via a telepresence robot ............................................................................. 34
Figure 2.2. The Interface on the screen of the robot .............................................................................. 35
Figure 2.3. The route of the activity ..................................................................................................... 35
Figure 2.4. Learning the meaning of the word quad .............................................................................. 41
Figure 3.1. Telepresence robots in foreign language learning ............................................................... 54
Figure 3.2. Focal spots in the arboretum and tour route ....................................................................... 63
Figure 3.3. Communication via the telepresence robot ........................................................................ 63
Figure 3.4. Conversation about the sundial ......................................................................................... 67
Figure 3.5. Pumpkins and Halloween ................................................................................................. 69
Figure 3.6. Potential learning sites ...................................................................................................... 71
Figure 3.7. Two parts of a wedding in the United States ..................................................................... 73
Figure 3.8. Embodied interaction ....................................................................................................... 75
Figure 3.9. Whiteboard ...................................................................................................................... 76
Figure 4.1. Features and the route at the Arboretum .......................................................................... 100
Figure 4.2. Communication with FL learners in China via a telepresence robot ................................. 101
Figure 4.3. Exhibit/display space near the gate ................................................................................... 106
Figure 4.4. Carrying the robot at a passage place ............................................................................... 116
Figure 4.5. Switching from exhibit/display space to special use space at the sundial ...................... 117
Figure 4.6. Switching from a passage space to an exhibit/display space near a birch tree........ 118

Figure 5.1. Relations among terms of place-based learning, FL learning, and Telepresence Technologies.......................................................... 142
The basis of this research is rooted in my struggle to improve my English-speaking skills. As a foreign language learner, I started to learn in middle school, and I spent more than ten years reciting vocabulary, reading passages, and memorizing grammar. However, I found that I couldn’t say a complete sentence fluently while speaking with others in English, even though I had decent reading skills. It seems that the way to improve speaking is so different from skills in reading and those in other disciplines such as math and physics. Some of my friends told me that the best way to improve speaking is to study abroad in a country where the language is mainly used. Thus, I found my way to the Learning, Design, and Technology program at Penn State University partly because I am passionate about building technical environments for education and partly because I hoped I could improve my speaking further.

Although I still struggled especially in the first several years, I did feel my speaking skills was growing gradually through the years of my Ph.D. studies in a way I had never experienced before. I gained deeper insight into the way to improve speaking from the courses I chose from the Department of Applied Linguistics including Dr. Xiaofei Lu’s Second Language Acquisition. I found the method of improving speaking skills was named language acquisition, which is so different from the method of reciting vocabulary and grammar in language learning. I also learned that place could play an important role in education especially from the studies of Dr. Susan Land and Dr. Heather Zimmerman in my program. I felt that place could also play an important role in language acquisition since it could explain why my friends said the best way to improve speaking was to study or live in a country where the language was used.

I then started to consider if I could find a way to have more foreign language learners across the globe gain access to the environments abroad with the support of technology. A spark
emerged when I recalled the movie, Avatar, directed by James Cameron. In that movie, the hero is a veteran with physically disability. However, he is able to connect his mind to an alien body (an avatar) and then finish his mission remotely on an alien planet. He also learned Na’vi, the alien language, through communicating with his alien girlfriend and other aliens in their daily life. Although this is a sci-fi movie, the way for the hero to acquire another language struck me. Could we find an equivalent technology in our reality to realize this idea? I then recalled a TV show, Big Bang. In one episode, Sheldon, one of the geek heroes, controls a machine, a telepresence robot, to walk outside his bedroom. Yes, I thought that would be a satisfactory answer to bring the sci-fi story to reality.

The topic of this dissertation was confirmed after I discussed it with my advisor and the committee members. I must thank them and their guidance for shaping the theoretical frameworks and research methodologies of this research. Also, I am thankful for their support, which helped me explore this topic further. Moreover, I would like to thank Dr. Rayne Sperling, who gave me the idea of using a three-article dissertation format since I had published two journal articles as the first author and the principal investigator of this project during the exploration of this topic.

The first article has been published in the Journal of Language Learning and Technology. My contributions are conceptualizing the research idea, making the literature review, creating the research design, recruiting participants, organizing the learning activities, analyzing data, and writing the first draft. Dr. Lu helped me refine the theoretical framework, especially on authenticity. He also made edits on the whole draft and made changes based on reviewer feedback. I used the research methods I learned from Dr. Zimmerman’s qualitative research methods class and Dr. Joan Kelly Hall’s conversation analysis class, who is a professor in
Applied Linguistics at The Pennsylvania State University. Zhi Zhou, my wife, helped me recruit learner participants, videotape the learning activities, and transcribe interviews.

The second article has been published in the Journal of Computer-Assisted Language Learning. My contributions are conceptualizing the research idea, writing literature review, creating the research design, recruiting participants, organizing the learning activities, analyzing data, and writing the first draft. Dr. Lu helped me refine the theoretical framework. He also made edits on the whole draft and made changes based on reviewer feedback. Dr. Katherine Masters and Dr. Jaclyn Dudek helped me refine the theoretical framework and made edits on the draft. Dr. Masters also made the listening materials used in the learning activities. Zhi Zhou helped me recruit learner participants, videotape the learning activities, and transcribe interviews. I also partly used the methodology I learned from Dr. Heather Zimmerman’s design-based research class.

The third article hasn’t been published yet. My contributions are writing literature review, creating the research design, recruiting participants, organizing the learning activities, analyzing data, and writing the first draft. Dr. Hollett guided me to conceptualize the research idea of geosemiotics. Also, Dr. Hollett and my advisor, Dr. Hooper, held several data sessions with me to discuss how to connect the video data to the theories. Also, I want to thank all of my committee members for providing your professional and valuable comments especially on Chapters 4 (the third article) and 5.

I appreciate all the help from my advisor, my academic committee members, and my colleagues for their excellent contributions to this project and the publications. I would not have finished them without your support.
ACKNOWLEDGEMENTS

This journey of completing this doctoral study would never have been completed without the guidance and support of my mentors, family, and friends. First, I would like to thank my advisor, Dr. Simon Hooper, who has always been supportive throughout my study. His rigorousness and generousness walked me every step of the way from coursework to graduation. I am so grateful to have the opportunity to be his student.

I must also thank my committee members, Dr. Heather Zimmerman, Dr. Tyler Hollett, and Dr. Xiaofei Lu. They have been very generous with their time and professional guidance. Dr. Zimmerman helped me shape the theoretical framework and research methods of the study. Dr. Hollett polished my research framework with the perspective of Geosemiotics. Dr. Lu provided me with tremendous professional advice on Applied Linguistics and Computer Assisted Language Learning since the beginning of this study.

It was a privilege to work with my fellow researchers, Dr. Jaclyn Dudek and Dr. Katherine Masters, who are the best team members I could have in research and writing. Although I cannot disclose the names of the other participants in this study, I am grateful to have all the teachers and students who devoted their time to the study with passion.

Special thanks go to Dr. Rayne A. Sperling, who suggested me using three-article dissertation format that perfectly matches my doctoral study. She has been always kind and considerate for all graduates who work or worked with her. I also appreciate Dr. Joan Kelly Hall, Dr. Stephen D. Looney, and their team who provided professional advice on conversation analysis in this study. In addition, the thoughtful support from the Arboretum of Penn State, especially from Dr. Kim C. Steine, Kelly Oleynik, Kate Reeder, and Shari Ederson has always been very appreciated.
During the long journey of pursuing my doctoral degree at Penn State, my cohorts, Shulong Yan, Yong Ju Jung, and Jaclyn Dudek were the most incredibly helpful friends I could ever have imagined. I will always miss the wonderful years we studied together in the LDT program at Penn State. I also appreciate my friends Yingliang He and Tao Xie, who gave me help with recruiting participants and finding professional editors.

Lastly, I want to thank my family. My parents have been incredibly supportive of me studying ten thousand miles away from home. My wife Zhi Zhou is the person who would always be there for me no matter how hard this journey was. She is also a team member of this study who helped with participant recruiting and data collecting. I have to mention my two-year-old daughter who was born during this long journey of doctoral study. Although she has been very active and usually destroys everything she can reach, she is definitely the best gift I have ever had in my life.
CHAPTER 1: INTRODUCTION

This study explores a new approach to learning a foreign language (FL) virtually in a target-language country using telepresence robots. A telepresence robot provides learners with the experience of being present at a remote location (Bell et al., 2016; Minsky, 1980). As an introduction to this study, this chapter will explain why learning in the daily-life settings in a target-language country is important for FL acquisition. It will also describe what telepresence robots are in detail and why it is necessary to use them to support FL learning. Finally, this chapter will introduce the research focus of this study.

Accordingly, this chapter contains six sections. The first section explains why it is important for FL learners to learn FLs in real-life situations. The second and third sections briefly introduce current teaching methods and technologies in support of FL learning in real-life situations and consider their limitations. The fourth section elaborates on the concept of telepresence and three generations of telepresence technologies. The fifth section details this study’s research purpose and introduces the following three chapters of this study. Finally, the sixth section summarizes the content of this chapter.

Why Place Matters

FL learners, especially in Asian countries, traditionally learn a target language with grammar-focused and teacher-centered methods in classrooms (Talebi et al., 2015; Tan, 2016). Specifically, an instructor reads and explains a passage in a textbook, sentence by sentence, to learners in a class setting. The grammar of a language will be taught as it occurs in the passages or as an independent grammar course. Hino (1988) also stated that the predominant way of

\textsuperscript{1} Target language is a term that has been widely used in Applied Linguistics (e.g., Ozverir & Herrington, 2011). The term refers to the foreign language that a FL learner is learning. A target-language country or community is a one where the language is mainly used.
teaching English in Japan was the Yakudosu method, in which an English sentence is first translated into Japanese word-by-word, and then the translated Japanese words are re-ordered as a sentence to match the grammar of Japanese.

Although the grammar-focused methods for classroom teaching can prepare FL learners well for standardized foreign language tests, scholars (e.g., Larsen-Freeman, 2013; Thorne et al., 2009) have argued that traditional classroom settings isolate the learners from the outside world and that the abstract content in textbooks makes little reference to the actual context of the spoken language (Holden & Sykes, 2011). As a result, the traditional teaching methods in classrooms are not adequate for FL learners to build well-rounded communicative competence, i.e., applying the language skills in real communicative situations (Savignon, 2018). The term communicative competence is first proposed by Hymes (1972) to provide a much broader view of language use in real situations. FL speakers need to know not only grammatical structures, but also the norms to use the language appropriately in any given social context. Savignon (2018) expanded the meaning of communicative competence from an interactive view to include the proficiency of using a FL in spontaneous interaction.

Based on the above definitions of communicative competences, Wei (2012) investigated the communicative competence of four international students in an American university. The four students were all from China and had studied English as a foreign language in schools for no less than 10 years before arriving in the USA. The study found that all participants had weak communicative competence, and this lack of communicative competence had been one of the biggest barriers preventing the participants from achieving both academic success and successful integration into non-academic lives in the USA. The participants reported in interviews that their self-consciousness of their weak spoken-English skills and unfamiliarity with the communicative
conventions in the USA made them lack confidence and become hesitant to engage in classroom activities (e.g., group discussion). Consequently, they had to rely more on their personal efforts such as reading PowerPoint slides and online materials after the class to compensate for their insufficient engagement in the class. Even so, their peers and instructors might consider them as unqualified or unprepared to be legitimate contributors to the class because of their passivity and silence. Moreover, the negative experiences with communication led to a sense of exclusion from the community of local students. The participants rarely contacted their American peers in daily life and tended to socialize only with their Chinese peers, despite the participants’ expectations of engaging in an extensive intercultural experience in the USA.

Many scholars (e.g., Vanegas, 2008; Volosinov, 1973) consider that the unskillful communicative competence of the FL learners is caused by an insufficient exposure to and engagement in the actual conversation situations in daily life. In native language acquisition, children spend several years listening to the language and practicing with their parents/caregivers, friends, and/or other people in various real-life settings: at home, shopping, on trips to museums, and in potentially all the places they find themselves (Vanegas, 2008). So, it is natural for native language speakers to gain precise and adequate meanings of words, phrases, and sentences and then use them in real life situations appropriately (Vanegas, 2008). However, most FL learners barely have the experience of using the FL in their lives since the it is not the language used in the FL learners’ communities.

De Guerrero (2006) and Hino (1988) further explained why sufficient exposure to and engagement in real communicative situations in daily life are important for FL learners’ communicative competence. Based on studies involving a large number of sampled Spanish English-as-a-second-language (ESL) learners, De Guerrero (2006) claimed that the FL learners
with decontextualized FL instruction (e.g., reciting vocabulary word by word, learning grammar explicitly) will formulate thoughts first in the mother tongue (L1) and then translate their L1 thoughts into the second language (L2). Hino (1998) also stated that FL learners with the Yakudosu instruction tend to employ a similar translation strategy in listening and speaking: for example, translating the Japanese words in a sentence in their mind word-by-word first and then instantly reordering the translated words based on the English grammar.

However, for a foreign language learner, translating a L1 to a L2 in their mind is a slow, tiring, imprecise, and ineffective process (Hino, 1998). Many words in the mother tongue do not have an exact meaning or have multiple meanings in the FL. For example, a Chinese word ‘香’ is used to describe something that smells good including flowers and food. Similar words in English are ‘fragrant’ and ‘aromatic’. However, both ‘fragrant’ and ‘aromatic’ are mainly used to describe the odor of flowers, rather than the smell of appetizing food. When a Chinese English-as-a-foreign-language (EFL) speaker uses the translation method for speaking, he or she might struggle to translate the Chinese sentence ‘这菜 (the food in the dish/this dish) (is) 真 (very) 香 (smells good)’. After translating the first two Chinese words word by word, the EFL speaker has part of the English sentence, ‘this dish is very…’ in his or her mind. However, when it comes to the third word ‘香 (smells good),’ the EFL speaker may hesitate to find a mental solution to translate it since there are two word-translating options in the learner’s mind, ‘fragrant’ and ‘aromatic.’ Finally, the EFL speaker may rush to choose ‘fragrant’ without more time to consider.

Both a foreign language (FL) and a second language (L2) refer to a language different from the mother tongue/first language (L1). A second language is another language used in the learner’s community. For example, Indian people use both an Indian language and English in their lives. The Indian language is their first language and then English is the second language (L2). By contrast, a foreign language is the language that is not mainly used in the learner’s community. For example, Chinese people usually use Chinese and barely use English in their daily lives, so English is the foreign language (FL) for people in China. Compared to L2 learners, FL learners have fewer chances to engage in the conversation in the target language. To simplify the terms, L2 and FL will be alternated in this paper.
in spontaneous speaking. As a result, the FL learner says an awkward sentence in English: ‘this dish is very fragrant’, which sounds like the plate containing the food has a flower-like smell. When these kinds of difficulties arise frequently, the FL learner’s speaking will not be fluent and natural.

Thus scholars argue that it is important for FL learners to build the connection between the conversational scenarios and the FL directly without using the mother tongue as a medium in the mind. In the above example, if a FL learner has already been exposed to a real situation in which a dish was on a table and a native speaker says ‘this dish smells good,’ the FL learner may say the sentence in similar situations immediately without the confusing translating process from the mother tongue to the target language. Based on a study involving 24 in-depth interviews and a web survey of 167 FL learners, Resnik (2018) consolidated this point of view that exposing to and engaging in real conversational scenarios will help FL learners form their thoughts in the FL. The finding of the study has shown that living in an English-speaking country and frequently being exposed to and engaged in naturalistic language use contexts led to a significant increase in the proportion of thinking in English.

Scholars in language socialization provide a broader view to explain why real-life situations are important for language learning. From the perspective of language socialization, language learning includes not just linguistic knowledge, but also other forms, such as cultural and social knowledge that are learned in and through real-life social interaction (Duff & Talmy, 2011). For example, Afghari and Karimnia (2007) compared the cultural differences in everyday conversation between English and Persian. The results cover a wide range of topics including greetings, farewells, terms of address, compliments, praise, and social amenities such as apologies and gratitude. The findings show that learning a foreign language is not just mastering
the pronunciation, grammar, and vocabulary, but also learning the ways in which the foreign language reflects the ideas, customs, and behaviors of the society.

The importance of culture to improve communicative competence was further highlighted in a comparative experimental study conducted by Li (1999). Li (1999) identified the interaction between two Canadians in English or two Chinese students in Mandarin Chinese as intra-cultural communication. The interaction between a Canadian and a Chinese person was identified as inter-cultural interaction. Forty Canadian and Forty Chinese university students were randomly assigned to intra-cultural or inter-cultural groups to play the role of physician and patient and have a conversation accordingly. The study found that only 50% of information was delivered in inter-cultural interaction, which was significantly less than in intra-cultural interaction (75%). The finding implies that a FL learner without understanding of the target-language culture would have lower competence to deliver information in conversation.

Moreover, many empirical studies have proved that the use of places outside the classroom can support the development of communicative competence effectively. Fajaria (2013) designed a series of learning activities at places outside classrooms for 20 English FL high-school students. The activities covered the following topics: hotel services, facilities in a hotel, traveling, and cuisine. The results of surveys on the participants indicated that the outside-classroom activities helped students speak English fluently and increased their social interaction, courage, self-confidence, and their understanding of the culture and social roles they played. Richardson and Murray (2017) found that children had more physical play, more enhanced sensory experience, and were more excited while learning language in a forest compared to learning in classroom.
In brief, a successful FL learner needs to learn a FL in situ, i.e., learn the FL on the site or in the place where the FL is naturally spoken to improve the FL learner’s communicative competence. “Well-rounded communicative proficiency, it seems, depends to a large extent on the learner’s efforts to use and learn the language beyond the walls of the classroom” (Benson & Reinders, 2011, p. 2). So far, a large number of language programs for studying a foreign language abroad (e.g., Intensive language program, 2019) have been available. After enrollment in these programs, FL learners can travel to the target-language countries, live with host families, volunteer with local organizations, study with native students, and explore the foreign language through cultural events (Intensive language program, 2019). However, only a few learners worldwide can afford these programs due to the high cost of travel, room and board, and tuition fees. For example, the cost of a one-year English communication program hosted by a university in the USA is estimated at over forty thousand dollars (Intensive English Communication Program Admissions, 2019). According to statistics from the International Association of Language Centers (IALC) (IALC Study Travel Research Report, 2016), only 0.25% of FL learners were able to travel to the target countries for FL learning.

**Teaching Methods to Improve Communicative Competence**

In order to increase FL learners’ exposure to and engagement in the naturalistic context of the FL without traveling to a target country, a variety of teaching methods have been explored and developed. An earlier teaching method focusing on the communicative use of a L2 is called Audio-lingual method or Army method (Alemi & Tavakoli, 2016; Mart, 2013). The Audio-lingual method emphasized the importance of listening to and speaking a FL based on the principle that ‘a language is first of all a system of sounds from social communication; writing is

---

3 In situ is a Latin phrase, which means ‘on site’ or ‘in position’.
a secondary derivative system for the recording of spoken language’ (Carroll, 1963, p. 1063).

The Audio-lingual teaching method typically follows a three-stage model called PPP: presentation-practice-production (Talebi et al., 2015; Tan, 2016). In PPP, the teacher first reads a dialogue in the presentation stage. Then students memorize the dialogue by repeating it from memory at the practice stage. Finally, the FL learners are given the opportunity to produce new speech in the L2 using various drilling techniques such as single-slot substitution, in which the teacher states a line in the dialog and gives a clue word; then students produce a new sentence by replacing a word in the given line with the clue word (Mart, 2013).

Although the Audio-lingual method is easy to implement in classroom settings and can build students’ habits in speaking and listening, it has received waves of criticism mainly because of its roots in Skinner’s Behaviorism theory. This theory claims that humans learn through training and reinforcement (Alemi & Tavakoli, 2016). It was discovered that the acquisition of a language is more complex than habit formation in Behaviorism, especially considering humans have the ability to create novel and unique sentences, which cannot be explained as habit formation (Chomsky, 1959; Savignon, 2018).

Furthermore, Savignon (2018) found that FL learners failed to use the structures and vocabulary they had rehearsed in the Audio-lingual drilling methods when they were asked to engage in spontaneous interaction. In a comparative experimental study, Savignon (1972) found that FL learners who were engaged in unscripted communication far excelled learners who were taught through the audio-lingual drilling method when it came to unscripted communicative tasks. FL learners engaged in unscripted communication demonstrated the same grammatical accuracy as the learners who had spent time drilling grammatical patterns. Therefore, Audio-
lingual methods are actually weak in improving FL learners’ communicative competence in real situations, although its goal is to improve the communicative use of the FL.

Terrell (1977) proposed the Natural Approach (NA) to overcome the shortcomings of the Audio-lingual method. The NA distinguishes language learning from language acquisition. Learning is the conscious process of understanding the grammar rules and memorizing the vocabulary or dialogue by rote memory and drilling. Acquisition, on the other hand, is the unconscious process of absorbing general principles of grammar through communication in real conversational situations. According to Krashen’s monitoring hypothesis (Krashen & Seliger, 1975; Krashen & Terrell, 1983), if a grammar rule has only been learned rather than acquired, then the FL learner cannot use the grammar rule in spontaneous speech. Rather, the FL learner can only use the learned grammar rule to notice whether his or her speech follows the rule. Krashen’s monitoring hypothesis explains why FL learners could not use the grammar drilled in the Audio-lingual method when the learners engaged spontaneous interaction — FL learners only learned the grammar rules in their consciousness, yet they hadn’t ‘acquired’ the grammar; thus, they couldn’t use the grammar rule naturally in spontaneous conversation.

Accordingly, Terrell (1977) claimed that FL adult learners should be given the opportunity to acquire language in a natural way, i.e., how a child naturally acquires his or her mother tongue in daily communication without formal education or a special course. Terrell (1977) assumed that the process of L2 acquisition in adults is quite similar although not identical to L2 acquisition by children. Accordingly, the following principles were proposed by observing the language acquisition of children: (1) all classroom activities especially for new FL learners should focus on meaningful communication rather than language form, such as grammatical perfection, considering children do not learn grammar explicitly, rather, they induce grammar
from speech implicitly; (2) errors in speech shouldn’t be corrected during the conversation since grammar errors, such as subject-verb and non-adjective agreement, are acceptable in real communication with native speakers and error-correction will discourage FL learners; (3) affective factors, such as emotion, the motivation to use the FL in real life situations, and avoiding embarrassment during a conversation, play an important role in L2 acquisition. In addition, the learning procedure of the natural approach follows three stages: listening comprehension, simplified response with few words or even partly in L1, and advanced response with more words and more complex sentence structures.

Although Terrell (1977)’s natural approach has emphasized the importance of communicative competence and gives general guidance, the specific methods and techniques of the natural approach haven’t been well developed (Pervaiz, 2013). Later, communicative language teaching (CLT), a more mature approach to build FL learners’ communicative competence, has been dominant in FL teaching-method research and widely promoted in many countries in recent years (e.g., Benati, 2009; Tan, 2016; Wang, 2007). For example, the Educational Ministry of China has designed New Curriculum Standards (NCS), in which one of the purposes is to improve students’ communicative competence by using CLT (Wang, 2007).

Similar to the Natural Approach, CLT emphasizes providing FL learners with authentic language input in real life context and encouraging FL learners to produce genuine and meaningful communication as a pair or in a group (Brown, 2000). Both the Natural Approach and CLT hold the basic theory that acquisition of a FL can only take place while FL learners are involved or involve themselves in meaningful and spontaneous communication. Accordingly, error correction should be avoided during conversations, since error corrections discourage FL learners, and the communicative activities focus more on meaningful interaction rather than the
language form, such as pronunciation, spelling, and grammar (Krashen & Terrell, 1983; Nunan, 1989; Richards, 2006; Terrell, 1997).

Different from the Natural Approach, CLT has developed more systematic and specific instructional methods and techniques. Richards (2006) grouped the instructional methods of CLT as content-based instruction, task-based instruction, text-based (or genre-based) instruction, and competency-based instruction as the foci of the FL instructions. Content-based instruction organizes the communicative activities based on topics (e.g., environmental protection, travelling); task-based instruction involves students in various real-life tasks or instructional tasks such as role-playing or information exchange either in or outside classrooms; genre-based instruction emphasizes how texts or speeches in different genres, such as notice, poem, receipt, news, informal or formal conversation, are used in various language use situations; finally, competency-based instruction identifies and develops FL learners’ competencies to use the language to complete tasks in real life contexts (e.g., ordering food, filling out a tax form).

Although CLT has been widely researched and applied, a main challenge in the current CLT practice is the lack of resources to build authentic language contexts. As Tan (2016) pointed out, most EFL learners barely have the chance to use English outside the classrooms since English is a foreign language in China. Teachers in Asia, such as in China and Indonesia, lack confidence to conduct communicative activities in English due to their insufficient proficiency and experience. Many EFL teachers, especially in middle schools, do not have a solid educational background for teaching English, and the majority of the EFL teachers have never been to the countries where the target language is used (Marcellino, 2005; Rao, 2013). Also, the big class sizes (usually from 40 to 60) in many Asian countries challenge the FL instruction in pair work and group work emphasized in CLT (Hiep, 2007; Tan, 2016). Therefore, it has been a
growing trend to use information communication technologies to provide FL learners with more resources such as connecting FL learners with native speakers in target countries.

**Technologies Applied in Language Learning and the Pushes for Authenticity**

So far, various information technologies have been applied in foreign language learning, such as video clips (e.g., Winke et al., 2010; Chai & Erlam, 2008), virtual reality technologies (e.g., Lan, 2015; Esmaeili, 2016), and augmented reality based on mobile devices (e.g., Holden & Sykes, 2011). Each technology has its strengths and weaknesses for language learning.

For instance, digital video clips organized by communicative functions enable FL learners to view raw linguistic data in rich language contexts in real life (Kitajima & Lyman-Hager, 1998). Various video processing software can be used to manipulate discrete portions of digital videos to show the video clips according to the needs of students (Kitajima & Lyman-Hager, 1998). FL learners can easily access video clips on websites such as YouTube (Burston, 2005). In addition, video subtitles can enhance FL learning by bridging the gap between reading and listening and helping FL learners acquire new vocabulary and idioms (Zanón, 2006).

However, video clips just allow foreign language learners to receive or input information such as watching videos, listening to audio and reading subtitles. According to the comprehensive output hypothesis (Swain & Lapkin, 1995), FL learners need to output the language through speaking to check for any problems in speech and to notice the gaps between what they want to say and what they are able to say (VanPatten & Williams, 2014). Therefore, video clips have limitations for improving FL learners’ communicative competence in speaking.

Studies have found that virtual language learning environments can motivate FL learners to engage in conversations with the virtual avatars of other FL learners (Chang et al., 2012) and to connect the topics learned in classrooms with those in the virtual environment (Gadelha,
2018). Also, Chen (2016) found that learning in virtual reality not only improves learners’ linguistic knowledge such as pronunciation and vocabulary, but also develops learners’ advanced cognitive abilities such as analysis, synthesis, and evaluation.

Yet due to the limitation of current 3D modeling techniques and expensive costs of constructing 3D models, Falconer (2013) argued that most virtual environments lack authenticity, i.e., the virtual environment is not as real as the real world. For example, the characters in most virtual environments are depicted in an unrealistic cartoonish style, the physical environment lacks details, the learners cannot control the gestures of the virtual avatars precisely, and tasks are designed artificially (e.g., killing a dragon with others in a virtual world). Some scholars (e.g., Clarke & Silberstein, 1983; Richards, 2006) claim that FL learning activities should be as authentic as possible, i.e., mirroring the real world to a large extent, so current virtual reality technologies have a limitation of providing authentic FL learning environments.

Another trend of technology-mediated language learning is using augmented reality (AR) technology on mobile devices to support language learning (e.g., Godwin-Jones, 2016; Solak & Cakir, 2015; Holden & Sykes, 2011). Augmented reality refers to technologies that extend the real world with 2D or 3D virtual objects and allow the users to interact with them (Dita, 2016). For example, Holden and Sykes (2011) used ARIS, an augmented reality interactive storytelling tool, to create a foreign language learning AR game, Mentira. The FL learners investigated a murder case at a neighborhood in Albuquerque by having conversations with fictional characters on their mobile devices. As the FL learners physically navigate the neighborhood in Albuquerque, different conversations with various characters pop up on the mobile device screen. The analysis on interview transcripts and students’ records of using the tool showed that
the AR FL learning environment motivated FL learners to use the target language in a specific
place where the conversation may happen (Godwin-Jones, 2016; Holden & Sykes, 2011).

Although AR allows the FL learners to acquire a target language in an authentic way, i.e.,
the AR technologies will show or engage FL learners in conversation in a real-life context, it
may be problematic for FL learners to access the places where the target language is mainly
used. For example, an English learner in China may not find an American football stadium in his
or her town for learning topics relevant to American football since American football in China is
not as popular as it is in America. Even if the English learner manages to find an American
football stadium in China, there are huge differences between the stadiums in terms of the layout,
the number of the seats, and the behaviors of the audiences due to the cultural differences.
Consequently, the language used in the American football stadium in China will be different
from what is used in the US. The language authenticity is then closely bonded with the place
physically and culturally. Just as Kramsch and Sullivan (1996) point out, what is authentic in
London might not be authentic in Hanoi. Since AR technologies add 2D or 3D virtual objects or
avatars only to a local place around the FL learners, AR has limitations in engaging FL learners
into the authentic language in a remote target language country.

**Telepresence Technologies and Telepresence Robots**

In this study, I propose to use a new technology, telepresence robots, to support FL
learners in accessing a remote place in a target country and learning the authentic language use
there. The term telepresence, first proposed by Minsky (1980), refers to a set of technologies that
gives remote users the feeling of being present at a distant location. Bell, et al. (2016)
categorized telepresence technologies as comprising three generations. The first generation of
telepresence technologies consisted of video-conference systems on computers, e.g., Skype,
Adobe Connect, Zoom, and Google Hangouts, which supported audio and visual communications between people in different places by showing and transmitting compressed digital audio and video streams in real time. (See Figure 1.1)

Figure 1.2D Telepresence approach using a fixed camera and monitor

The second generation of telepresence technologies encompasses video-conference systems based on mobile devices, e.g., iPad, iPhone, and Android phones, with movable screens and cameras that make it possible to adjust the viewpoint of transmitted videos. Thus, instructors or students in the physical classroom are able to move the mobile devices manually. Figure 1.2 shows a classroom using a “TriPad,” which is an iPad and a tripod, as both a screen and a camera in order to display and capture audio and video streams. Other devices and systems of the second-generation telepresence technologies include selfie holders and live-broadcasting mobile apps such as Ustream Broadcaster, Air Cam Live Video, JumiCam Webcam Streamer, etc. Compared to the first telepresence technologies, second generation technologies offer an online student an adjustable viewpoint in the classroom. It is also possible to bring these mobile devices
outside the classroom, due to the fact that the weight of current mobile devices is light enough that they may be held for an extended period of time.

**Figure 1.2. 2.5D Telepresence approach using the “TriPad”**


However, the second-generation technologies still have issues due to the lack of control over camera position and view angle by remote users. According to Bell, et al. (2016) it is challenging for the local instructors and/or students holding the tripod mentioned above to perceive what the remote students are experiencing, so the local instructors and/or students might not be able to offer a better experience for the remote students. The necessity of adjusting the position and view angle of mobile devices, while simultaneously participating in the learning experience, is also a burden for the local instructors and/or students. Furthermore, the lack of control reduces the feeling of remote students’ social presence, which is the degree a remote person is perceived as a “real person” in the virtual communication (Gunawardena, 1995). Evidence has shown that social presence is positively related to students’ satisfaction and is a predictor of their perceived learning (e.g., Richardson & Swan, 2003; Szeto & Cheng, 2016).
Compared to the first two generations of telepresence technologies, telepresence robots, (see Figure 1.3) the third generation of telepresence technologies, allow remote users to control the movement of the robots and the view angle of the camera on the robots, which provide remote language learners with the ability to explore environments around the robots actively. There is no necessity for the local instructors and/or learners to adjust the position and view angle of the robots since remote students can easily reposition the robots themselves according to their needs. Bell et al. (2016) reported that students expressed satisfaction in being able to move the robot on their own, rather than relying on some else in the remote classroom. The autonomous action of controlling the robots provided students with a better experience of participating in class. Additionally, with the development of mobile technology, some telepresence robots such as Double Robotics and Keebots have mobile devices like the iPad or iPhone as part of their core components, which makes the cost of those robots significantly more affordable and thus can promote widespread use. In summary, telepresence robots have the most appealing potential to support FL learners to learn authentic language use at a remote place.
Research Purpose and Structure of the Dissertation

Although there are already some telepresence-robot studies in foreign language learning (e.g., Tanaka et al., 2014), the use of telepresence robots has been confined to a small indoor space, such as a single classroom or office. It is not clear what benefits and challenges the instructor and the learners will perceive while using telepresence robots in real-life settings outside the classroom. Moreover, few studies have discussed what learning theories and design concerns can be applied in this form of learning.

Therefore, the research purposes of this study are fourfold: (1) to investigate the perceived benefits and challenges of telepresence robots to support foreign language learning in a remote place outside the classroom, (2) to identify possible learning theories and constructs that can be applied to this form of learning, (3) to understand the learning process of the FL learning activities using telepresence robots outside the classroom, and (4) to examine the design concerns of the learning activities and provide sights for future design of the telepresence robots.

This dissertation is structured in a three-article-dissertation (TAD) format to meet the research purposes of the study. The first chapter is the introductions to this study. Chapters 2, 3, and 4 are three articles wrapped in this dissertation, two of which have been published in peer-reviewed journals. The last of the three articles will be revised and submitted to a peer-reviewed journal after submitting the dissertation. The final chapter will be the summary of the whole study.

As a pilot study, the first article examines the perceived benefits and challenges of using Romo, a small telepresence robot, around one of the main buildings on an American university campus. Three FL learners in China were recruited to have a one-on-one session with an English...
instructor at the American university’s campus. The instructor played the role of a tour guide to
discuss the objects and activities around the building with the remote FL learners, who controlled
the robots to follow the instructor.

The second article framed the FL learning via telepresence robots in real life settings as
telepresence-place-based foreign language learning (TPFLL). Three design principles were
proposed to guide the design of the learning activities. A case study including 11 FL learners in
China and four instructors at the American university was conducted to examine the
appropriateness of these principles at a campus arboretum at the university.

The last article further analyzes the learning process through a lens of geosemiotics,
which systematically analyzes how the place and other relevant factors, such as people with
different social roles and their social actions, shape the discourse at a certain location. The results
of the study are expected to explain how the features of a place influence the FL learning and
what instructional strategies can be used at different places. In addition, the demands of the
telepresence robot at different places are discussed for guiding the future design of the robots.

**Conclusion**

Learning outside the classroom, especially in the target-language countries or
communities, is important for FL learners. However, few FL learners can afford the cost of
travelling to the target-language countries due to expensive boarding and travel costs. This study
proposes an approach to the use of telepresence robots to connect and immerse a FL learner in a
place in the target-language countries. Telepresence robots will allow the FL learner to
communicate with native English speakers in natural contexts. This study also proposes a new
term, telepresence-place-based FL learning, to refer to this new approach. The perceived benefits
and challenges of telepresence-place-based FL learning are investigated. The possible design
concerns of learning activities and telepresence robots will be discussed in detail to address the challenges experienced by FL learners. Finally, the findings of this study are expected to inform theories in this area and to understanding the learning process of TPFLL.

References


https://www.slideshare.net/mubashar92/natural-approach-merits-and-demerits


In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction* (pp. 399-406). ACM.


Vanegas, C. (2008), Similarities and Differences between First and Second Language Acquisition, *Multilingualism Wiki*, Retrieved from


CHAPTER 2: EXPLORING THE AFFORDANCES OF
TELEPRESENCE ROBOTS IN FOREIGN LANGUAGE LEARNING

This chapter has been published as an article in the Journal of Language Learning and Technology. The content of the article is reprinted with the permission. The layout of the article has been updated to be aligned with the requirements of dissertation format.

You can cite the article as follows:

Abstract

The importance of authentic communicative practices in foreign language (FL) learning has long been recognized. However, most FL learners lack adequate access to authentic communicative environments in the target language. In this article, we propose the use of telepresence robots as a potential solution to bridge this gap. Telepresence robots can be controlled by remote language learners online, enabling them to gain virtual access to authentic environments in the target language and to interact with native speakers in those environments in real time. In this exploratory study, three English learners and a native-speaker of American English participated in a campus tour activity using a telepresence robot. We examined the experience of our participants and the conversational features of their telepresence interactions through analyses of the interview data, field notes, and transcripts of conversations captured on video. Our findings show that telepresence robots have substantial potential for promoting FL learning by providing authentic communicative practice for remote language learners. The findings have useful implications for informing future research design.

Keywords: Computer-Mediated Communication, Language Learning Strategies, Distance and Open Learning and Teaching, Telecollaboration
Introduction

The importance of authentic communicative practices in foreign language (FL) learning has long been recognized. Authentic communicative practices provide meaningful contexts that motivate learners and assist them in understanding the meaning of knowledge (Edelson & Reiser, 2006). In the absence of such practices, learners demonstrating adequate language skills in standardized testing situations may not be able to successfully apply those skills in authentic communicative situations (Larsen-Freeman, 2013). Unfortunately, most FL learners lack adequate access to authentic communicative environments in the target language. Researchers have used various virtual presence technologies to facilitate interaction between FL learners and speakers in target language communities, such as video conferencing systems (e.g., Yen, Hou, & Chang, 2015) and virtual reality (VR) environments (e.g., Lan, 2015). However, these technologies fall short in enabling learners to feel physically present in and to explore diverse real-world environments of target language use. To bridge this gap, we propose a solution that uses a telepresence robot—a remote-controlled, wheeled device with a display and camera—to enable FL learners to gain virtual access to authentic communicative environments in the target language and engage in real-time interaction with speakers in those environments.

The term telepresence, first coined by Minsky (1980), refers to a set of technologies that give remote participants the feeling of being present at a different location. Compared to other robots, telepresence robots provide opportunities for interpersonal communication with people at different places, rather than interaction between humans and robots with artificial intelligence. Telepresence robots have been used in various domains, such as providing home care assistance for elderly people (Michaud et al., 2007), changing the dynamics of a blended learning design studio class (Bell, Cain, Peterson, & Cheng, 2016) and facilitating virtual school attendance by students with physical disabilities (Newhart & Olson, 2017). In the field of FL learning, a few
studies reported that the application of telepresence robots in classroom settings significantly improved learner interest, confidence, and motivation (e.g., Kwon, Koo, Kim, & Kwon, 2010; Tanaka, Takahashi, Matsuzoe, Tazawa, & Morita, 2014). Research into the application of telepresence robots in FL learning in real-life settings, however, has not yet emerged. As a first step toward understanding the potential of using telepresence robots for providing authentic communicative practices for FL learners, this exploratory study examines the experiences of four participants in a campus tour activity facilitated by a telepresence robot and the features of their telepresence interactions.

**Theoretical Background**

*Communicative Practices*

From the dialogic view of language and learning (Bakhtin, 1981), language is organized dialogically at the level of utterance, which is both context shaped and context renewing (Goodwin & Heritage, 1990). As Volosinov (1973) put it, “the real unit in language that is implemented in speech … is not the individual, isolated monologic utterance, but the interaction of at least two utterances—in a word, dialogue” (p. 117). He further argued that “language acquires life and historically evolves precisely here, in concrete verbal communication, and not in the abstract linguistic system of language forms, nor in the individual psyche of speakers” (p. 95). Based on this view, communicative practices play a critical role in language learning, as it is through engagement in such practices that learners acquire the meanings and functions of language forms in context. In reality, however, most FL learners have few opportunities to engage in communicative practices with speakers in the target language community. According to the International Association of Language Centers, only 0.25% of FL learners are able to travel to target-language countries for educational purposes (International Association of
It is thus important to find alternative ways to provide communicative practices for FL learners.

**Authenticity in Language Learning**

Authenticity in language learning refers to the resemblance between what learners are exposed to in learning and what their future language use practices will be like (Gilmore, 2007). Taylor (1994) distinguished three facets of authenticity (i.e., authenticity of language, task, and situation), which were subsequently elaborated on by other researchers. In general, authentic language, tasks, and situations refer to language input that serves real-life communicative purposes (Lee, 1995), tasks that bear real-world relevance and reflect professional skills students will need after graduation (Strobel, Wang, Weber, & Dyehouse, 2013), and situations that resemble real situations of social interaction in daily life (Strobel et al., 2013), respectively. These facets should not be viewed in isolation, as authenticity is a function of many factors all at once, including not only the language but also “the participants, the use to which language is put, the setting, the nature of the interaction, and the interpretation the participants bring to both the setting and the activity” (Taylor, 1994, p. 4).

**Virtual Presence Technologies and Computer-Assisted Language Learning**

Technologies that can create virtual presence or experiences, including video conferencing systems, VR environments, and telepresence robots, have been utilized to support FL learning. For example, Lan (2015) found that the usage of virtually immersive contexts in English as a foreign language (EFL) learning provided ubiquitous learning opportunities and game-like scenarios and led to enhanced learner performances. Jauregi and Bañados (2008) employed Adobe Connect to enable virtual interaction between Spanish as a foreign language
learners and native-speaker Spanish teachers and identified positive impacts on learning outcome, learner motivation, and cultural understanding. Yen et al. (2015) reported that using Skype to facilitate the application of role-playing strategy significantly enhanced EFL learners’ speaking skills. A few studies applied telepresence robots in FL classrooms and observed significant improvement in learner interest, confidence, and motivation (Kwon et al., 2010; Tanaka et al., 2014).

Despite the reported benefits, current applications of virtual presence technologies have their drawbacks when it comes to providing learners with authentic communicative practices. Falconer (2013) reported that characters in VR learning environments are typically depicted in an unrealistic cartoonish style, that the physical environments lack details, and that tasks are designed artificially. Research on language learning via video conferencing systems and telepresence robots has to date been restricted to classroom, office, or home settings (e.g., Kwon et al., 2010; Tanaka et al., 2014; Yen et al., 2015). This limits the types of real-world environments learners are exposed to and the conversational topics that could have naturally arisen from more diverse contexts. Collectively, these drawbacks result in the lack of authenticity. In our view, telepresence robots have the potential to bring about a greater degree of authenticity in all three facets discussed above if used to allow learners to interact with speakers in the target language community in more diverse real-world environments for genuine communicative purposes.

**Research Questions**

This exploratory study constitutes a first step towards understanding the potential of using telepresence robots to provide authentic communicative practices for remote FL learners. The specific research questions addressed are as follows:
1. What are the perceived benefits of using telepresence robots in FL learning outdoors?
2. What are the perceived challenges of using telepresence robots in FL learning outdoors?
3. How can we refine the design of tasks to address the perceived challenges?

Methodology

Participants

The data in this study were collected over one week in April 2015. The purposeful sampling strategy was adopted to select adult EFL learners with the plan to study overseas and with no prior experience in living abroad, as the campus tour activity (described below) would be directly relevant to them. Three Chinese EFL learners residing in China (Xiu, Xian, and Gou) were recruited as remote participants (see Table 2.1); Coleman, a male native speaker of American English enrolled in a graduate program at the university chosen as the research site, was recruited as the local participant (all names are pseudonyms).

Table 2.1. Demographic Information of Remote Participants

<table>
<thead>
<tr>
<th>Pseudonym</th>
<th>Gender</th>
<th>Age</th>
<th>Occupation</th>
<th>English Proficiency</th>
<th>Years Studying English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gou</td>
<td>Male</td>
<td>24</td>
<td>Graduate student</td>
<td>Intermediate</td>
<td>12</td>
</tr>
<tr>
<td>Xian</td>
<td>Female</td>
<td>28</td>
<td>Government employee</td>
<td>Intermediate</td>
<td>9</td>
</tr>
<tr>
<td>Xiu</td>
<td>Female</td>
<td>24</td>
<td>Graduate student</td>
<td>Advanced</td>
<td>12</td>
</tr>
</tbody>
</table>

Technical Settings

Romo, a small telepresence robot, was used in this study. Despite its relatively small size, Romo has most functionalities that other telepresence robots have. Romo allows learners to
control its movement via a mobile app, to adjust the view angle of its camera by tilting the smartphone, to see what it captures on the screen of the smartphone, and to start a live video chat with the camera. We used the 4G network on an iPhone to connect Romo to the Internet.

**Procedure**

The research site was the campus of a large public university in the Eastern United States. The general goal of the activity was to have the native speaker participant introduce the buildings, history, and culture of the campus to the remote EFL learners as a campus tour guide; the EFL learners would then report what they learned at the end of the activity. The actual activity consisted of three phases: pre-task, task, and review. Each learner completed the task with the native speaker one-on-one in approximately two hours. In the pre-task, the native speaker and the learner spent half an hour video-chatting on QQ, a popular chat app, to get familiar with each other; the native speaker also oriented the learner to the whole task. In the task phase, the learner controlled Romo to participate in a 1-hour campus tour guided by the native speaker, who introduced the history and culture of the buildings along the tour route.

*Figure 2.1. Communication via a telepresence robot*
Figure 2.1 and Figure 2.2 demonstrate the robot interface and the mode of communication between the native speaker and the learner. Figure 2.3 shows the campus tour route. In the last phase, the learners orally reported what they had learned at the end of the tour. They were then interviewed about their experiences with the task. The native speaker participant was interviewed after all three learners completed the task.

*Figure 2.2. The Interface on the screen of the robot*

*Figure 2.3. The route of the activity*
Data Collection and Analysis

Throughout the data collection period, four interviews (see Appendix I) were conducted, recorded, and transcribed by the first author: three with the EFL learners in Chinese and one with the native speaker in English. The interviews with the learners were translated into English and the translations were crosschecked by the two authors to ensure accuracy. 10 pages of field notes were taken by the first author to document researcher observation of the benefits and challenges manifested in the activity and to triangulate the interview data. The interview transcripts and field notes were analyzed qualitatively by the two authors using Atlas.ti, following the method McMillan (2012) describes. Specifically, the transcripts and notes were read through and words, phrases, or sentences relevant to the research questions were marked up as codes. The coding scheme went through several iterations of merging and splitting, until we were confident that each code could not be further split or merged with other codes. Finally, all codes were categorized into different themes.

In addition, six hours of video interactions were recorded, among which three hours were interactions between the EFL learners and the native speaker via the telepresence robot during the campus tour task. The telepresence conversations were transcribed and analyzed by the first author using the Computerized Language Analysis programs (MacWhinney, 2000). The results of this analysis were used to triangulate those from the coded data.

Findings

Five themes emerged from the 21 codes identified in the field notes and interview transcripts, as shown in Table 2.2.
### Table 2.2. Themes and codes

<table>
<thead>
<tr>
<th>Theme</th>
<th>Code</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Emotions</strong></td>
<td>Exciting</td>
<td>Participants feeling excited</td>
<td>The real cool thing is obviously the interaction and the use of technology to actually experience real things.</td>
</tr>
<tr>
<td><strong>Gaming</strong></td>
<td></td>
<td>Participants feeling game-like experiences</td>
<td>This robot makes me feel like manipulating a virtual person walking in a virtual world. This is somewhat similar to manipulating a character in a game hanging around in a virtual world.</td>
</tr>
<tr>
<td><strong>Relaxing</strong></td>
<td></td>
<td>Participants feeling relaxed</td>
<td>I feel the way we used it today is more relaxing...</td>
</tr>
<tr>
<td><strong>Challenging</strong></td>
<td></td>
<td>Participants feeling challenged</td>
<td>I feel it is more challenging considering my English ability.</td>
</tr>
<tr>
<td><strong>Authentic Learning Experience</strong></td>
<td>Real environment</td>
<td>Experiencing another real world via the robot</td>
<td>Meanwhile, I know the world where the robot is walking is a real world.</td>
</tr>
<tr>
<td><strong>Social presence</strong></td>
<td></td>
<td>Feeling present in a remote place</td>
<td>I feel as if I was hiding in a corner watching the foreigners and experiencing many things.</td>
</tr>
<tr>
<td><strong>Emerging topics</strong></td>
<td></td>
<td>Naturally emerged topics</td>
<td>I feel that there are more topics to talk about when immersed in the environment</td>
</tr>
<tr>
<td><strong>Cultural learning</strong></td>
<td></td>
<td>Learning about another culture</td>
<td>(Learning in this way) provides an opportunity to learn something about foreign culture.</td>
</tr>
<tr>
<td><strong>Learner-Centered Activities</strong></td>
<td>Learner agency</td>
<td>Learning in an active way</td>
<td>When watching the introduction videos or photographs of campuses online, I can only receive information passively. With the robot, I can explore the campus actively...</td>
</tr>
<tr>
<td><strong>Self-impact</strong></td>
<td></td>
<td>The activity is related to students’ personal experience</td>
<td>Especially for college students, it will be very attractive to know something about the lives of students of their age abroad or something related to their majors.</td>
</tr>
<tr>
<td><strong>Learner difference</strong></td>
<td></td>
<td>Inter-learner differences</td>
<td>The first student seemed more interested in seeing the campus ..., and the other two students seemed really enjoying driving around by themselves</td>
</tr>
<tr>
<td><strong>Disorientation</strong></td>
<td></td>
<td>Difficulty in locating position</td>
<td>When I was operating the robot, it was hard for me to find the destination since I’m not familiar with the campus.</td>
</tr>
</tbody>
</table>
Technical Issues

<table>
<thead>
<tr>
<th>Network connection</th>
<th>Issues caused by poor network connection</th>
<th>Also I mean the lag is there, right? It’s a bit of a problem…</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robot size</td>
<td>Small size of the robot</td>
<td>Colman had to crouch down to talk to Xiu.</td>
</tr>
<tr>
<td>Robot speed</td>
<td>Slow speed of the robot</td>
<td>As the robot walks slower than human beings, Colman had to hold the robot while we were heading there.</td>
</tr>
<tr>
<td>Robot sound volume</td>
<td>Low sound volume of the robot</td>
<td>There were times in the tour when I had to talk pretty loudly to make sure the student can hear me.</td>
</tr>
<tr>
<td>Reflections on the screen</td>
<td>Display issue caused by strong sunlight</td>
<td>When the robot is being ridden in strong sunlight, it is a little hard to see Xiu’s face on the screen.</td>
</tr>
<tr>
<td>Body language</td>
<td>Difficulty to convey body language</td>
<td>I could use gesture and body language when I talk to other foreigners face-to-face. I feel it lacks a way to communicate.</td>
</tr>
</tbody>
</table>

Practical Concerns

<table>
<thead>
<tr>
<th>Privacy issues</th>
<th>Concern for privacy issues</th>
<th>If you don’t videotape, you can conduct this study indoors or in a lot of places outdoors.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher resource</td>
<td>Concern for the lack of native-speaker teachers</td>
<td>The first issue is if there are enough native-speaker teachers.</td>
</tr>
<tr>
<td>Difference from classroom learning</td>
<td>Comparison to learning in the classroom</td>
<td>When listening to a lesson in the classroom, I focus on something else, like grammar or vocabulary. When using the robot, I need to focus on how to express my ideas more accurately.</td>
</tr>
</tbody>
</table>

Emotions and Motivation

The EFL learners and the native speaker experienced positive emotions, as expressed using such words as excited, cool, and interesting during the activities. Coleman noted the following in his interview:

*It was really cool to see the students’ reactions… Students’ reactions were really impressive. I think all three of them. There were moments that they just seemed really excited to interact with me or just interact with the environment. And it really shocked me that their interaction with the environment made it really interesting.*
Gou explained the reason he experienced positive emotion from the perspective of gaming in his interview.

*This robot makes me feel like manipulating a virtual person walking in a virtual world. This is somewhat similar to manipulating a character in a game hanging around in a virtual world. Meanwhile, I know the world where the robot is walking is a real world. It feels like a game experience but what I am experiencing is a real world. This makes me excited."

Nevertheless, some learners experienced some stress during the activity. For example, Xian mentioned that she felt her English ability was not good enough to communicate with the native speaker fluently. However, both Xian and the native speaker indicated that this challenge could be alleviated by better preparing learners before the activity (e.g., by having them watch some introductory videos about the campus).

**Authentic Learning Experience**

Authentic learning experiences were created by making learners feel they engaged in the activity as if they were physically on campus. Xian indicated in her interview that she felt as if she were physically present on campus as a student talking to Coleman.

*When he (Coleman) took me around, it felt like a teacher guiding a student, and I paid attention to everything he said.*

Gou mentioned a similar experience in his interview.

*As for my other experiences, the most interesting thing is to see the real environment abroad, which is very different from what we have experienced domestically.*
In terms of the benefits of learning in an authentic environment, Xiu’s comment in her interview below illustrates how the environment provided a rich and unique language context to facilitate her acquisition of the precise meaning of a culture-related word.

Xiu: In our English class at school, we also have a native speaker teacher here, but in China, it’s more difficult to come up with examples to illustrate the meaning of something specific to her culture. For example, when Coleman told me something about the word quod? I forgot the exact word...

Interviewer: Quad?

Xiu: When he wanted to introduce something like that, he would say, “see it’s over there”. It was more situated and more precise. Because of the large cultural gap, it’s just more difficult to explain things like this in China.

The following exchange between Xiu and Coleman (transcribed from the video interactions) can triangulate Xiu’ points mentioned above (see Figure 2.4). Jefferson’s (2004) transcription system is used here to mark pauses, voice pitch, and non-verbal gestures.

1 COL*: so:(.){what can you see there = can you see:
2 }((RH points at the quad and then gazes at XIU))
3 XIU*: a v.. a a: (1.1) a plaza↑ or a square?
4 >plaza↑, square?<
5 COL*: yeah, yeah, that’s a good way >to describe it<,
6 a pla:za or square like what i said
7 we we usually call this a qua:d right? kweu yew ey dee=
8 XIU*: =oh↑, [quad
9 COL*: [yeah, i told you >about this before<
10 so, this is uhm this is really normal for: university
11 campuses right?
Xiu’s example of the word *quad* explicates her view on the usefulness of the physical context for the acquisition of the precise meaning of this word, which could be harder to learn in the traditional classroom setting. Therefore, the physical context around the telepresence robot is not just the background of the conversation, but provides a way to help learners understand the precise meaning of culture-specific idiomatic expressions in the target language.

*Figure 2.4. Learning the meaning of the word quad*

Another finding is that, along with the learning activities, the authentic environment around the telepresence robot triggered more natural conversational topics and helped the interlocutors organize the topics, as Coleman observed in his interview:

*I feel like the tour itself was pretty straightforward. When we were talking about doing this, I was worried that the students would need like metrics or a worksheet or something to engage them in the tour, but actually it turned out it’s not necessary at all. Once we were moving around, it was very easy to have a conversation.*

Coleman’s comment indicates that the authentic physical environment, including the buildings and other objects encountered on the campus tour, allowed for conversational topics to emerge naturally as they moved along the route, making the activity more natural and closer to
real-life communication than conversations in formal learning settings, where the conversational topics are often prescribed.

Learner-Centered Activities

The learning activities were largely learner-centered, as the telepresence robot allowed learners to decide how they wanted the activities to proceed and what they wanted to talk and learn about. Gou’s comment in his interview below illustrates how he could actively choose the learning content.

The most impressive part of the experience was that I could control the robot myself. I could find a lot of videos and photos about university campuses online, but this project provides a robot that I can control, which gave me the freedom to see what I wanted to see. It is highly autonomous. When watching the introduction videos or photographs of campuses online, I can only receive information passively. With the robot, I can explore the campus actively, and that’s the most fun part of today’s activity.

Gou’s comment shows that the ability to control the movement and view angle of the robot allowed him to actively explore the environment and choose conversational topics that he was interested in, making the learning process highly engaging. Gou’s interview transcript indicates how the match between the learning scenario and his personal interests was a motivating factor for him.

Seeing how foreign college students of our age live, where they live, how they study, and how they eat, I feel as if I was hiding in a corner watching the foreigners and experiencing many things. ... Especially for college students, it will be very attractive to know something about the lives of students of their age abroad or something related to their majors.
Coleman also commented on the match between the activity and the learners’ interests and the importance of taking learners’ background and interests into account in activity design.

*It was a really good match between their background and interests and what we showed them today.*

*But I think that would really depend on who you’re taking around.*

A drawback of the activity was that learners might feel disoriented when exposed to authentic learning environments without adequate guidance, as illustrated by Xiu’s comment below.

*When I was operating the robot, it was hard for me to find the destination since I’m not familiar with the campus.*

**Technical Issues**

The learning activity generally proceeded smoothly, but there were a few minor issues caused by technical limitations. The most salient one was the occasional lags in streaming video and audio caused by unstable network connection, as observed by Gou in his interview:

*As for future improvement, the only thing I can think of is the Internet speed. I don’t know if it was because of the Internet connection on my side. The whole activity was well-organized, but due to the lag, I had some problems hearing the teacher, and occasionally the delay was severe. Other than that, the tour was really interesting.*

Other issues noted by the participants included the size, speed, sound volume, and the screen of the robot, largely due to Romo’s relatively small size. Most participants hoped the robots used for future studies could be taller and faster, with higher volume and a clearer display screen.
A special issue caused by the size of the display screen was about the conveyance of body language. In particular, the small size of the screen made it difficult to see the gestures and body language of the interlocutor clearly, as illustrated by Xian’s comment in his interview below:

_The communication in our activity lacks one thing. I could use gesture and body language when I talk to other foreigners face-to-face. I feel it lacks a way to communicate [while using the telepresence robot]._

**Practical Concerns**

The participants also reported a few practical concerns regarding the use of telepresence robots for FL learning, including the privacy of other people who may be captured by the camera, limited availability of native speaker interlocutors, and the integration of this technology in traditional classroom settings. These concerns are discussed in the next section.

**Discussion**

The findings from our analyses of the field notes, interview transcripts, and telepresence conversations have well answered our research questions.

*Perceived Benefits of Using Telepresence Robots in Foreign Language Learning Outdoors*

Our results show that the telepresence robot provided a more authentic and interactive environment for communicative practices than other researched technologies such as VR environments (Falconer, 2013). Such an authentic conversational environment provided multiple benefits to the learners. First, exposing learners to and engaging them in such an authentic conversational environment facilitated their acquisition of the precise meanings of culture-related expressions in the language. Second, this environment allowed for conversational topics to
emerge naturally, making the flow of the learning activity easy to organize for the participants. Third, the use of the telepresence robot motivated the learners by providing them with a game-like experience and learner-centered activities. As such, the learners all experienced positive emotions during the activities. Finally, the ability to control the movement of the robot and select learning content based on their own interests and learning needs allowed the learners to align the learning activity with their own zone of proximal development (ZPD; Lantolf & Thorne, 2006; Vygotsky, 1978), that is, the area between what they are able to do with and without expert guidance.

Perceived Challenges of Using Telepresence Robots in Foreign Language Learning Outdoors

The learners reported two major challenges posed by the outdoors activity with the telepresence robot. One challenge was the disorientation experienced by some learners in the new, open environment, and the other was the stress felt by some learners when communicating about certain impromptu topics that they felt they lacked the language skills to handle. On the one hand, these challenges can be seen as useful learning opportunities brought about by the activity. Authentic contexts of language use will necessarily include new environments, in which impromptu conversational topics will naturally emerge. Authentic communicative practices should thus ideally provide opportunities for exposing learners to such environments and impromptu topics. On the other hand, it is also important to help learners learn within their ZPD (Lantolf & Thorne, 2006). Giving learners the ability to control the movement and view angle of the telepresence robot and to self-select the learning content helped address some of the disorientation and stress. More importantly, the mediation (Lantolf & Thorne, 2006) or assistance the learners received from the native speaker supported both their learning about the
new environment and their learning of new culture-specific expressions that were useful for
talking in and about the environment.

In terms of technical limitations, issues such as coverage and speed of wireless Internet
connection and the size, speed, sound volume, and display quality of the telepresence robot could
be largely resolved in the near future, considering the continued development of communication
technology. Additionally, in cases where intrusion of other people’s privacy is of major concern,
sites with active pedestrian traffic could be avoided, more contained sites could be used and
permission from those present could be sought, and the video capturing and recording function
of the telepresence could be disabled. In this last case, observations or audio recording of
participant interactions could be used as alternative data collection methods for research
purposes.

Implications for Future Research Design

The findings of this study have useful implications for informing the design principles for
future research on using telepresence robots for language learning, some of which we summarize
tentatively below.

- The activity site should be chosen based on instructional goals and student interest, with
  attention to privacy concerns.

- To the extent possible, language learning activities should involve authentic tasks that
  happen naturally in actual environments of language use and that meet students’ learning
  needs.

- A pre-task activity should be included to orient learners to the features of the telepresence
  robot and expectations of the language learning task and to build rapport between the
  learners and activity facilitators.
• Activity facilitators should be aware of the instructional goals and students’ learning needs in choosing to maintain or change conversational topics during the activities. They should also be trained to pay attention to learner abilities and difficulties and to offer appropriate support to the learners when difficulties arise.

• The learning activity may be gamified to maximize learning motivation and engagement.

Conclusion

As an emerging technology, telepresence robots have substantial potential for promoting FL learning by providing remote language learners with virtual access to authentic physical and sociocultural contexts in a target language community. This study constitutes the first step toward understanding the ways in which authentic communicative practices facilitated by telepresence interaction may enhance or hinder language learners’ learning experiences. Our findings provide useful information for future designs of language learning activities that integrate telepresence interaction to promote FL learning. As a small-scale exploratory study, the number of participants was small. In own future work, we will conduct larger-scale studies to more systematically examine the affordances of telepresence interaction to support FL learning with more participants, more diverse environments and tasks, and richer data on learner perceptions and learning outcomes.

Acknowledgements

We wish to thank the participants of this study for their contribution and the anonymous reviewers of the manuscript for their helpful comments. Special thanks go to Zhi Zhou, who helped us recruit participants and videotape the activities.

Notes

1. For an introduction to Romo, see its Kickstarter page.


CHAPTER 3: TELEPRESENCE-PLACE-BASED FOREIGN LANGUAGE LEARNING AND ITS DESIGN PRINCIPLES

This chapter has been published as article in the Journal of Computer Assisted Language Learning. The content of the article is reprinted with the permission. The layout of the article has been updated to be aligned with the requirements of dissertation format.

You can cite the article as follows:
Abstract

Real world places are important for foreign language (FL) learning since they provide rich and authentic language and cultural contexts for learners. However, numerous FL learners are unable to travel to target-language environments for language learning purposes. This article develops the concept of telepresence-place-based FL learning and examines its design principles through a mixed-method case study. Eleven adult English learners in China and four native-speaking English instructors participated in a series of learning activities at an arboretum on the campus of a university in the US. Based on the analysis of surveys, interviews, field notes, and learning activity videos, the findings show that using telepresence robots outside the classroom provides unique learning affordance for learners, especially when proper design principles are followed in designing the learning activities.

Keywords: place-based learning; technology-mediated communication; telepresence robot; foreign language learning; design principles
Introduction

Real world places are important for language learning since they provide rich language contexts, especially when the learning objectives involve environmental, social and cultural aspects (Holden & Sykes, 2011). Meanwhile, target-language environments remain inaccessible to many language learners due to various (e.g., financial or physical) constraints. In a globally networked society, learners can now view and navigate other parts of the world remotely via telepresence technologies.

Telepresence robots present a promising tool to mediate time, space and access for learners on a global scale. The term telepresence refers to technologies that allow users to communicate with people at a remote place and to feel present there (Minsky, 1980). A language learner can control the movement and view angle of a telepresence robot remotely through a mobile device, while video chatting with a native speaker or instructor via the robot, as illustrated in Figure 3.1.

Figure 3.1. Telepresence robots in foreign language learning

Building on a pilot study (Liao & Lu, 2018) that explored the perceived benefits and challenges of using telepresence robots in foreign language (FL) learning in outdoor settings, this
study further develops the concept of telepresence-place-based FL learning (TPFLL) and proposes three design principles that can be used to guide the design of TPFLL activities.

**Literature review**

*Language learning outside the classroom*

While first language (L1) acquisition occurs naturalistically, FL learning often takes place through classroom-based instruction, which alone does not offer FL learners adequate exposure to and engagement with real-world contexts (Holden & Sykes, 2011). Without such exposure and engagement, learners with good lexical and grammatical knowledge of the target language may find it difficult to apply that knowledge in real-life situations (Larsen-Freeman, 2013). Additionally, learning in real-life settings not only helps learners consolidate their linguistic knowledge and skills, but also allows them to obtain other forms of knowledge relevant to language learning, such as cultural knowledge and the capacity to position oneself in a new community (Godwin-Jones, 2013).

In light of these considerations, research on FL learning outside the classroom has recently gained momentum. For instance, Wagner (2015) showed that “learning in the wild” allowed newcomers of a language community to make sense of the target language in daily life and argued that FL learners are unlikely to attain native-like proficiency without exposure to and engagement with authentic language-use scenarios. Richardson and Murray (2016) examined the impact of physical interaction with environments and social interaction with other people on children’s second language (L2) development. Compared with those who learned in classrooms only, children with weekly access to a forest field on school grounds displayed more verb, adjective, and exclamation usage as a result of greater physical play and excitement as well as enhanced sensory experiences.
These studies highlight the benefits of language learning outside the classroom. Unfortunately, most adult FL learners have no or limited access to use the target language in authentic, real-life contexts. According to the International Association of Language Centres (2016), only 0.25% of FL learners are able to travel to target-language countries for educational purposes. Recent advances in information and communication technologies make them increasingly effective for facilitating communication between FL learners and L1 speakers in the target-language community for language learning purposes.

Technology-mediated language learning

Various technologies have been applied in FL learning, such as video clips (e.g., Winke, Gass, & Sydorenko, 2010), video conferencing (e.g., Jauregi & Bañados, 2008; Yen, Hou, & Chang, 2015), and virtual reality (VR) (e.g., Jauregi, Canto, de Graaff, Koenraad, & Moonen, 2011; Lan, 2015). However, these technologies lack affordances to support FL learning in real-life settings robustly. For example, video clips do not allow for two-way interpersonal communication. Studies on video conferencing focus primarily on indoor activities (e.g., Yen et al., 2015). VR environments lack authenticity due to limitations of current 3D modelling techniques, with characters depicted in an unrealistic cartoonish style and tasks designed rather artificially (Falconer, 2013).

Telepresence robots hold much promise in overcoming the limitations of the above-mentioned technologies with the ability to provide VR-like experiences in authentic real-world environments and to mediate face-to-face, real-time communication simultaneously, making them ideal for placed-based learning (PBL) scenarios. Our pilot study (Liao & Lu, 2018) yielded preliminary evidence for the advantage of telepresence robots over VR, as illustrated by the learner comment below:
This robot makes me feel like manipulating a virtual person walking in a virtual world. 
This is somewhat similar to manipulating a character in a game hanging around in a virtual world. Meanwhile, I know the world where the robot is walking is a real world. It feels like a game experience but what I am experiencing is a real world. This makes me excited. (Liao & Lu, 2018, p.26)

Language learning mediated by telepresence robots

Basoeki, Dalla Libera, Menegatti, and Moro (2013) classified educational robots as wheeled robots and humanoid robots. Currently, most telepresence robots, such as Vgo, Cisco iRobot Ava 500, Double Robotics, and Keebots, are wheeled robots, which cost less than humanoid robots and are capable of moving around via wheels (see Figure 1).

Telepresence robots have been used in various situations such as supporting blended class (Bell et al., 2016) and facilitating students with physical disabilities to attend school virtually (Newhart & Olson, 2017). However, few studies have examined their use in FL learning. Building on our prior research on the perceived benefits and challenges of using a small telepresence robot to support FL learning (Liao & Lu, 2018), this study further frames this type of learning as Telepresence-place-based foreign language learning (TPFLL) and proposes three principles for designing TPFLL activities.

Framing the study as TPFLL

Defining TPFLL

The use of technology does not guarantee better learning outcomes (Amiel & Reeves, 2008). The complex interaction between the prior knowledge, skills, learning goals, and accessible resources of learners makes well-considered instructional design vital when
determining the educational efficacy of technology. The instructional design of this study is informed by the place-based education framework. Place-based education was originally proposed by Lane-Zucker (Elder, 1998) and further developed by Sobel (2005):

Place-based education is the process of using the local community and environment as a starting point to teach concepts in language arts, mathematics, social studies, science and other subjects across the curriculum. Emphasizing hands-on, real-world learning experiences, this approach to education increases academic achievement, helps students develop stronger ties to their community, enhances students’ appreciation for the natural world, and creates a heightened commitment to serving as active, contributing citizens. (p. 6)

While Sobel’s definition applies to all subjects, the unique nature of language acquisition requires a more specific definition for FL learning. For example, language learning in real-life settings not only supports learners’ linguistic development but also helps them obtain other forms of knowledge highly relevant to language learning, such as cultural knowledge and the capacity to position oneself in a new community (Liao & Lu, 2018; Godwin-Jones, 2013). Additionally, Sobel (2005) addresses the potential for engagement in local community activities, yet for FL learning, it is important for FL learners to engage in remote target-language community activities. Given these considerations, we extend Sobel’s definition of place-based education to TPFLL as the following: TPFLL is the process of using the target-language community to teach and learn a FL via telepresence technologies. This approach emphasizes real-world learning experiences and aims to enhance learners’ linguistic development, cultural awareness, and engagement with social interaction in the target-language community.
Design principles for TPFL

Design principles, i.e., the characteristics or procedure of a planned learning design in a specific educational context, constitute an effective means to bridge the gap between theory and practice in educational research (Euler, 2017). Design focused research attends to what the most influential theories for the educational practice in a specific context are and how they can be practically and collectively applied in the context from a systematic view (Dede, 2005).

Previous research has proposed different design principles for PBL in different contexts (e.g., Land & Zimmerman, 2015; McClain, 2016). For instance, Land and Zimmerman (2015) proposed and examined four design principles for place-based science learning with the support of mobiles: (1) support learners to make scientific observations outdoors; (2) use child-centered designs for mobile, informal learning; (3) use conceptual models; and (4) scaffold complex disciplinary practices in natural settings. Design principles for TPFL, however, are still lacking.

In this study, we propose three design principles for TPFL based on the PBL literature and findings of our pilot study (Liao & Lu, 2018).

1. Situate learning in real-life settings

Based on the rationale that learning is effective when embedded in cultures, contexts, and activities (e.g., McClain, 2016), meaningful social interactions situated in diverse real-life settings can be leveraged as authentic learning spaces and opportunities in TPFL. Compatible with the competency-based paradigm of FL pedagogy (e.g., Leung, 2005), TPFL should focus on pragmatic applications of language in real-life contexts (Johnson, 2008) and provide opportunities to help learners acquire and demonstrate the ability to complete specific tasks using the target language in real-life settings (Griffith & Lim, 2014). TPFL is also uniquely positioned to focus on sense-making, i.e., the ability or attempt to make sense of a concept in
ambiguous and complex situations (Klein, Moon, & Hoffman, 2006), as its real-life settings provide rich contextual information for FL learners to acquire the meanings of concepts through situated use.

2. Scaffold the learning process

To help learners obtain uptakes from the authentic social interactions in real-life settings, it is crucial to scaffold the learning process (Liao & Lu, 2018). Scaffolding, rooted in the concept of zone of proximal development (Vygotsky, 1978), refers to instructional strategies that guide learners to move progressively from current skillsets to desired skillsets (Kim & Hannafin, 2011; Vygotsky, 1978). Computer tools and online resources have been increasingly employed as scaffolds in educational research (Delen, Liew, & Willson, 2014; Pea, 2004).

It should be noted that scaffolding is not simply synonymous of support in teacher-initiated and directive instructions (Pea, 2004; Puntambekar & Hübscher, 2005). An essential feature of scaffolding is contingency, which refers to responsive and tailed support (Van de Pol, Volman, & Beishuizen, 2010). A teacher acts contingently when he or she continually diagnoses students’ current level of competence and learning problems, and then provides adapted support. Another key feature of scaffolding is fading. Instructors or facilitators need to gradually ‘fade’ off from the learning as learners are eventually expected to apply the learned skills independently (Collins, Brown, & Newman, 1989). TPFLL thus needs to consider both factors of scaffolding.

3. Enhance learner agency

Learner agency refers to the ability to take an active role and make intentional choices in the learning process (Mercer, 2012; Xiao, 2014). Learners with a personal sense of agency have more interests, motivation, self-efficacy, engagement, and autonomy (Mercer, 2012; Xiao, 2014). Learner agency is crucial for TPFLL as learners with high agency can more actively seek
learning opportunities relevant to their interests and experiences with the mobility of telepresence robots.

However, it is not enough to enhance learner agency by directly asking learners to behave more actively. Mercer (2011; 2012) claims that learner agency is a complex system and can be influenced by affective factors. For example, learners’ perception of others’ attitudes towards them will influence their self-confidence and behaviour of agency. It is thus important to provide affective support in TPFL.

While we elaborated the three design principles separately as they emphasize different factors of the learning design, they are in fact closely intertwined. For example, real-life settings can help instructors to scaffold the learning process, and successful scaffolding will positively affect learner agency (Mercer, 2011).

Research questions

This study examines how the proposed design principles can be applied in TPFL situated in a campus arboretum. We investigate the applicability and significance of these principles by addressing the following research questions.

1. What strategies can be used for situating learning in real-life places in TPFL?
2. What strategies can be used for scaffolding the learning process in TPFL?
3. What strategies can be used for enhancing learner agency in TPFL?

Methodology

Participants

Four native English speakers, all doctoral candidates at the university chosen as the research site, were recruited as language instructors. Three majored in Applied Linguistics and
had some FL teaching experience; one majored in Learning, Design, and Technology and no prior FL teaching experience.

Eleven adult EFL learners in China were recruited as remote participants by sending recruitment letters to college students and young professionals by email or other online channels. The demographic information of the learners is summarized in Table 3.1.

Table 3.1. Demographic information of the learners

<table>
<thead>
<tr>
<th>Items</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Female: 11 (The only male participant didn’t finish the survey)</td>
</tr>
<tr>
<td>Age</td>
<td>Range: 22-32; Mean: 25.6; SD: 2.9</td>
</tr>
<tr>
<td>Occupation</td>
<td>College students: 3; Young professionals: 8</td>
</tr>
<tr>
<td>Degree</td>
<td>Bachelor: 7; Master: 4</td>
</tr>
<tr>
<td>Years of learning English</td>
<td>Range: 8-16; Mean: 10.7; SD: 2.4</td>
</tr>
<tr>
<td>English proficiency</td>
<td>Intermediate (College English Test Band 4): 4; Advanced (College English Test Band 6): 7</td>
</tr>
</tbody>
</table>

1. **Physical settings**

The research site was a campus arboretum at an American university, selected because of its social, cultural and mobile affordances to situate the learning. The arboretum has numerous aesthetic features that could trigger conversational topics, is a central social gathering space, and hosts various regional events. It was also a good fit in terms of mobility and logistics: with even paved paths and light pedestrian traffic, it provided a safer and quieter setting than other campus places.

As part of the design, we selected focal spots of the arboretum to structure the conversation activity around. These included a water fountain, bamboo walkway, tree and shrub collections, a stone sculpture sundial, the main circle lawn, a witness tree, and the event pavilion. Figure 3.2 shows the arboretum map and pictures the focal spots.
2. **Technical settings**

As shown in Figure 3.3, we used a Keebot, a mid-size telepresence robot, to mediate the learning. The iPad on the robot was connected to the Internet either through the campus wi-fi network or the 4G network on the researcher’s iPhone hotspot, depending on speed and satiability. A supplementary website (http://rill360.com) was set up for the EFL learners to orient themselves to the arboretum and collect demographic information and survey data.

*Figure 3.3. Communication via the telepresence robot*
Procedure

The learning activity was organized into four parts: orientation, warm-up, guided arboretum tour, and final debrief. The activity design closely followed the design principles proposed above, as summarized in Table 3.2.

Table 3.2. Design principles and specific design of the study

<table>
<thead>
<tr>
<th>Design principles</th>
<th>Study design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Situate learning in real-life settings</td>
<td>• Learners and instructors discuss the objects and relevant social activities along the route</td>
</tr>
<tr>
<td></td>
<td>• Learners play the role of a tour guide for encouraging pragmatic language use</td>
</tr>
<tr>
<td>Scaffold the learning process</td>
<td>• Orient learners to the remote environment and the learning tasks with pre-task activities</td>
</tr>
<tr>
<td></td>
<td>• Instructors check and facilitate learner understanding</td>
</tr>
<tr>
<td></td>
<td>• Instructors help learners express what they want to say</td>
</tr>
<tr>
<td></td>
<td>• Learners play the role of a tour guide as a strategy of fading</td>
</tr>
<tr>
<td>Enhance learner agency</td>
<td>• Make the conversation open and pay attention to topics that learners are interested in</td>
</tr>
<tr>
<td></td>
<td>• Encourage learners to control the robot and its view angle to explore the environment</td>
</tr>
<tr>
<td></td>
<td>• Learners play the role of a tour guide for engaging learners in the activity more actively</td>
</tr>
<tr>
<td></td>
<td>• Encourage learners and minimize their frustration</td>
</tr>
</tbody>
</table>

All learners participated in an orientation one week before the guided arboretum tour. The researcher trained them to control the robot through their mobile devices. They then reviewed introductory material to the arboretum on a supplementary website, signed a consent form, and completed a demographic information survey. The instructors also familiarized themselves with the arboretum and the focal spots one week before the tour activity. A guideline was provided on the supplemental website to aid the instructors in developing scaffolding
techniques. The guideline proved especially helpful to the instructor without prior FL teaching experience.

The warm-up happened before the tour activity and lasted 15-20 minutes. During that time, the learners connected to the robot, reviewed 10 English words related to the focal spots, listened to a recorded conversation about hosting a wedding ceremony in the arboretum, and summarized the listening material orally in two minutes. The learners and instructors also introduced themselves to each other.

During the one-hour arboretum tour, each learner controlled the robot and communicated with an instructor one-on-one. Each tour covered five focal spots (see Figure 3.2; the fountain and bamboo walk were combined as one focal spot) with 10-15 minutes per spot for learner-instructor conversation. At the end of the conversation around each focal spot, each learner played the role of a tour guide to describe that spot.

At the end of the tour, the learners debriefed what they had learned, were interviewed by a researcher, and completed an exit survey regarding their perceptions of the learning outcome of the study.

Data collection

This study used a mixed methods approach to collect and analyse data. Four types of data were collected, including activity videos, field notes, interview transcripts, and surveys. Two cameras were used to video-record each of the 11 tour sessions, one installed on the robot to capture the video from the learner’s perspective, another held by a researcher to capture both the robot and the instructor. A second researcher took field notes throughout each tour. Altogether 11 hours of video interactions were recorded, and 45 pages of field notes were produced.
Eleven learners were interviewed in Chinese by one researcher, a native speaker of Chinese, for 15 minutes at the end of each tour. The four instructors were interviewed in English for approximately one hour after they had completed all sessions that they facilitated. The research team transcribe all interviews and translated the Chinese transcripts into English. A total of 108 pages of translated interview transcripts were produced.

All eleven learners completed a demographic information survey, and ten learners completed the exit survey on their perceived learning outcome of the activity.

Data analysis

A qualitative analysis (McMillan, 2012) was used to generate an open coding scheme. Specifically, all transcripts and field notes were read through first. Chunks of text related to the research questions were marked up as codes in Dedoose, an online qualitative research tool. After several iterations of merging and splitting until each code could not be split or merged further, all codes were categorized into different themes based on our research questions. Finally, 24 themes were summarized from 95 codes and were further grouped into 3 categories based on our research questions.

Survey results and the videos were used to triangulate the findings from the interviews and fieldnotes. We strategically sampled video clips relevant to the coding themes and transcribed them using Jefferson’s (2004) transcription system, which marks voice pitch, pauses, laugh, and non-verbal gestures, to show the learning scenarios at the micro level. Content that have little relevance to the research questions is removed from this paper and noted as ‘…’.
Findings

Research question 1: Strategies for situating learning in real-life places

The physical environment of the arboretum allowed for conversational topics to emerge naturally as the robot moved along the route, as illustrated by the episode generated around a stone sundial (see Figure 4.4), one of the focal spots below. Note that all names used in the paper are pseudonyms.

Figure 3.4. Conversation about the sundial

Josephine (Instructor): {do you see that sculpture over there?
{ ((right index finger points at the sundial))
Xiaoyun (Learner): yes
Josephine: okey (.) can you guess what that is?
...
Xiaoyun: okey, is that-

(1.1)
Xiaoyun: em
Josephine: it has to do with time (1.0) and the sun

(1.8)
Xiaoyun: <↑o:h> (.) e:mh (0.4) hh em
i↑know: it- it ca:n (.) make us know the times. {in chinese,
Josephine: {((smiles and nods head))
Xiaoyun: we call it ri gui, but i don’t know emh (.)
< what can we call em> {i:ts english name
In this episode, the instructor, Josephine, naturally initiated the conversation by pointing to the sundial and then asking a question. Xiaoyun saw the sundial through the camera of the robot and interacted with the instructor. When the facilitator found that Xiaoyun struggled to answer, Josephine offered contextual clues such as the sun and time. After Xiaoyun deduced the meaning and learned the pronunciation of the sundial, Josephine provided further information by interacting with the sundial, i.e., walking around it to show its size and scale.

In the interviews, Qing, another learner, discussed how she perceived the environment as supporting their language comprehension by connecting words with physical features in a rich and authentic language context.

*Qing: I can see the stuff and make connections. I think it has better relevance. It’s not like when I read a word in a book but have no idea what it looks like.*

Additionally, the arboretum facilitated the display and acquisition of cultural knowledge, as illustrated by the following episode around the seasonal pumpkin display in the arboretum (see Figure 3.5), a conversationally rich feature that was not covered in the initial design but became available later.
Zhiwei (Learner): em so em ↑why(.) why there(.) why there are so much pumpkin there?
Brad (Instructor): that’s a good question. so tomorrow in the united states is a em ↑holiday
(2.5)
Zhiwei: ↑oh oh oh, i(.) i know the, &haha, i don’t know how to “express”
(1.8)
...
Zhiwei: they- they wake up for(.) em they &ha they drink drink and the kids(.)
little kids they will em:: get the: candies &ha
Brad: {that’s right, that’s exactly right(.) that’s ↑halloween
{((smiles and nodes head))
(1.6)
Zhiwei: oh <halloween>
Brad: you got it(.) {and ↑also, {we will take some pumpkins
{((left hand pats on a pumpkin))
Zhiwei: {okey
Brad: and we will make a face(.) in the pumpkin {with a knife (0.8)
{((mimics carving))
cut it up, and put in a candle.
{((mimics cutting up and putting candle into the pumpkin))
(2.9)
Zhiwei: oh ↑okey we have ↑some i i see some pictures em
in the: in the shopping ↑mall. the the pumpkin likes the ghost
In this episode, the pumpkin display was of great interest to Zhiwei as it was culturally unique. Zhiwei initiated the topic by asking why there were so many pumpkins. Brad, the instructor, hinted that it was because of an American holiday. Although Zhiwei could describe some traditional activities of the holiday, she did not know the name of it. The instructor then told her it was Halloween and they were able to engage in conversation about the traditions relevant to pumpkins. Then Zhiwei connected Halloween to her prior cultural experience: pictures of carved pumpkins in a shopping mall.

Therefore, the tour provided not only arboretum-related but also cultural-related topics. These topics were emergent and varied from learner to learner. For example, when Meimei saw an oak tree, she connected the wood of the tree to vampire folktales and pop-culture, spurring an emergent conversation with the instructor about oak stakes and vampire slaying. In other conversations, some learners explained that the bamboo near the arboretum entrance was “food of the panda”. Additionally, the learners observed how the arboretum visitors engaged in social activities such as playing Frisbee and walking pets. The arboretum was also a popular wedding venue, especially in the fall. These social and cultural activities provided a unique and rich socio-cultural context for the learners.

The results of the survey on perceived benefits, summarized in Table 3, confirmed that situating the learning in the place can benefit learners in listening, speaking, vocabulary, and cultural awareness. However, learners were neutral in regard to whether the activity helped improve their grammar. This could be because the activities focused on pragmatic language use and sense-making. Consequently, the instructor did not correct learners’ grammar errors unless they impeded understanding.
Table 3.3. Survey results on perceived learning outcomes

<table>
<thead>
<tr>
<th>Questions</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>This activity is helpful for improving listening.</td>
<td>8</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>This activity makes me more confidence when speaking in English.</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I don’t feel this activity is helpful for increasing vocabulary.</td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>I don’t feel this activity is helpful for improving grammar.</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>This activity is helpful for improving accuracy of my speaking.</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I don’t feel this activity is helpful for improving fluency of my speaking.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>I don’t feel this activity is helpful for improving cultural awareness.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

Learners also noted on the exit survey the extent which they considered various places to be potential learning sites for TPFLL. As shown in Figure 3.6, they seemed to agree that such places as libraries, popular spots on campus, apartments, and bus stops, etc., could work well for TPFLL.

Figure 3.6. Potential learning sites
However, a challenge of situating the learning in real-life settings was the flexible learning content. Although we had given the learners 10 key words before the tour, many conversational topics were unplanned but emerged spontaneously. For example, the pumpkin display around Halloween was not considered initially. The emergent topics also varied from learner to learner. While the dynamic and emergent aspect of conversion was pedagogically and affectively rich, it was challenging for learners with low proficiency to improvise on the emerging topics without sufficient preparation, as Xiu mentioned in their interviews:

Xiu: Actually, there was a lot of stuff that I wanted to talk about, but I couldn’t because of my limited vocabulary.

Research question 2: Strategies for scaffolding the learning process

As mentioned earlier, two essential components of scaffolding TPFL are fading and contingency, which require the instructors to frequently detect learners’ current level and emerging issues during the conversation and adaptively support them. Therefore, we investigated the second research question from three aspects: detection of learner level and issues, adaptive support, and fading.

1. Detection of learner level and issues

Before the activity, we gave the learners a 10-word quiz and asked them to listen to a brief description about booking a place for a wedding ceremony and then orally summarize it (see Figure 5) to gauge their English proficiency. We also asked the instructors to identify and support learners’ issues emerging in the conversation. Emmerich, one of the instructors, demonstrated an effective strategy for detecting learner issues, as shown in the following episode.
Emmerich (instructor): {can you guess (. .) <what people do> in this pavilion?}
{((right index finger points to the pavilion))}
(2.1)
Meimei (Learner): to: em: (. .) em it is used like (. .) wedding place
Emmerich: very good. {↑yeah: a lot of people use this to get married
Meimei: { and:
Emmerich: they have a {wedding (. .) ceremony.
Meimei: {get married
(1.1)
Meimei: oh::
Emmerich: do you know the word ↑ceremony
(1.4)
Meimei: ceremony em yes {celebrate for
Emmerich: {so, in a
(1.2)
Meimei: em (. .) people together talk together and eat together &ha
Emmerich: exactly right. so in a ↑wedding in the united states,
we have ↑two: parts (. .) first we have {the ceremony (. .)
Meimei: {two parts
Emmerich: {↑ceremony (.) right and then we have the↑reception
Meimei: {ceremony em okey
                    (1.9)
Meimei: ↑reception
Emmerich: reception
Meimei: what’s what
Emmerich: so for for the americans the ↑reception (.) is a big party
it is where you eat food
Meimei: °oh°
Emmerich: and sometimes you drink (. ) wine ( .) and you celebrate people
and dance (0.5) but the ↑ceremony for us ( .) is the very serious part
Meimei: °em° (1.1) ↑oh::
Emmerich: for us the ceremony is very {everybody sits very serious ( .) and some people cry ( .)
{((mimics sitting very seriously))
                    (1.1)
Meimei: oh:

In this episode, Emmerich employed pauses, direct questions, and phonological means as strategies to check learner comprehension. For example, when he mentioned the pavilion could be used for a wedding ceremony, he stressed the word ceremony, and then paused for about 1.1 seconds. When Meimei mentioned nothing more about the word, he asked her if she knew the word. In explaining the two parts of a wedding, i.e., a ceremony and a reception, he used a rising tone with reception stressed, and then paused again to check on Meimei’s comprehension.

Weinan, another learner instructed by Emmerich, commented how she felt about Emmerich’s detection strategy and adaptive support.

Weinan: The instructor was flexible. Actually, he (the instructor) knew when I was unable to express myself or when I didn’t get what he said based on my reaction ... Then he would adjust his language and way of talking to help me understand.
2. Adaptive support

When any misunderstanding was noted, the instructors strived to provide adaptive support to help the learners understand. The instructors frequently resorted to simple-word explanations and employed body language and embodied communication (acting things out) as necessary to help learners make sense of new words and concepts, as illustrated in Figure 6. For example, in the episode above, Emmerich mimicked how people sit seriously in a ceremony to help Meimei understand its different meanings. These strategies become more effective with environmental support. In Figure 3.8, Josephine pantomimed a golf swing on the lawn, and the turf provided context for the learner to distinguish lawn sports from others.

Figure 3.8. Embodied interaction

In her interview, Meimei mentioned the benefit of the adaptive support she received, particularly with respect to the instructor’s use of simpler words and environmental support to explain new vocabulary items.
Meimei: What was most impressive was his (Emmerich) explanations of the words that I didn’t know using words that I knew. This helped me understand and remember the new words much better. Normally I feel it difficult to really understand and remember new words out of context. Today when he explained the words to me by pointing at things and using easier words, I got them quickly and felt very good. That good feeling also helped me remember the words better.

Another strategy used by the instructors to support new vocabulary learning was through spelling out long, difficult words on a small, portable whiteboard, as Emmerich did in Figure 3.9.

Figure 3.9. Whiteboard

Xiu, who participated in the current study as well as our pilot study (Liao & Lu, 2018), compared the experience with and without the support of whiteboard.

A: What do you think of the use of whiteboard?

Xiu: I think it’s quite good. Last time (the pilot study) we didn’t involve it. When I asked a question about grass this time, I recalled that I had asked a similar question last time, but I couldn’t recall what the word I learned (in the pilot study) was.

A: Do you remember the word you learned this time?

Xiu: Yes, I remember the word (lawn) this time after he (Emmerich) wrote it down for me.

A: Could you tell me more about the word in the last activity (the pilot study)?
Xiu: It was like “lawn” or what, I only remembered it when he (Colman, the instructor in the pilot study) explained to me and I knew it was not “turf.” You see I couldn’t remember what the word (I learned in the pilot study) was. But I remember this time after (Emmerich) wrote it down. Writing the words down is useful.

Actually, the word Xiu had learned in the pilot study but couldn’t recall is “Quad” (Liao & Lu, 2018). In the pilot study, Colman, the previous instructor, introduced the quad at the main gate of the campus to Xiu and spelled it orally. Although the place provided Xiu with contextual information and Xiu still remembered that quad was related to grass or “lawn,” she had forgotten the spelling of the word. Writing a word down helped Xiu obtain its spelling, pronunciation and meaning in context.

3. Fading

While all instructors attempted to gradually withdraw input from the conversation, this was not always possible given the limited time of each session. As a trade-off strategy, we asked learners to play the role of a tour guide immediately following the activity at each focal spot to force the instructors to fade off. Most learners indicated in the interview that the role-play helped them use the language independently. For example, Xunian expressed how she became more confident throughout the activity culminating in the role play.

Xunian: I think it’s necessary ... I was nervous when I did it the first time. Later I became more and more relaxed. ... I had more time to speak (during the role play), so I could spend more time organizing my speech. If he (the instructor) didn’t arrange this for me, I might speak very little because of being shy to talk (in the FL), so this step really encouraged me to express my thoughts.

Ruanning further noted that the role-play could help the instructor check whether the learner understood everything.
Ruanning: This would help the instructor know how much I understood her explanations. Then we could better exchange our thoughts. Without this (role play), she might not fully know how much I understood or what I thought.

Kerianne, one of the instructors, however, felt that for learners with lower proficiency, she would have liked to not fade off fully, as they appeared to need further scaffolding or support during the role play.

Kerianne: Maybe, it could be something like, if role play is going really well for a student, then great, that worked great. But if role play, if the student isn’t talking a lot, maybe instead just going “OK, great” and moving on. Maybe the teacher can ask questions, too, like “OK, well, that’s a good start, but remember about this tree? Do you remember anything about this tree?” and can prompt them to try to talk more.

Research question 3: Strategies for enhancing learner agency

As mentioned earlier, the conversational topics and learning content were flexible, as different learners had different backgrounds, prior experiences, and interests. Additionally, the environment actually varied among the tour sessions, with the change of weather, season, decorations and displays, and social events. It was critical to enhance learner agency, as learners with higher agency took more advantage of the learning opportunities in this dynamic environment.

We identified learners’ ‘agentic behaviors’ as incidents where they appeared to take control, take action, or play an active role in the learning process (Mercer, 2011). The following field note details a scenario in which Xunian, a learner, actively controlled the robot to find what she could talk about during the role-play without instructor guidance.
After finishing the conversation with Xunian in the pavilion, Brad crouches down near the robot and tells Xunian to do the role-play again. Xunian says “Ok, the pavilion is …” and then she starts turning the robot around without Brad’s guidance. While the robot has turned around almost half a circle, she says ‘pavilion is a big place’ … and continues to turn the robot. After she turns the robot more than a full circle, she says “The view is very great. We can see the hill far away from here… and the pavilion has a big roof… so maybe the raining day the people like to play or sit here.”

Another example is that when we intentionally facilitated conversations between the learners and arboretum visitors, some learners handled the impromptu conversations well, while others were rather quiet. The following field note details how Xiu actively engaged in an impromptu conversation.

Now there are several people walking while Xiu is driving the robot to the next spot. We meet some ladies and tell Xiu she can say hi to them. After that, they have more small talks about the arboretum, the weather, and what we are doing. The ladies introduce themselves to Xiu and one of them says she was in China in May. And they talk about Shanghai and Shuzhou cities, and the difference between gardens in China and America. They smile during the conversation.

These two episodes demonstrated how learners with higher agency behaved agentively during the activities. Such activities as controlling the robot, choosing conversational topics, and talking to visitors not only supported learners’ output, but also made the learning process more learner-centered and better adapted to their personal interests, background, and prior experience. To enhance learner agency, we encouraged learners to leverage the mobility of the robot and the ability to adjust the view angle to capture what interested them. Learner feedback indicated that
most learners held a positive attitude towards the mobility feature, as illustrated by the interview transcripts below.

_Xiu: The advantage is that I can control my view angle, without telling the other person: “please turn me around.” This is good. So that’s what the feature is, more automobility._

_Mei: I could see the instructor, and the instructor could see me. Therefore, we interacted behaviourally and reciprocally, which is better than the way in which an instructor just asks me to do this or that. For example, Josephine (the instructor) let me look around first before asking me what I was interested in. I feel pretty good about this experience._

Secondly, the role-play was used to motivate the learners to take a more agentive role in seeking and understanding information in the activities, as they would need to apply the information afterwards. For instance, after Qing debriefed what they had observed and discussed at the last focal spot, Kerainne, an instructor, thanked Qing and said “Thank you for teaching me about this again. I will certainly pay more attention next time.”

Lastly, we asked the instructors to give the learners more affective support since learner agency is influenced by learners’ personal emotion and how they perceive other people’s evaluation on their learning ability (Mercer, 2011). Our instructors used positive responses such as “it is a good start” and “you did a good job in the role-play.” When learners paused to figure out a word, the instructors provided positive reinforcement, such as “take your time.” We also asked instructors to only correct grammar errors that affected understanding, since excessive error correction would likely break the flow of the conversation and cause learner frustration. Learner feedback to the instructors were overwhelmingly positive, with such interview comments as “The instructor was very nice and responsible,” and “I feel very relaxed.”
Discussion

The results of the study shed useful light on our research questions as well as the applicability of the three design principles we proposed for TPFLL activities.

*Design principle 1: Situate learning in real-life places*

Three strategies worked well for situating learning in real-life places. First, the instructors and learners used objects in the environment as an anchor to start a conversation. The environment also helped instructors use embodied interaction such as gesture, body language, or interaction with the objects to explain concepts. Second, they expanded the topics to relevant cultural traditions. Finally, they used their prior knowledge and past experiences to expand conversation. The survey and interview results showed that the TPFLL activities enhanced their lexical knowledge as well as listening and speaking. These findings are in line with the finding of our pilot study (Liao & Lu, 2018) and the literature (e.g., Hummels & Van Dijk, 2015). The survey results also show that TPFLL could be situated in various other teaching contexts, such as popular spots on campus and bus stops.

Some learners with lower language proficiency found it challenging to handle unplanned topics and learning content. Given the dynamic nature of the real-life settings that TPFLL presents and variation in learner proficiency, interests, and experiences, it is important for instructors to effectively scaffold the learning process, as our second design principle states.

*Design principle 2: Scaffold the learning process*

We adopted the view that the learning process could be scaffolded by instructors, learning materials, and tools (Pea, 2004) and examined three aspects of scaffolding in our study: detection of learner level and issues, adaptive support, and fading.
Our findings indicated it was useful to gauge learner proficiency before the learning activities with such materials and instruments as surveys, word quizzes, and listening or speaking tests. During the activities, our instructors frequently checked on learners’ comprehension or production problems with pauses, questions, phonological means, etc., and used different strategies to help them as necessary, such as explaining a concept using simpler words, slowing down, using embodied interaction, interacting with objects in environment, and using a whiteboard to spell out a new word. It was clear that compared to classroom settings, the real-life settings in the TPFL activities allowed for a greater range of adaptive supporting strategies. As the activities moved along, the instructors needed to fade off as the learners were expected to use the language to perform tasks independently. We used role-play as a strategy of fading, which required the learners to act as a tour guide at the conclusion of the conversation around each focal spot. Most learners felt that the role-play encouraged them to engage in useful language output in a more authentic environment than classroom settings and allowed the instructors to check their linguistic, conceptual and cultural understanding. Some learners with lower proficiency, however, felt the need for further scaffolding, calling for more individualized activity design to support varied learner needs.

*Design principle 3: Enhance learner agency*

Learner agency is a crucial factor in TPFL as more agentive learners seek learning opportunities more actively in the TPFL environment. Existing research on learner agency has focused on classroom or online settings (e.g., van Lier, 2008; Xiao, 2014). We examined scenarios that demonstrated learners’ ‘agentic behaviour’ (Mercer, 2011) during the TPFL activities.
To enhance learner agency, we encouraged learners to control the movement and view angle of the telepresence robot throughout the activities. Learners felt that the mobility of the robot not only helped them find conversational topics, but also made the learning process more learner-centered. Other strategies included emphasizing learners’ active role in the role-play and providing learners with more affective support during the activities. The field notes and interview results showed that learners held positive attitudes towards these strategies.

**Conclusion**

In this paper, we framed the act of FL learning mediated by telepresence technologies outside classroom settings as TPFLL, proposed three design principles for designing TPFLL activities, and examined their applicability in a case study in a campus arboretum at an American university. Our results show that the learning activities designed based on the principles effectively supported TPFLL.

While we implemented the TPFLL activities one-on-one, multiple learners can indeed participate in a TPFLL activity via the same telepresence robot. Small group activities come with the advantages of improved scale and collaborative learning but will require more coordination among the learners and the instructor. Furthermore, while we situated our TPFLL activities in a specific outdoor setting (i.e., a college campus), activities designed following the principles proposed here can be implemented in other indoor or outdoor pedagogical contexts as well. For example, an instructor may introduce and orient engineering students to the facilities in an engineering laboratory by walking through the laboratory with the students and a telepresence robot. The instructor can initiate topics based on what the students see, detect and help address their problems and challenges in expression and comprehension, and encourage them to actively explore, take note of, and ask questions about the environment. The instructor also has the option
to make the discussion more structured and formal, if necessary. Finally, TPFL activities can be easily combined with teaching in normal classroom settings. With careful planning, small group activities may be implemented either as part of a regular class (assuming adequate class time) or as after-class simulated field work projects to allow students to learn within and from real-world environments that they would otherwise not have immediate access to.

In our future research, we intend to examine how the design principles may be further refined for other pedagogical contexts and for classroom learning and to more rigorously consider the role learning tasks play in the design principles. We also intend to conduct comprehensive analyses of longitudinal learning process and outcome data to gain a deeper understanding of how the desired types of learning (e.g., situated learning, scaffolded learning, and agency-driven learning) occur during the activities, how critical learner awareness and behavior (e.g., contextual awareness, proactive and reflective behaviour, and cues-based conversation) change over time, and the specific learning gains TPFL activities lead to.

References


CHAPTER 4: GEOSEMIOTIC ANALYSIS OF TELEPRESENCE-PLACE-BASED FOREIGN LANGUAGE LEARNING

Abstract

Many studies have demonstrated that learning a foreign language in a setting outside of the classroom can improve the communicative use of the target language. However, many places remain inaccessible to learners, due to physical limits of mobility and health, socioeconomic factors, or political or temporal restraints. Our previous studies have shown that telepresence robots can be used to immerse learners in a remote place for the purpose of learning a foreign language. This study further uses geosemiotics as a theoretical lens to examine the foreign language learning process via a telepresence robot.

Keywords: Telepresence robot, foreign language learning, place-based learning, geosemiotics
Introduction

The idea of “place” is important for foreign language learning, since places in real-life settings provide rich context for not just linguistic knowledge but also social and cultural aspects in the target language countries (Holden & Sykes, 2011). Learning in a place can comprise a unique, memorable, and motivating experience that is difficult to replicate elsewhere (McClain, 2016). At the same time, many places remain inaccessible to learners, due to physical limits of mobility and health, socio-economic factors, political, or temporal restraints. For example, it may be difficult for a foreign language (FL) learner to travel to a target language country to learn the target language, because of the cost of travel, accommodations, living expenses, and tuition.

In order to help FL learners gain access to given places in the target language countries, our previous studies (Liao & Lu, 2018; Liao et al., in press) proposed a solution called telepresence-place-based foreign language learning (TPFLL), in which telepresence robots are used to simulate the experience of being in a target language country and, in that way, virtually support more immersive learning experiences. The term “telepresence” refers to technologies that provide remote users with the feeling of being present at another location (Minsky, 1980). Telepresence robots allow users in one location to control the movement and camera angle of a robot at another site, thus providing a FL learner the ability to explore the environment around the robot and communicate with native speakers at that location.

In this study, we further examine the influence of place by using a theoretical lens of geosemiotics. The findings may enhance understanding of the sense-making process; they may also help guide the design of telepresence robots and associated learning activities for foreign language learning use.
Literature Review

In this section, research on the use of telepresence robots in foreign language learning is investigated first. Next, various approaches to foreign language learning via telepresence robots are discussed. Finally, the framework of geosemiotics and its potential use in analyzing learning activities is introduced.

Foreign language learning via telepresence robots

Robot-assisted language learning has gained more interest among educators and researchers in recent decades, with most studies focusing on the use of artificial-intelligence robots as learning partners or teachers. For example, Mazzoni and Benvenuti (2015) explored whether a humanoid robot acting as a language learning partner could help preschool children learn English as a second language. Hong et al. (2015) investigated how a programmable humanoid robot supported foreign language learning in elementary schools. However, Kwok (2015) found that more Hong Kong secondary school students preferred a human teacher to a robot teacher since human teachers can understand students’ complex needs and current robots are limited in their ability to communicate with learners naturally and dynamically, because of technological limitations on speech recognition and artificial intelligence.

Telepresence robots support a remote learner or a teacher to move the robot around at a given place and to have live-streamed communication with people at that site. The telepresence robot becomes an avatar of the remote teacher or learner at the local place since the teacher or learner can hear what the robot hears and see what the robot sees. While many studies have been conducted on the application of telepresence robots in distance education (e.g., Gomoll et al., 2016) and STEM education (e.g., Lee et al., 2017), only a few studies have focused on using telepresence robots specifically in foreign language learning. For example, Tanaka et al. (2013)
explored the use of child-operated telepresence robot systems in international communication between two distant classrooms, one in Japan and the other in Australia. The results showed that the child-operated telepresence robot can effectively support remote communication, especially if the robot is equipped with two mechanical arms. In a later study, Tanaka et al. (2014) further compared the behaviors of children in Japan who used either Skype or a telepresence robot to communicate with an English teacher in Australia. The results showed that the use of telepresence robots effectively promoted communication between the teacher and the learners. In addition, children using telepresence robots were less likely to “freeze” when communicating with the English teacher, compared to their tendency to become motionless when facing a teacher over Skype.

However, the research in these studies (e.g., Tanaka et al., 2013; Tanaka et al., 2014) took place primarily in classrooms. The mobility of telepresence robots makes it feasible to create learning environments outside of classrooms, in order to engage FL learners with real-life contexts. Liao and Lu (2018) examined the perceived benefits and challenges of using Romo, a small telepresence robot, for FL learning outdoors. Three adult FL learners in China had a one-on-one session with an English instructor. In each session, a learner controlled the robot to simulate moving around at important buildings on a campus of an American university. The instructor played the role of a tour guide, discussing with the learner the objects and activities around one of the campus’ main buildings. The results showed that the telepresence robot provided a more authentic learning environment than other technologies, such as virtual reality, for FL learners. The environment also motivated the learners and the instructor to generate more conversational topics. However, some learners had difficulty with the impromptu topics that would emerge from natural conversation.
To address these problems, a second study (Liao et al., in press) framed the FL learning via telepresence robots outside classrooms as telepresence-place-based foreign language learning (TPFLL) and proposed three design principles to guide the design of learning activities. Specifically, the design principles were: (1) situate the learning in real-life settings, (2) scaffold the learning process, and (3) enhance learner agency. A case study was conducted to examine the applicability of these principles at an arboretum on an American university campus. The results showed that these three principles can effectively improve FL learners’ learning experience.

Form-focused and meaning-focused language instruction approaches

It is important to distinguish two foreign-language instruction approaches, focus on forms and focus on meaning, to understand the design and the process of TPFLL. *Focus on forms* (Long & Robinson, 1998) is a traditional foreign-language instruction approach, which divides the learning goals into units such as phonemes, vocabulary, intonation patterns, and grammar (Schmidt, 1995). These units are presented to FL learners in a predetermined sequence based on the difficulty of the learning goals. FL learners learn and practice the units consciously, such as discussing the grammar rules, memorizing dialogues, comparing the FL with the native language, and making error corrections (Herbert, 2004).

Although a *focus on forms* approach can prepare FL learners well for FL tests, it has been found that FL learners who focus only on discrete vocabulary and grammar have trouble using the linguistic knowledge appropriately in spontaneous interaction in real conversational situations. Communication in real-life settings not only necessitates discrete linguistic knowledge, but also requires the learners to be exposed to and practice the language in a given social and cultural context ( ). In addition, a *focus on forms* does not cater to learners’ individual
needs, learning styles, and preferences (Long & Robinson, 1998). As a result, the process of a focus on forms can be boring and may decrease learners’ motivation (Celik, 2019).

In contrast, a focus on meaning, i.e., meaning-focused language instruction, involves FL learners in rich and meaningful use of the target language in real-life contexts (e.g., Celik, 2019; Long & Robinson, 1998). This approach is based on the idea that people learn a foreign language by following the principles of learning their mother tongue (Long & Robinson, 1998). Accordingly, the purpose of meaning-focused language instruction is not to learn the language forms such as vocabulary and grammar based on a structured syllabus, but instead improving the communicative language use in real-life contexts (Celik, 2019). From the perspective of focus on meaning, language will be acquired implicitly and incidentally through meaningful communication, in the same way that a child learns the mother tongue through communication with his or her parents, caregiver, and friends at various places (Krashen & Terrell, 1983).

In addition, a focus on meaning is more flexible than a focus on forms, since a focus on meaning doesn’t have a fixed syllabus and thus can motivate FL learners by catering to the learners with adaptive content to meet their individual needs (Celik, 2019). In our previous studies (Liao & Lu, 2018; Liao et al., in press), the FL learning activities using telepresence robots were designed as a focus on meaning, because the telepresence robots allowed the remote FL learners to control the robots to move around at a remote place outside classrooms and have a conversation with native speakers at that location. The environment at the remote location provides rich context for naturalistic conversation and a meaning-focused language instruction.

Geosemiotics

The term “geosemiotics” was proposed by Scollon and Scollon (2003) for studying “discourses in place.” (p. 2) The discourses in most geosemiotic studies focus on signs in the
world, such as road signs (Hamid, 2015), shop signs (Lou, 2007), advertisements, posters (Scollon & Scollon, 2003), and signs in foodscapes (Abas, 2019). These studies investigated how places, as well as other relevant factors such as the people, social events, and the culture in those places, influence the design and meaning of the signs.

Some geosemiotics studies have expanded discourses to include speaking, writing, and behaviors in either physical or virtual places. For instance, Bortoluzzi and Trevisan (2009) explored how the Second Life, an online virtual environment, shaped e-users’ identity by analyzing such forms of discourse as chatting via a keyboard, speaking through a microphone, and looking at images on the walls in the virtual environment. Wohlwend et al. (2011) used the framework of geosemiotics to investigate the web play of children in Webkinz, a virtual online community for kids who have bought Webkinz toys. Moreover, Pierce (2012) examined how the setting and materials such as desks, bookshelves, and PowerPoint slides, influence the interactions among students and the teacher in an English as a Second Language classroom.

A core concept in geosemiotics is indexicality. The term as originally used refers to how the meaning of words such as “this,” “that,” “here,” “there,” “now,” and “then” will be shaped by the spatial, temporal, and social contexts of the language (e.g., Hodges et al., 1986; Scollon & Scollon, 2003). For example, if a professional explorer stands on a hill with a pointing-down gesture and asks “What is this?”, the answer might be the geographic name of hill. By contrast, if a child asks a same question in the same situation, a more appropriate answer might be “the ground.” The meaning of “this” in the question is then indexed differently by the place and the people in the place.

Scollon and Scollon (2003) expanded their interests in indexicality from pronouns and adverbs to general discourse, especially to signs. They systematically studied the factors that
influence the meaning of signs in a specific place and proposed a semiotic aggregate framework. The term “geosemiotic aggregate” is defined as “multiple semiotic systems in a dialogical interaction with each other” (Scollon & Scollon, 2003, p. 12), with multiple factors interacting to influence the meaning of the discourses. They specifically identified three semiotic systems: *interaction order, visual semiotics, and place semiotics.*

Interaction order in geosemiotics refers to “the current, ongoing, ratified (but also contested and denied) set of social relationships we take up and try to maintain with the other people who are in our presence” (Scollon & Scollon, 2003, p. 16)—in other words, the social relationships of people with different social roles at a place. For instance, an old man who kneels on a street with a hat in front of him plays the social role of beggar. Another man giving money to the beggar plays the role of philanthropist. The social roles of the people and their relationships influence what and how they will interact. Scollon and Scollon (2003) examined in detail the interpersonal distance of the people, the people’s appearance and behavior (*personal front*), and the unit of people (*single*, one person *with* another, etc.) at a given place.

The second semiotic system, *visual semiotics,* refers to “all of the ways in which meaning is structured within our visual fields” (Scollon & Scollon, 2003, p. 11). From the perspective of visual semiotics, the way discourse or materials are presented will influence the meaning of the discourse (Lou, 2017). For example, Scollon and Scollon (2013) found that the layout of signs influences the meaning of the text in the signs. Lou (2015) further included a person’s eye/body vectors and body movement in visual semiotics, since eye gazing, body posture, and body movement have certain implicit meanings in a given context and can be visually perceived by the individuals at that setting.
The third semiotic system in geosemiotic aggregate is *place semiotics*, which is used to connect spatial organization with language use at given places (Lou, 2015; Scollon & Scollon, 2003). Places are classified as either public/frontstage spaces or private/backstage spaces. Public spaces are further classified as exhibit/display, passage, special use, or secure, based on the functions of the space. Similarly, discourses can be further classified as contextual, decontextualized, or transgressive by examining the match between the places and the discourses.

In addition, Scollon and Scollon (2013) claimed that the discourses are highly relevant to the *social actions* of the people at a given place. Social actions are behaviors of people who play different social roles at a place. For instance, again, a beggar in a street might ask the people passing by to give him money or food as charity. The discourse of asking and the actions of begging are then socially connected. Similar to discourses, social actions are also influenced by the three semiotic systems—interactional order, visual semiotics, and place semiotics—from the perspective of geosemiotics.

In this study, the discourses as well as the social actions between FL learners and the instructor in TPFLL activities will be examined through a lens of geosemiotics analysis for understanding the meaning-making process in a *focus on meaning* approach to instruction. The results of the study are expected to guide the design of telepresence robots and of TPFLL activities at different places. Specifically, the research questions of the study are as follows:

1. How do the three semiotic systems as well as social actions in geosemiotics shape the discourse and the meaning-making process in TPFLL activities?

2. What are the appropriate instructional strategies at different places to support the meaning-making process in the TPFLL activities?
(3) What are the principles to improve the design of the telepresence robot for facilitating the meaning-making process at different places in TPFLL activities?

**Methodology**

**Participants**

Five English native speakers were recruited as instructors from a university in the United States; three were doctoral students in the Department of Applied Linguistics and the other two were doctoral students in the College of Education. All of them had teaching experience and four of them had FL teaching experience.

Sixteen adult foreign language learners in China were recruited by sending a recruitment letter to college students at a university in China and working professionals who were recommended by friends of the researcher via email and WeChat, an online communication tool. The demographic information of the FL learners is described in Table 4.1. Of the 16 participants, 15 filled in all of the demographic information, and one male learner participated in the learning activity but did not complete the information.

*Table 4.1. Demographic Information of the Learners*

<table>
<thead>
<tr>
<th>Items</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Female: 14, Male: 2</td>
</tr>
<tr>
<td>Age</td>
<td>Range: 21 – 32; Mean: 25.9; SD: 3.2</td>
</tr>
<tr>
<td>Occupation</td>
<td>Student: 7; Employee: 8</td>
</tr>
<tr>
<td>Degree</td>
<td>Bachelor: 7; Master: 6; Doctor: 2</td>
</tr>
<tr>
<td>Years of learning</td>
<td>Range: 8 – 16; Mean: 11.2; SD: 3.0</td>
</tr>
<tr>
<td>English proficiency</td>
<td>CET4 (Intermediate Level): 5; CET6 (Advanced Level): 10</td>
</tr>
</tbody>
</table>
The study received approval from the Institutional Review Board, and all participants signed a consent form before engaging in the learning activities. Pseudonyms are used here when referring to the participants.

**Design of the TPFLL activity**

*Figure 4.1. Features and the route at the Arboretum*

The site of the study was a botanic garden located on the campus of an American university. The site was selected for three reasons. First, the garden provided the learners with a variety of FL learning opportunities relevant to the natural features and social events of the garden. The garden’s natural features include a fountain, a gate, a bamboo walk, a stone sculpture sundial, a “witness tree” (the first tree planted in the garden), a central event lawn, and a pavilion, as well as trees and flowers along the touring route (see Figure 4.1). The garden is also a social gathering space for students and community members for casual events, such as picnics, and formal events, like weddings and concerts. Second, the garden provided even and uninterrupted paved paths without the dangers of normal street traffic such as curbs,
intersections, and heavy pedestrian use; this enabled the user to easily control the telepresence robot in order to move it around the garden. Third, the garden provided a quiet and secluded setting that was more conducive to conversation and privacy than many traditional university campus spaces.

A Keebot, a mid-size telepresence robot, was used in the learning activities. FL learners in China can remotely connect to the telepresence robot via the internet by using an app on a mobile device. Nine learners used an iPhone; one used an iPad; and five used an Android phone as the mobile device. The instructors and the robot were located at the learning site, the arboretum at the university. Figure 4.2 shows a remote FL learner communicating with an instructor at the arboretum via the Keebot. The telepresence robot was connected to the Internet over either 4G or the Wi-Fi provided by the garden.

*Figure 4.2. Communication with FL learners in China via a telepresence robot*

The design of the learning activity followed the principles of a *focus on meaning* instruction approach: (1) the learning goal was to expose the learners to the naturalistic use of English at the arboretum; (2) the content of the conversation was flexible to meet the learner’s interests, background, and language skills; (3) the instructor supported the learners to understand the conversation and express their thought in English when difficulties arose; and (4) the
instructor minimized error corrections during the conversation, since language forms, such as pronunciation and grammar, were not the focus of the instruction.

In order to make the conversation close to the language that is used at the arboretum, a learning activity was designed to imitate the process of a tour guide hosting a visitor. The learner was expected to learn the language implicitly while understanding what an arboretum garden in America is like. The instructor played the role of a tour guide first to introduce garden features to the learner, as the learner controlled the robot walking around at the garden. The instructor and the learner then had further discussions based on the learner’s interests and background. Finally, the learners debriefed by discussing what they had learned at the end of the activity. The duration of each activity session was about one hour.

Data collection and analysis

Demographic information was collected via a supplemental website (http://rill360.com) before the learning activities began. Two video cameras were used to record the 16 sessions of activities; one was installed on the robot to capture video from the learner’s perspective, while the other was held by the researcher to capture the activity of both the learner (as seen on the robot’s screen) and the instructor. The researcher followed closely behind the instructor and the Keebot, but out of sight from the learner as much as possible.

After collecting the videos, the researcher and his colleagues held two data sessions to demonstrate the video clips and discuss how the video data can connect to the theories of geosemiotics. Then the researcher viewed all of the videos shot from the learner perspectives in InqScribe, a video analysis tool. Memos were written while watching the videos to build connections between the video clips and the geosemiotics framework for analysis. Specifically, the instructors’ appearances as personal front in interaction order were noted at the beginning of
each video. The social actions such as demonstrating pumpkins and carving pumpkin faces were noted in detail. Then the instructor’s and the learner’s social roles such as tour guide and pumpkin carver were identified based on the social actions. In addition, the instructors’ salient body language, gestures, and the layout of objects and the instructor in the learner view were written in the memos to make connections with visual semiotics.

After all memos had been written, the different areas at the arboretum were identified as different space types in place semiotics. The discourse patterns were then extracted at each space type based the discourse’s geosemiotic features in interaction order, visual semiotics, and social actions. The memos were then re-examined as a whole after all patterns were extracted to confirm all video clips that were relevant to the learning activities could be classified in proposed patterns.

Multimodal discourse analysis (MDA) was then applied to give a typical example of the patterns at each space. MDA theorizes and analyzes how speakers leverage language and other semiotic resources, such as gestures and body movement, to make meaning of discourse (Jewitt, 2009; O’Halloran, 2011). MDA is appropriate to be applied in the research of geosemiotics since the main factors in geosemiotics—including the frames from the learner’s view, mediated actions such as gestures and body movement, and social roles—are integrated as semiotic resources in the multimodal discourse analysis.

So far, various MDA methods have been applied in the study of geosemiotics. For instance, Wohlwend and her colleagues (2011) used tables that included screen images, mediated actions, game meanings and strategies, talk, and player identities and participation goals to demonstrate the specific discourse scenarios in the world of Webkinz. Bortoluzzi and Trevisan (2009) analyzed students’ identities in the virtual environment of Second Life from the
perspective of geosemiotics by presenting the visual frame and image, camera position, kinetic action, sound track, and meta-functional interpretation of conversational scenarios. Liao (2019) used MDA and geosemiotics to translate multimodal texts at a museum. The multimodal analysis divided the museum spaces into four ranks including the museum surroundings, the museum building, the museum exhibition, and the museum objects. Photos and the layouts of the museum were used to demonstrate how the four-rank museum spaces influenced visitors’ meaning-making and translation of the texts at the museum.

In this study, a revised MDA method used by Wohlwend and her colleagues (2011) was applied as it is most relevant to the research purpose to demonstrate the relationships among the spaces, the discourse, participants’ social roles, and social actions. Specifically, excerpts that had significance reflecting the geosemiotics features of each pattern were chosen from the learning activity videos. The key frames were chosen from each clip first. Then the content of conversation along with the mediated actions including the body language and gestures were elaborated. In addition, the social roles of the participants were identified accordingly.

Findings

After watching all videos and writing notes, the results show that the spaces at the arboretum follow the classification of spaces in place semiotics, including exhibit/display spaces, special use spaces, and passage spaces. In addition, different patterns of discourses at each space emerged according to the features of geosemiotic systems, such as interaction order, visual semiotics, and social actions. In terms of interaction order, social actors, unit, personal distance, and personal front were examined. In terms of visual semiotics, the layouts of the instructor and the objects in the learner view were examined. The overarching findings are listed in Table 2 and will be further explained in the later part of this section.
Table 4.2. Overarching findings of the geosemiotics analysis

<table>
<thead>
<tr>
<th>Space Type</th>
<th>Pattern</th>
<th>Social Actions</th>
<th>Interaction Order</th>
<th>Visual Semiotics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exhibit/display space</td>
<td>Presenting</td>
<td>Present objects</td>
<td>Social actors: guide/presenter and audience</td>
<td>Instructor outside or at one side or center of the learner view. Pointing at or holding an object.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unit: with</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal distance: intimate to social</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal front: formal or informal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Micro-role-playing</td>
<td>Role-play activities relevant to the objects</td>
<td>Social actors: relevant to the social activities</td>
<td>Instructor at one side or center of the learner view. Using gestures to action things out.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unit: with</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal distance: intimate to social</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal front: formal or informal</td>
<td></td>
</tr>
<tr>
<td>Special use space</td>
<td>Etic engagement</td>
<td>Observing social activities at the space</td>
<td>Social actors: guide and audience, roles relevant to the social activities</td>
<td>Instructor at one side of the learner view. Gazing at others’ social actions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unit: with &amp; unit relevant to the social activities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal distance: intimate to public</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal front: informal</td>
<td></td>
</tr>
<tr>
<td>Emic engagement</td>
<td>Engage in social activities at the space</td>
<td>Social actors: relevant to the social activities</td>
<td>Instructor at center of the learner view. Engaging in the social actions.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unit: with</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal distance: intimate to social</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal front: informal</td>
<td></td>
</tr>
<tr>
<td>Leading</td>
<td>Lead the learner to walk (decontextualized conversation)</td>
<td>Social actors: guide and visitor</td>
<td>Instructor at center of the learner view. Walking ahead, looking back, and waving hands.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unit: with</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal distance: intimate to public</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal front: formal or informal</td>
<td></td>
</tr>
<tr>
<td>Passage</td>
<td>Carrying</td>
<td>Carry the robot</td>
<td>Social actors: guide and visitor</td>
<td>Instructor outside or at one side or center of the learner view. Carrying and walking</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unit: with</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal distance: intimate</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal front: formal or informal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Talking while walking</td>
<td>Talk about the scenarios while walking (contextual conversation)</td>
<td>Social actors: guide and visitor</td>
<td>Instructor at center of the learner view. Gazing at robot/learner and walking backward.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unit: with</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal distance: intimate to social</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Personal front: formal or informal</td>
<td></td>
</tr>
</tbody>
</table>
Presenting and micro-role-playing at exhibit/display spaces

Exhibit/display spaces are “…spaces in which we live and act are simply to be looked at as we do other things in them or as we pass through them” (Scollon and Scollon, 2013, p. 170). In this study, the gate and the sundial can be considered exhibit/display spaces, since the space and the ornaments are used for simply looking while passing through. For instance, some pumpkins were placed on the ground as ornaments near the garden gate, because the learning activities took place within a few weeks of Halloween (see Figure 4.3). Visitors can simply look at those pumpkins while they pass through the gate; thus this area as well as its ornaments, benches, and walls constitute an exhibit/display space.

Figure 4.3. Exhibit/display space near the gate

Two patterns were identified in the exhibit/display space. The first pattern is named presenting: It would be natural for a tour guide to introduce the objects in an exhibit/display space to a visitor based on what the visitor has seen. Table 3 shows a typical presenting scenario between an instructor and an FL learner who controlled the robot near the garden of the arboretum.
Table 4.3. Presenting pumpkins at an exhibit/display space near the arboretum gate

<table>
<thead>
<tr>
<th>Learner View</th>
<th>Mediated Actions</th>
<th>Conversation</th>
<th>Social Roles</th>
</tr>
</thead>
</table>
| ![Image]     | The instructor stands outside of the learner view. | *I*: You tell me if you see anything  
*L*: I knew vegetable  
*I*: Wow, vegetables! Do you know what kind of vegetables?  
*Learner*: Em  
*I*: Have you seen these before?  
*L*: It looks like this. Oh, you can come closer.  
*I*: Do you want to move the robot forward?  
*L*: OK  
*I*: Yeah. Good. Good, good. | The instructor plays the role of a guide to orient the learner to what the learner is interested in. |
| ![Image]     | The instructor waves the right hand from the robot to the pile of the pumpkins. | *I*: Have you seen this kind before?  
*L*: Pumpkin?  
*L*: So, pumpkin. OK  
*...*  
*I*: This kind is the most popular in the United States.  
*...*  
*I*: Can you tell me about what it looks like?  
*L*: Just like orange?  
*I*: Yeah, good!  
*...*  
*I*: What does it look like?  
*L*: It’s like…  
*I*: I’m gonna show you all the different sides.  
*Lr*: Err…, star.  
*I*: Yeah! That’s what I was thinking, like a star or a flower. | The instructor plays the role of a presenter to present different pumpkins. |
| ![Image]     | The instructor crouches down and grabs a green pumpkin.  
Then the instructor puts the pumpkin back. | *I*: Can you tell me about what it looks like?  
*L*: Just like orange?  
*I*: Yeah, good!  
*...*  
*I*: What does it look like?  
*L*: It’s like…  
*I*: I’m gonna show you all the different sides.  
*Lr*: Err…, star.  
*I*: Yeah! That’s what I was thinking, like a star or a flower. | The instructor plays the role of a presenter to present different pumpkins. |
| ![Image]     | The instructor grabs an orange pumpkin. | *I*: This kind is the most popular in the United States.  
*...*  
*I*: Can you tell me about what it looks like?  
*L*: Just like orange?  
*I*: Yeah, good!  
*...*  
*I*: What does it look like?  
*L*: It’s like…  
*I*: I’m gonna show you all the different sides.  
*Lr*: Err…, star.  
*I*: Yeah! That’s what I was thinking, like a star or a flower. | The instructor plays the role of a presenter to present different pumpkins. |

Table 4.3 gives an example of conversation as well as the geosemiotics analysis in the exhibit/display place near the garden gate. The social action with this conversation is that a tour
guide, played by the instructor, introduces the pumpkins to the visitor, played by the FL learner. In place semiotics terms, the place type is exhibit/display, and the conversation between the instructor and the FL is contextual, since all content of the conversation in the clip is relevant to the objects in the place.

In terms of interaction order, the social actors—the instructor and the FL learner—are a tour guide/presenter and a visitor/audience, respectively. The unit of the people is with, a term defined by Scollon and Scollon (2013) as “two or more who are perceived as being together with each other as the main focus of their mutual attention” (p. 61). The instructor had an “intimate (18 inches or under)” (Scollon & Scollon, p. 71) interpersonal distance with the robot controlled by the FL learner. Finally, the personal front of the instructor is that she is wearing a relatively formal dark brown suit with a marker pen and a notebook in hand.

In terms of visual semiotics, Table 3 above lists the key frames in the learner view and the layouts of the frame images during the conversation. The relation between the visual semiotics and the conversation can be seen as follows.

(1) The instructor asks the FL learner to identify the objects the FL learner can see. From the FL learner’s view, various objects can be seen in the background, such as flowers in pots, pumpkins, and a bench. The learner actively selects pumpkins based on her interests and names it as a vegetable. However, when the instructor asks the learner to tell the specific name of the vegetable (“pumpkin”), the learner hesitates to respond.

(2) The instructor then, with a waving gesture, asks the learner to control the robot to move forward. The instructor puts her right hand at the center of the screen, so the learner can notice her gesture.
(3) After the learner drives the robot forward, the learner view is zoomed in. At the same time, the instructor crouches down at the left side of the learner view and picks up a green pumpkin at the center bottom of the learner view, so the learner can focus on the pumpkins. At this point, the learner identifies it successfully as a pumpkin.

(4) The instructor expands the topic to the appearance of the pumpkins. She first grabs an orange pumpkin and puts it at the center bottom of the learner view. Then the instructor asks the learner to describe it. After the learner says that it is like an orange, the instructor gives positive feedback to the learner.

(5) The instructor expands the topic to compare the different pumpkins by picking up another shape of pumpkins and again putting it at the center bottom of the screen. She asks the learner questions and gives the learner feedback again.

In general, this example of presenting shows how the three semiotic systems and social actions interplay to shape the discourse of presenting the appearance of different pumpkins at an exhibit/display space.

Another pattern that can be found in exhibit/display spaces is called micro-role-playing, in which instructors tend to expand the discourse from the appearance of the objects to the social and cultural functions or activities of the objects. Table 4.4 below shows an example in which an instructor is describing the cultural meaning of pumpkins in western countries by creating the action of carving pumpkin faces.
Table 4.4. Micro-role-playing for expanding discourse to sociocultural aspects of an object

<table>
<thead>
<tr>
<th>Learner View</th>
<th>Mediated Actions</th>
<th>Conversation</th>
<th>Social Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td>The right hand of the instructor points to the pumpkins. The learner moves up the camera. The instructor then grabs the coffee cup from the ground.</td>
<td><strong>I:</strong> You said this plant is for Halloween. What do we do with this plant in Halloween? <strong>L:</strong> Um, use it, like… um, huh huh huh. I can’t explain.</td>
<td>The instructor plays the role of a guide to orient the learner to the pumpkins.</td>
</tr>
<tr>
<td><img src="image2.png" alt="Image" /></td>
<td>The right hand of the instructor makes a gesture of cutting and then points to the learner and opens the palm.</td>
<td><strong>I:</strong> So, people they cut them, right? <strong>L:</strong> Cut them? Oh, huh huh <strong>I:</strong> And they may…what do they … <strong>L:</strong> Make faces. <strong>I:</strong> Exactly. They make faces. <strong>L:</strong> Ok, um. <strong>I:</strong> So, they use the pumpkin, and they cut it, and they empty it, they take everything outside. And they cut a face. <strong>L:</strong> Oh, Um. <strong>I:</strong> So, it looks like a person. <strong>L:</strong> I know, I know.</td>
<td>The instructor plays the role of a pumpkin face maker.</td>
</tr>
<tr>
<td><img src="image3.png" alt="Image" /></td>
<td>The right hand of the instructor makes gestures of cutting and emptying, then makes gestures of cutting again.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the perspective of geosemiotics, the instructor’s behavior in demonstrating the process of carving a pumpkin face actually changes his social role from a tour guide to a pumpkin face maker. In addition, the space of the social action has been switched from the physical arboretum gate to an imaginary place of making a pumpkin face, such as home. We called this micro-role-playing because the role-play happens and finishes very quickly, i.e., the instructor turns his role from the pumpkin face maker back to a tour guide once he had finished the conversation about pumpkin face making.
The instructor also positions himself at a further distance compared to the example in Table 3, so that the learner can observe his body movement and gestures in a wider view. The learner, accordingly, moves the camera on the robot upward slightly during the conversation, so that the learner can see instructors better while the pumpkins are still in the learner’s view.

Etic and emic engagement at special use spaces

Special use spaces—such as shopping malls and restaurants—have a purpose for public use (Scollon and Scollon, 2013). In this study, the event lawns and pavilion can be considered special use spaces since they are designated for particular purposes, such as hosting a ceremony or concert. We observed that there are two patterns in the discourse at a special use space: One is called etic engagement and the other is emic engagement. In etic engagement, the instructor and the learners position themselves as outsiders to observe the social actions or objects at a special use space. In emic engagement, the instructor and the learner play the social roles and engage in the social actions themselves in the special use space.

Table 4.5 demonstrates a typical example of etic engagement on the event lawn in the arboretum garden. The social action here is watching two other visitors playing baseball at the center of the lawn. Here there are three types of social actors: the instructor as a tour guide, the learner as a visitor, and the other two visitors as baseball players. The units of the people are two with: The tour guide is with the visitor, and the two baseball players are with each other. The instructor and the two visitors wear informal shirts. The instructor mainly keeps a personal distance (close phrase, or 18 inches to 30 inches) with the learner/robot. The instructor and the learner/robot have a public distance (far phase, or 4 feet to 12 feet) with the two visitors.
Table 4.5. Etic engagement with baseball game at a special use space

<table>
<thead>
<tr>
<th>Learner View</th>
<th>Mediated Actions</th>
<th>Conversation</th>
<th>Social Roles</th>
</tr>
</thead>
</table>
| ![Learner View](image1) | The instructor gazes at the learner while talking. | *I:* You can’t put the robot on the grass. Why don’t you tell me what you see and what is going on?  
*L:* OK. I see… | The instructor plays the role of a tour guide as the learner plays a visitor. Meanwhile, the other two visitors on the lawn play the role of baseball players. |
| ![Learner View](image2) | The instructor turns to the left and watches the two visitors playing baseball. | *L:* Oh, I see a man. Play with his dog. He throws something, a ball or something.  
*I:* Yeah, they are playing. They are playing baseball.  
*L:* Oh. Baseball! Oh… oh, with his friend?  
*Instructor:* Yeah  
*Learner:* Oh… There are two people, I just see (saw) one.  
*Instructor:* Have you ever played baseball? | |
| ![Learner View](image3) | The instructor walks closer to the robot and turns it slightly to the right. |  |  |
| ![Learner View](image4) | The learner controls the robot to turn further right. |  |  |

During the conversation, the instructor positions herself at the left side of the learner’s view, then asks the learner to describe what the learner sees. Since the robot cannot walk on the grass, the instructor just asks the learner to observe at the side of the lawn. The learner answers that she sees a man with a dog playing with a ball. The instructor notices that the learner doesn’t find another visitor because the visitor is out of the learner’s view, so the instructor turns the robot a little to the right and tells the learner they are playing baseball. The learner then notices another visitor and confirms this with the instructor. After getting positive feedback from the instructor, the learner turns the robot further right to get a better view of both visitors at the same time. Then the learner and instructor talk more about baseball.
By contrast, Table 4.6 below demonstrates how an instructor and a learner participate in the social actions at the event lawn in an emic way. The instructor first asks the learner a question about a picnic on the grass field. However, the learner doesn’t get the point, so the instructor role-plays a picnic by taking off her coat and putting it on the grass. Then the instructor sits on her coat and start to talk more about a picnic in America, to help the learner to understand the meaning of word “picnic.”

Table 4.6. Emic engagement into a picnic at a special use space

<table>
<thead>
<tr>
<th>Learner View</th>
<th>Mediated Actions</th>
<th>Conversation</th>
<th>Social Roles</th>
</tr>
</thead>
</table>
| ![Image](image1.png) | The instructor uses two hands to imitate holding a blanket. | *I:* Would you bring a blanket to put on the grass?  
*L:* Grass?  
*I:* Would you…  
*L:* Would you need a blanket for picnic?  
*L:* Blanket? | The instructor plays the role of a guide to orient the learner to a picnic on the lawn. |
| ![Image](image2.png) | The instructor takes off the coat and puts it on the ground. |  
*L:* Are you cold?  
*I:* Would you put a blanket on a grass so you can sit?  
*L:* Yes, I can… I can sit  
*I:* You would need a blanket, you would bring food.  
*L:* Yes, bring food. | The instructor plays the role of a friend of the learner. |
| ![Image](image3.png) | The instructor sits on her coat and keeps talking. |  |  |

In this example, the instructor changes her role from a tour guide to a friend of the learner as they partake in a picnic together. Accordingly, the learner becomes another picnic participant in the discourse. The instructor sits and talks to the learner, which imitates a scenario in which two friends or family members have a picnic on the event lawn. The instructor manages the interpersonal distance from intimate (18 inches and under 18 inches) to personal (far phase, or 30
inches to 4 feet) to make the role play more real and authentic. Also, the instructor’s informal clothing enhances the feeling of being friends in a picnic on a lawn.

*Three instructional strategies at passage spaces*

Passage spaces are spaces that allow people to travel through an area; examples include roads, pathways, stairs, and escalators. Scollon and Scollon (2013) pointed out that passage spaces can “foster a kind of slow, ‘civilized’ passage on foot” and allow people to “walk together with others in a common social space without having to keep a constant eye out for automobiles” (p. 170). Using this definition, the pathway at the garden in this study can be considered a passage space.

Table 4.7 shows one of the typical strategies, *leading*, at a passage space from the witness tree of the garden to the pavilion (see Figure 1). The instructor guides the learner to control the robot walking towards the destination; as a result, the content of the conversation is decontextualized, i.e., is not related directly to the scenarios in the environment, but instead is relevant only to the action of moving the robot. This decontextualized content can thus be applied to any scenario. The instructor uses his body movement, gestures, and phrases such as “keep going” to show the correct direction to the learner. The instructor maintains the distance from personal (close phase, 18 inches to 30 inches) to public (12 feet to 25 feet) and looks back at the robot from time to time. It takes a relatively long time to move from one spot to the next, because the robot can’t walk very fast.
Table 4.7. Leading the learner at the passage space between the witness tree and the pavilion

<table>
<thead>
<tr>
<th>Learner View</th>
<th>Mediated Actions</th>
<th>Conversation</th>
<th>Social Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Image" /></td>
<td>The instructor waves hand to the front. The instructor turns around and walks. The learner controls the robot to follow the instructor.</td>
<td><em>Instructor:</em> Ok. So, we’re almost done, but there is one more important thing for us to see. <em>Learner:</em> Ok. Em.</td>
<td></td>
</tr>
<tr>
<td><img src="image2" alt="Image" /></td>
<td>The instructor walks and looks back from time to time or waves in the direction of the destination. The learner follows without talking.</td>
<td><em>I:</em> Keep going. <em>L:</em> Want close? Emm. <em>I:</em> We have a long walk for this section. <em>L:</em> Oh. Huh heh. <em>I:</em> It’s actually not very far. Just take some time in the robot. <em>L:</em> Just go ahead and…</td>
<td>The instructor plays the role of a guide as the learner plays a visitor.</td>
</tr>
<tr>
<td><img src="image3" alt="Image" /></td>
<td>The instructor walks ahead and looks back several times until finally standing at the end of the passage. The learner controls the robot to follow.</td>
<td><em>L:</em> It’s a long way to go. Huh huh.</td>
<td></td>
</tr>
<tr>
<td><img src="image4" alt="Image" /></td>
<td>The learner finally reaches the destination.</td>
<td><em>Instructor:</em> It’s the last thing we have to see. Don’t worry.</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.4 shows the second strategy, *carrying*, which is being used by the instructor in the passage space. Instead of waiting for the learner to control the robot, the instructor carries the robot to pass through the passage space. The instructor and the learner can still talk with each other. However, the learner may lose his or her ability to look around the environment by themselves.
The third strategy we observed in the activities is *talking while walking*. In Table 4.8, the instructor asks some questions relevant to the next garden feature while walking backward. The learner answers the question while following the instructor. Although it takes the same time as the first strategy, *leading*, the instructor and the learner can have more conversation about the scenarios, so the content of the discourse is contextual. However, the sound of the robot moving is noticeable during the walking, so it might influence learner’s hearing of the conversation.

*Table 4.8. Talking while walking from the witness tree to the pavilion*

<table>
<thead>
<tr>
<th>Learner View</th>
<th>Mediated Actions</th>
<th>Conversation</th>
<th>Social Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Learner View" /></td>
<td>The instructor bows down somewhat.</td>
<td><em>I:</em> Did you listen to a conversation before?</td>
<td><img src="image" alt="Social Roles" /> The instructor plays the role of a guide as the learner plays a visitor.</td>
</tr>
<tr>
<td><img src="image" alt="Learner View" /></td>
<td>The instructor shows two fingers while walking backward. The learner controls the robot to follow the instructor, with salient moving sound.</td>
<td><em>I:</em> between two peoples? <em>L:</em> Oh, yeah. <em>I:</em> What did they talk about? <em>L:</em> Em, they are, they are planning their wedding.</td>
<td></td>
</tr>
</tbody>
</table>
Switch among different space types

In the study, we observed that the space type at an area may change according to the behavior of the instructor and the learner. An exhibit/display space may become a special use space. For example, a sundial is located near the event lawn (see Figure 5, below). If an instructor treats the sundial as an ornament of the garden and introduces the appearance and function of the sundial, then the space around the sundial is more like an exhibit/display space. However, if an instructor reads the time from the sundial (see the right-hand picture in Figure 4.5), the sundial has an actual purpose for public use, i.e., getting the current time. This space then becomes a special use space.

Figure 4.5. Switching from exhibit/display space to special use space at the sundial

Another instance of changing the space type is to switch a passage space to an exhibit/display space or special use space according to the instructor’s behavior. For instance, when the instructor in Figure 4.6 below carries the robot from the witness tree to the pavilion, the learner sees a birch tree along the way with special white and grey colors. The instructor puts the robot on the ground and starts to introduce the birch tree; in this way, the passage space here turns to an exhibit/display space.
Although the space type may change, the patterns in the findings are still meaningful for understanding the discourse at the changed space type. For example, after the instructor changes the passage space in Figure 6 above to an exhibit/display space, the conversation, social actions, and social roles follow the patterns of either presenting or micro-role-playing.

Discussion

The existing studies in geosemiotics have demonstrated how the framework of semiotic systems or resources in geosemiotics can be applied in various settings. For example, Hamid (2016) analyzed how people make sense of traffic signs in urban cities to inform the design of traffic policies and manuals from both the theoretical lens of geosemiotics and mobility. Lou (2015) examined how the interaction order, visual semiotics, and place semiotics varied in three Hong Kong markets and how these differences influenced customers’ sense of the places and their preferences of communication modes in those markets. Pierce (2012) investigated how spaces and materials influence the interactions in ESL classrooms from the perspective of geosemiotics including interaction order, visual semiotics in materials and participants’ interaction photographs, classroom design as place semiotics, and the recurring themes in research data.
However, few studies have examined the influence of places in language acquisition from a perspective of geosemiotics. This study bridges the gap between sense-making in language acquisition and geosemiotics by pulling out the interactional patterns at each type of space and by demonstrating the specific scenarios with MDA. In addition, this study also provided the insights on how these findings can guide instructional design and educational robot design in TPFLL learning activities.

*Understand meaning-making in TPFLL activities*

Because the TPFLL activities in the study used a *focus on meaning* instructional approach (Krashen & Terrell, 1983; Savignon, 2018), the focus of the learning activities helped FL learners make sense of the meaning of the discourse rather than the language forms, such as grammar and pronunciation. The learners are expected to acquire the FL while making sense of the discourse. Based on the findings of this study, learners will comprehend the FL better if they focus on understanding the meanings of the following terms:

1. The place (e.g., what does an arboretum mean?)
2. The objects in a place (e.g., what do pumpkins mean?)
   a. The appearance of the object (e.g., size, color, shape)
   b. The function of the object (e.g., as food or ornament)
   c. The cultural activities relevant to the object (e.g., carving pumpkin face)
3. The social events at a place (e.g., playing baseball or having a picnic on a lawn)
   a. The social actors in the social event (e.g., baseball players)
   b. The behavior of the social actors (e.g., playing baseball)

From the perspective of geosemiotics, three semiotic systems—interaction order, visual semiotics, and place semiotics—will shape the discourse and the meaning of the discourse. In
terms of interaction order, this study consolidates the findings of Wohlwend’s (2011) study that kids or learners plays various social roles such as shopper in a virtual store while they are playing in the online world of Webkinz. In this study, the instructor and the learner were playing different social roles other than teacher and student such as tour guide, pumpkin-carver, or friends in a picnic. The content and the communication modes of the conversation will be changed by the social roles of the speakers accordingly. If the social actors played by the instructor and the learner are a garden administrator and a worker, the topic of the conversation may change to the position of, and the correct way to pile, the pumpkins. If the people unit is *files and processions*, i.e., a tour guide leading a group of visitors (Scollon & Scollon, 2013, p. 61), the instructor may need more strategies to interact with the multiple learners who play the role of visitors. Moreover, the *interpersonal distance* and *personal front* will influence how the learner perceives the social role of the instructor. If an instructor wears more formal attire in a far social distance (7 feet - 12 feet) with the learner, the instructor is likely to be perceived as a teacher at a school, while an instructor dressed more informally and having a closer distance with the learner may seem more like a social actor, such as a tour guide or a friend of the learner.

Visual semiotics also influence how the learner makes sense of the discourse in a particular space. This study focused on the learner-view layout since camera positioning will influence the learner’s perception of the space (Bortoluzzi & Trevisan, 2009). Three patterns of learner-view layout can be found: (1) the instructor is out of view; (2) the instructor is at the left or right side while the object is at the other side or center of the view; and (3) the instructor is at the center of the learner’s view, with or without an object. The first layout usually implies that the learner needs to find an object that interests him or her. The second layout implies that the learner needs to figure out the meaning of an instructor’s discussion by observing the object at
the other side or at the center. The last layout usually means that the learner should focus only on
the instructor’s conversation or the object that the instructor is holding. This study also adopted
Lou’s (2015) method that included body language and gestures in visual semiotics to
demonstrate how learners can visually perceive the instructor’s social actions at a space.

In terms of place semiotics, this study used the classification of space types proposed by
Scollon and Scollon (2013) that shows different patterns of the meaning-making process at
different space types. The meaning-making of an object at an exhibit/display space usually
follows this procedure: a learner observes an object at the space, with or without the instructor’s
guiding; then the instructor asks the learner to identify the object with the gestures of pointing at
or holding the object; and finally, the instructor expands the topic to the appearance, functions,
and cultural traditions of the object with the gestures of acting things or activities out. The
meaning-making in a special use space focuses more on the social events or social actions, with
the instructor and the learner participating in the social events in either an etic or emic way based
on their social roles in the social actions. Accordingly, the instructor will either gaze at others’
social actions or engage in the social actions directly. Moreover, three strategies—leading,
carrying, and talking while walking—were observed at the passage space. Leading is the simple
way to guide a learner to control the robot to go from one spot to the next usually with the body
language of walking ahead, looking back, and waving hands. The conversation usually focuses
only on the moving of the robot, so the content of the conversation is decontextualized. Carrying
the robot while walking can save time in traversing the passage space, and the instructor can
have a contextual conversation with the learner, but this strategy doesn’t give the learner the
ability to observe the environment himself or herself. Talking while walking, the third strategy,
allows the instructor to have a contextual conversation with the learner while walking backwards
through the passage space. However, the sound of the robot moving is obvious and might influence the learner’s ability to hear and understand.

It also is important to note that the space type at a given place might change according to the behaviors of the instructor and the learner. For example, an exhibit/display space can be changed to a special use space if an instructor wants to use the space for a special purpose rather than simply observing it. The reverse is also true: A special use space may become an exhibit/display if the instructor and the learner only observe the appearance of the space rather than engaging in social actions in that space. Similarly, a passage space could turn to an exhibit/display space or a special use space, if the instructor and learner stop walking and start to talk about the objects or engage in social activities there.

In brief, the findings of this study provide additional evidence that sense-making is a complex and ongoing process that is socially and culturally related to humans’ social actions at a given place (De Jaegher & Di Paolo, 2007; Weick, 1995).

*Instructional design from the perspective of Geosemiotics*

Alignment with the viewpoint that the goal of meaning-focused FL learning is to expose the FL learner to, and engage the learner in, the naturalistic language use (Oura, 2001; Ozverir et al., 2017), the language used in the TPFL activities should be as authentic as possible. From the perspective of geosemiotics, the instructor can design the TPFL activities in accordance with the three semiotics systems as well as the social actions to make the language used in the activities closer to the real-use language.

Prior to the learning activities, an instructor can prepare the conversation based on the features of the learning site. Although the conversation content could be flexible, the preparation can help the instructor in facing different situations. Specifically, the route of the learning
activities should be planned ahead, considering the route distance and the number of talking spots. Then each talking spot can be classified as either exhibit/display space or special use space. A pathway in between two talking spots can be considered passage space.

For exhibit/display space, the instructor needs to observe what objects can be used for the conversation, taking into consideration the objects’ appearance, functions, and cultural meanings. For special use space, the instructor can list the possible social activities and consider whether to use either an etic or emic engagement during the conversation. If it is emic engagement, in which the instructor and the learner participate in the social actions at a special use space directly, it would be ideal to prepare relevant equipment or tools for role playing, for instance, preparing a real blanket for demonstrating a picnic on a lawn. Finally, for passage spaces, it is necessary to plan in advance the strategies that will be adopted. For example, if the route in the passage space is short and there are no interesting objects around, then it might be appropriate to use the leading strategy. If the passage space is short but there are interesting objects, such as a special tree, then talking while walking can be used. If the passage is too long, it will be better to carry the robot in order to save time or add an exhibit/display space partway through the long passage space.

During TPFL activities, an instructor needs to manage his or her distance from the learner and the position in the learner’s view. After guiding the learner to an exhibit/display space, the instructor can stand outside of the learner view and ask the learner to observe the environment first to find an object that the learner is interested in. If a learner can’t find one, then the instructor can stand near the object, such as pumpkins, and use pointing or other gestures to guide the learner’s focus to the object. For special use space, the instructor can stand aside the social activities at the place, such as watching two visitors playing baseball on an event lawn, if
the instructor wants to use an etic way to introduce the social actions taking place. By contrast, the instructor can position himself or herself at the center of the screen in order to use an emic way to engage the learner in the social activities at the place. Finally, for passage space, the instructor can use his or her position in the learner’s view to guide the learner to move in either the leading or talking-while-walking strategy.

Telepresence robot design from the perspective of geosemiotics

To date, few studies have investigated the design of telepresence robots for FL learning use. From the perspective of geosemiotics, the use of a telepresence robot can be examined at each of the three space types. In exhibit/display spaces, the design of current robot basically meets the needs of orienting learners to the focused objects or social activities prior to presenting or micro-role-play. The learner can adjust the view angle to obtain a better view by turning the robot around or tilting the camera on the robot head up and down (see the learner view in Table 4.4 and Table 4.5).

However, for special use spaces, one limitation is the mobility of the robot. Although the robot can move on a flat concrete ground, it cannot walk on grass. This is a particular drawback when an instructor wants to use an emic way to engage the learner in the social actions at the space, such as a picnic, because the robot cannot move to the grass to give the learner a better experience. A second limitation is that the current telepresence robot, KeeBot, lacks mechanic arms, so it is not possible to engage in social actions that require the use of hands, such as playing baseball on a lawn.

For passage space, the speed of the current telepresence robot, KeeBot, is still slow, so some instructors choose to carry the robot through the passage, which might hinder the learner’s ability to observe the environment himself or herself. Although talking and walking is a good
strategy for instruction in these instances, the sound of robot moving might be too noticeable and might hinder the learner’s ability to hear the instructor clearly. In general, future telepresence robots need to be faster and quieter, with better mobility on different terrains, and with mechanical arms for the purpose of FL instruction.

Conclusion

This paper examines the meaning-making process in telepresence-place-based foreign language learning through a theoretical lens of geosemiotics. The results show that conversations within TPFLL activities are shaped by three geosemiotic systems—interaction order, visual semiotics, and place semiotics—as well as social actions. The findings suggest that TPFLL activities should be designed based on the different space types. In addition, the geosemiotic analysis of TPFLL shows that telepresence robots should be improved to meet instructional needs in different settings.

It is not clear yet whether the findings of this study are applicable to TPFLL activities at learning sites other than an arboretum. Moreover, the learning outcome of TPFLL, i.e., the degree to which the TPFLL activities can improve learners’ communicative use of the FL, has not yet been measured. We plan to conduct more studies at different learning sites, such as a library, a parking lot, or a restaurant, to examine the instructional strategies and learning outcomes of TPFLL further.

References


Coulthard (Eds), *Identity Construction and Positioning in Discourse and Society* (pp. 199-244). Tilgher-Genova s.a.s.


CHAPTER 5: DISCUSSION AND CONCLUSION

This chapter discusses the findings of the three articles and makes conclusions in six sections. The first section summarizes the findings of the three articles. The second section provides the insights on relevant theoretical topics such as place, authenticity, etc. The third section discusses the implications of the study in place-based learning, communicative language teaching, and telepresence robot design for the purpose of foreign language learning. The fourth section proposes more concerns about the practical application of telepresence robots in education. The fifth section describes the boundaries of the study and lists research questions and topics for future studies. Finally, the last section makes conclusions about this study.

Summary of the Three Articles

This dissertation explores how telepresence robots can be used in foreign language learning in real life contexts, especially to improve FL learners’ communicative competence. The three articles gradually narrow down the topics from generally examining the perceived benefits and challenges of FL learners, to general design principles of the learning activity design, and then to specific design concerns on different types of places through a theoretical lens of geosemiotics.

The first article (Liao & Lu, 2018) is a pilot study in which a small telepresence robot, Romo, was used to support three English learners in China as they communicated with a native English speaker (an instructor) at one of the main buildings of an American university campus. The goal of the learning activities was to provide authentic communicative practice for the FL learners. Based on the analysis of interview transcripts, fieldnotes, and the transcripts of video clips, the results of this study show that the learners felt the learning environment while using a telepresence robot was similar to a virtual reality game, but it was more authentic than virtual
reality. FL learners were motivated especially when they found that there are connections between their own campus life and the American campus life they experienced through the robots. It was easy for the instructor and FL learners to start spontaneous conversations since the objects and people around the campus buildings naturally provided rich conversational topics. Moreover, the learning activities were learner-centered since the FL learners had the ability to control what they wanted to observe via the telepresence robot. However, FL learners with lower English proficiency felt pressure when asked to have spontaneous conversation.

Based on the findings of the first article, three design principles were proposed and examined in the second article (Liao et al., 2019) including situating the learning in real-life settings, scaffolding the learning process, and enhancing learner agency. The second article defines the learning form via telepresence robot in real-life settings as telepresence-place-based foreign language learning. Eleven English learners in China were recruited as FL learners, and four native English speakers were recruited as instructors. A one-hour learning activity was conducted for each FL learner at an arboretum on an American university campus via Keebot, a larger telepresence robot than Romo. All learning activities were videotaped with two cameras, one held by the researcher and another fixed on the robot’s head to capture the learner view. The researcher made field notes during the activities, and all participants were interviewed after the activities. The results show that the participants had positive attitudes towards the learning activities, which were designed based on the three design principles.

The last article proposed more specific design concerns in different spaces through a lens of geosemiotics. Geosemiotics considers that discourses at a place are shaped or indexed by people’s social actions and three semiotic systems including: interaction order, visual semiotics, and place semiotics. Interaction order refers to the social relationships of people at a place who
play different social roles and behave differently regarding their social actions. Visual semiotics is the way that the discourse and materials are visually presented. For example, the layout of the objects and the instructor in the learner view will influence the conversation between the FL learners and the instructor. Finally, place semiotics connects spatial organization with language use at a given place. The third article classified the spaces at the arboretum into three types: exhibit/display space, special use space, and passage space, according to Scollon and Scollon’s (2003) classification in place semiotics. Then seven patterns including presenting, micro-role-playing, etic engagement, emic engagement, leading, carrying, and talking while walking, were extracted based on the features of geosemiotic systems and social actions at each space. Accordingly, the suggestions on instructions and robot design at each type of space were given based on the findings of the patterns.

In brief, the three articles in this dissertation explore the approaches to apply telepresence robots in FL learning, and the results show that telepresence robots have the potential to support FL learning that has not been realized in other studies.

**Discussions in Theoretical Topics**

*Place and language authenticity*

The understanding of the concept of ‘place’ has grown through the journey of the three studies in this research. Generally, place refers to a physical environment or a locality used for a special purpose (Place, 2019). From this perspective, place could be anywhere that has either educational or non-educational purposes for human beings. In the context of education, the meaning of place more specifically refers to a physical environment where teaching and learning occur, which has a similar meaning to learning environment, learning settings, or learning space (Cook, 2010).
When it comes to place-based learning or place-based education (Sobel, 2005), the concept of place is more related to the locations outside the classroom or to an outdoor space. However, outdoor places may exclude indoor places outside classrooms, such as the space inside a library or a stadium. Then, the terms such as learning in real life settings (e.g., Amaral & Meurers, 2011), in everyday life (Kramsch, 1995), or in the wild (Wagner, 2015) become closer to the places that are outside the classroom and have the potential to support FL learning. The first and the second articles adopted this idea that telepresence robots can support the FL learning occurring at the places outside the classroom.

The understand of place evolved in the third article with the theoretical lens of geosemiotics. From the perspective of geosemiotics (Scollon & Scollon, 2013), places are the spaces that involve people who are conducting social actions, playing social roles, having conversation, and making artifacts. Accordingly, places are highly relevant to human being’s social activities, cultures, and languages. A discourse cannot be fully explained or understood without considering its semiotic resources at a place including the people’s social roles, social actions, the physical and social features of the place, and how we visually perceive the scenarios at a place.

Therefore, places have strong connections with language and its authenticity. Authenticity has been used to refer to the resemblance of the language in FL classroom teaching to the language used in real situations (Gulikers et al., 2004; Tisdell, 2003). As Chapter 1 mentioned, the language used in the classroom should be as authentic as possible to improve the FL learners’ skills to use the target language in real situations (e.g., Brown, 2000).

However, the components of authenticity still remain vague in previous studies, and different scholars proposed different point of view. For example, Strobel et al. (2013) defined
four categories of authenticity: context authenticity (the language context is real), task authenticity (the tasks are genuine), impact authenticity (language activities have real impact for the learners’ lives or careers), and personal/value authenticity (the language activities convey the values that the learners agree upon). Buendgens-Kosten (2013) divided authenticity into three domains including cultural, functional, and linguistic authenticity. Falconer (2013) identified nine positive factors that are relevant to the authenticity dimension of virtual worlds including facilitation, presence and authority, visual realism, socialization, comparative reality, engagement, active learning, generalizability, and enabling learning from mistakes.

The findings of this research provide a new insight on authenticity from the perspective of place, i.e., place authenticity could be a new component of authenticity for language teaching. In particular, the third article has demonstrated how discourse is related to a place as well as the people with social roles, social actions, and the features of spaces. The acquired language will then be more authentic when social roles include acting appropriately at a particular place. Although place authenticity has some overlap with context authenticity, task authenticity (Strobel et al., 2013), and cultural authenticity (Buendgens-Kosten, 2013) since these authenticities also involve social activities, place authenticity emphasizes the physical place itself.

In traditional FL teaching, the place of learning is limited to the classroom in schools or a local place that a learner has access to. However, it is now possible to have learners to be exposed to and engaged in the real language use at a remote place in the target-language countries with the support of telepresence technologies. Accordingly, the degree of place authenticity in learning activities grows when the learning locations as well as the social roles and social actions are closer to real ones. From this perspective, the following learner situations
are sorted based on the degree of place authenticity from low to high: learning in a classroom with learning materials of situational dialogues in text; learning in a classroom but with roleplay activities in which learners play the roles of a certain social action at an imaginary space, such as imitating ordering food at a restaurant; learning in a virtual world that involves conversational scenarios in the FL; learning at a place in the learner’s local community; learning at the target-language community through telepresence technologies, and finally, learning at a place in the target-language community physically.

In conclusion, place is a crucial factor that is relevant to language authenticity. Just as Kramsch and Sullivan (1996) mentioned, what is authentic in London may not be authentic in Hanoi. The reason why the authentic language in London may not be authentic in Hanoi is not only because of the difference of the physical places, but also because the people, objects, social activities, and cultures are different. Telepresence robots, then, play an important role to bring the authentic place along with the authentic language to the FL learners.

*Scaffolding learners for improving communicative competence*

Communicative competence has been defined as the ability to apply language in real and spontaneous conversational situations in given social context (Hymes, 1972; Savignon, 2018). Chapter 1 of this dissertation further introduced two different teaching approaches to improve communicative competence, the Natural Approach (NA) (Terrell, 1977) and Communicative Language Teaching (CLT) (Richards, 2006). Both the NA and CLT hold similar rationales that the acquisition of a FL can only occur while the FL learners were exposed to authentic language use in real life contexts and were encouraged to produce meaningful communication. Based on the findings in the first and the second articles, FL learners in the study do enjoy the learning activities supported by telepresence robots to engage them in authentic and meaningful
communication at a main building and an arboretum of an American university campus. The survey results also show that the FL learners perceived that TPFLL could improve their communication skills including listening, speaking, and their cultural awareness.

However, the findings of the first article also showed that FL learners with low English proficiency had difficulty understanding authentic conversations and expressing their thoughts in an authentic way. Then, the second article proposed scaffolding learners as a design principle for the learning activities. Two key components of scaffolding were summarized including contingency (responsive and tailed support) and fading (learners are expected to gradually apply the learned skills on their own). Various specific scaffolding techniques were identified in the learning activity videos such as using pauses to check learners’ understanding and difficulties in the conversation, rephrasing the sentences and words to simple ones, slowing down speech, and acting a word out to help learners understand. Although the third article doesn’t cover the topic of scaffolding, the multimodal discourse analysis still shows how the instructors use gestures and equipment such as cloth to imitate the blanket in a picnic for helping the learner understand the meaning of the concept ‘picnic’.

These findings show the differences of the pedagogical strategies used in current TPFLL studies from those in the Natural Approach (NA) and previous Communicative Language Teaching (CLT) methods. As introduced in Chapter 1, the NA usually starts with plenty of listening comprehension of dialogues in classroom; then it moves to the stage of responding in simplified words or even partly in L1; finally, learners are expected to respond to more complex sentences (Terrell, 1997). The existing TPFLL studies haven’t provided plenty of listening materials before the learning activities. The learners were only provided a short audio including conversation about the social activities at the pavilion in the arboretum, and the learners reported
positive attitudes towards the listening materials. However, it is still not clear how many the listening materials should be given to the learners before activities. From the perspective of scaffolding, each learner has a different zone of proximal development (Kim & Hannafin, 2011; Vygotsky, 1978), i.e., the boundaries between learners’ current skillset and desired skillset, so it would be better to provide the learning materials adaptively based on learners’ needs. More research on this topic is still needed to disclose details about the design of the listening materials before TPFLL activities.

The second difference between the NA and TPFLL is that FL learners barely have the chance to use their mother tongue/L1 to support their speaking since the instructors in the TPFLL activities are native English speakers and had insignificant knowledge about Chinese. In other words, English speakers in authentic conversational scenarios, in most cases, may not know the learners’ first language. Thus, the learner and the instructor have to use the learner’s FL to communicate with each other. This is challenging for the instructor in TPFLL activities — before an instructor starts to scaffold the learner to express what the learner hopes to say, the instructor needs to know what the learner wants to express. In this case, the learner’s body language, simplified description, and/or a whiteboard would be helpful to scaffold the instructor, instead, to understand the learner first.

In terms of the relation between CLT and TPFLL, TPFLL could be considered as one kind of CLT approach, since CLT is a broad approach that can include various specific teaching methods that focus on improving learners’ communicative competence (Richards, 2006). However, most CLT studies were situated in classroom settings. The role of teacher is that of a facilitator, who creates a classroom climate to encourage learners to use and practice the foreign language with other learners (Richards, 2006). Therefore, few studies in CLT have investigated
how a teacher can scaffold FL learners’ learning directly. In this research, the role of instructors was neither a facilitator nor a teacher in traditional form schools. The third article in particular analyzed the social roles of the instructor switching among a tour guide, a pumpkin carver, a friend of the learner, etc. Thus, the general role of the instructors was closer to a private tour guide. However, it is possible to expand this role to be a facilitator for future studies if the learning activities involve multiple learners and multiple telepresence robots in a session.

In summary, although scaffolding has been a hot research topic for decades, the application of scaffolding in the teaching approaches to improve communicative competence haven’t been investigated thoroughly. The findings of this research reveal that learners in the TPFLL learning activities need to be scaffolded to understanding the listening process and to generate authentic sentences to express their thoughts. More studies need to be conducted to explore what other scaffolding techniques or relevant pedagogical strategies could be used in TPFLL activities.

**Mobility, social presence, and learner agency**

Mobility is a concept that involves not only the movement from A to B, but also the human activities that take place during the movement (Laurier et al., 2008). Scholars argue that the paradigm of mobilities should integrate place, movement, meaning, culture, people’s embodied practice, and personal experience (e.g., Cresswell, 2006; Ernste et al., 2012). Vilhelmson and Thulin (2008) further proposed that the method of mobilities had been changing from the real world to the virtual world with the development of information and communication technology (ICT). Even from home, individuals can still be virtually mobile and interact with others at a distance. The virtual mobility offered by ICT is then blurring the boundaries between in-home and out-of-home activities.
As one kind of ICT, telepresence robots offered the learners a new way to be virtually mobile in a remote place and interact with the people there. Instead of only being present at a remote place, learners who stay at home can actually control a robot to move around at a remote place with the mobility of robots. Chapter 1 of this dissertation mentioned that Bell et al. (2016) classified telepresence technologies in three generations: 2D telepresence approach using a fixed camera and screen; 2.5D telepresence approach using mobile devices such as an iPad and a tripod; and 3D telepresence approach using telepresence robots. Compared to the first two generations, telepresence robots allow learners to control the movement of robots and the view angle of the camera on the robot to experience social activities and make meaning of the foreign language by communicating with the native speakers at the remote place.

Furthermore, Bell et al. (2016) reported that the mobility of telepresence robots offered learners better social presence than the other two generations. Social presence is a concept proposed by Gunawardena (1995) to refer to the degree that a remote learner feels perceived as ‘a real person’ and has a positive influence on remote learners’ attitudes and learning outcomes (Richardson & Swan, 2003; Szeto & Cheng, 2016). In the first article of this dissertation, the learner participant reported a similar experience of social presence, such as ‘the robot makes me feel like manipulating a virtual person walking in a virtual world’ (p. 38); ‘I feel as if I was hiding in a corner watching the foreigners and experiencing many things’ (p. 38); and ‘when he (the instructor) took me around, it feels like a teacher guiding a student, and I paid attention to everything he said’ (p. 39).

The second article further connects the concepts of mobility to learner agency. Learner agency refers to the learners’ ability to take an active role and make choices intentionally in the learning process (Mercer, 2012; Xiao, 2014). Compared to the concept of social presence,
learner agency emphasizes more of learners’ activeness, rather than passively being present at a remote place. The findings of the second article show that the mobility of telepresence robots provided the learners the chances to have more agentic behaviours, such as observing the objects in the remote environment actively and choosing an observed object as the starting point of the conversation based on the learners’ background and interests. The instructor also played an important role in encouraging the learners to move and look around to choose the topics they were interested in.

The third article further demonstrated how different types of spaces influenced the learner roles that might have connections with learner agency, since the definition of learner agency involves the learners taking an active role. Specifically, learners usually play the social role of a visitor at an exhibit/display space at which the instructor plays the role of a tour guide to demonstrate and describe the objects in the environment. The role of the learner, in general, tended to be more passive, although the learner could ask more questions, expand the topic, and debrief what the instructor has demonstrated. The learner played a more active role when the instructor and the learner stayed at the special use space, especially in an emic way, i.e., the instructor and the learner engaged in the social activities at the space directly. For example, the learners would play a role of a friend of the instructor when the instructor and the learner pretended to have a picnic on the lawn. At the passage spaces, learners played more active roles when the instructor adopted the strategy of talking while walking compared to other two strategies, leading and carrying.

In brief, the mobility of telepresence robots provided a new way for learners to be mobile at a remote place in the target-language countries for a foreign language learning purpose. The mobility in turn improved learners’ social presence and could potentially improve learner agency
during the learning activities. However, the instructors played an important role in this process. More studies still need to be conducted to examine the relationships among mobility, social presence, and learner agency.

**Implications of the Studies**

*Implications for Place-Based Learning*

Telepresence-place-based foreign language learning (TPFLL) is one kind of place-based learning (PBL). The term, place, in both TPFLL and PBL refers to a physical environment outside classroom that has the potential to support learning or teaching (Sobel, 2005, Liao et al, 2019). Both TPFLL and PBL share the same rationales that the real-life settings outside classrooms provide learners a rich context for learning. Learners can acquire not only knowledge relevant to the curriculum, but also other knowledge forms or skills such as cultural awareness.

However, TPFLL allows the learning sites to be remote places with the support of telepresence robots, which gives learners access to places that may not have been previously accessible to them. The traditional PBL only emphasizes using a local place to improve learners’ academic achievement and develop stronger ties to a local community (Elder, 1998; Sobel, 2005). Few previous PBL studies have investigated the learning situated at a remote place. The studies in this dissertation supplement the current PBL theories and practices by demonstrating the affordance of remote places for foreign language learning.
More extra research areas can then be identified by analyzing the relationship between PBL and TPFL. As shown in Figure 5.1, TPFL is an area that integrates place-based learning, foreign language learning, and telepresence technologies, since PBL may not use telepresence technologies and PBL can cover wider disciplines than FL learning such as environmental education, history, and geography (Elder, 1998; Sobel, 2005). Three extra research areas arise in the Venn diagram in Figure 5.1 accordingly. Table 5.1 gives more detailed information about these three areas.

**Table 5.1. Three overlapping research areas relevant to PBL, FL, and telepresence technologies**

<table>
<thead>
<tr>
<th>Area</th>
<th>Abbreviation</th>
<th>Description</th>
<th>Existing research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place-based foreign language learning</td>
<td>PBFL</td>
<td>Foreign language learning taking place outside the classroom</td>
<td>Fajaria (2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wagner (2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Holden and Sykes (2011)</td>
</tr>
<tr>
<td>Telepresence-place-based learning</td>
<td>TPBL</td>
<td>Learning via telepresence technologies for disciplines other than foreign language</td>
<td></td>
</tr>
<tr>
<td>Telepresence-supported foreign language learning</td>
<td>TFLL</td>
<td>Foreign language learning via telepresence technologies</td>
<td>Tanaka et al. (2014)</td>
</tr>
</tbody>
</table>
Firstly, PBFL refers to the foreign language learning activities which take place outside the classroom at a local or remote place without using telepresence technologies. For example, Fajaria (2013) conducted a study in which 20 English FL high-school students participated in a series of FL learning activities outside classroom. Wagner (2015) investigated how newcomer of a language community learn the new language in daily life. PBFL would also include studies using technologies other than telepresence technologies. For example, Holden and Sykes (2011) used an augmented reality app on a mobile phone to help students learn French at a neighborhood in Albuquerque, New Mexico.

It would be reasonable to transfer the findings in this study to PBFL studies since PBFL and TPFLL both take place out of classroom for language learning purpose. For example, the second article of this study has proposed three design principles ‘situate the learning in real life settings’, ‘scaffold the learning process’ and ‘enhance learner agency’. The first principle of TPFLL would be naturally compatible with PBFL. The principle of ‘scaffold the learning process’ would be also applicable especially when novice language learner is with an expert or native language speaker. The third principle could be also applied in PBFL to improve the learners’ motivation and engagement. The main difference between PBFL and TPFLL is that TPFLL use a place at the target language country as the learning site, so FL learners have the chances to engage in more authentic conversations in a relatively lower cost compared to travel to the place physically.

However, PBFL and TPFLL might have significant difference from place-based learning in other disciplines. For example, Land and Zimmerman (2015) proposed and examined four design principles for place-based science learning with the support of mobiles including: (1) support learners to make scientific observations outdoors; (2) use child-centered designs for
mobile, informal learning; (3) use conceptual models versus discrete factual knowledge tied to singular object; (4) scaffold complex disciplinary practices in natural settings. Although these findings are solid for science education, it needs further investigation to examine if the four principles can be transferred to PBFL or TPFL learning activities, considering that the FL learners might not be adult learners and factual knowledge is also important for FL learners to describe a fact in real conversational situations.

Secondly, it is possible to use telepresence-place-based learning (TPBL) to refer to the learning activities for disciplines other than foreign language, such as history, geography, and biology with the support of telepresence robots (see the top right center portion of Figure 5.1). For example, a learner in America can learn the history of China by controlling a telepresence robot walking at a historical spot, say, the Great Wall, and talking to a tour guide or instructor there. However, few studies have been conducted in this area, so it could be a potential research topic for future studies.

Thirdly, TFLL (telepresence-supported foreign language learning) refers to the learning activities using telepresence technologies to mediate the foreign language learning. However, TFLL may happen in classrooms. For example, Tanaka et al. (2014) examined how children at a classroom in Japan communicated with a telepresence robot controlled by a remote English teacher in Australia. A main difference of TFLL and TPFL is that TFLL doesn’t have a place-based component, i.e. TFLL is situated in classroom settings. Although Tanaka et al.’s study showed that telepresence robots could significantly improve learners’ responses during the learning activities, a downside was that the most of the conversation (58%) were relevant to robot itself rather than diverse topics relevant to the remote environment in the TPFL learning activities. Another subsequent downside could be novice effects that take place in the application
of new technology (Liu et al., 2009), i.e., learners’ interests and responses may decrease after they have got familiar with the new technology. Thus, if most conversational topics are relevant to the robot itself, the effect to improve learners’ responses may disappear after learners get familiar with the robots enough. However, there might be other benefits to use telepresence robots for foreign language in classroom settings. For example, it would be easier to have multiple learners to learn together. More studies need to be conducted to compare the differences between TFLL and TPFL.

Implications for communicative language teaching

The design of the TPFL activities follows the approach of communicative language teaching since the goal of TPFL is to improve FL learners’ communicative competence by exposing and engaging FL learners to authentic language use in the remote target language countries. Compared to traditional language teaching approaches that are focused on language forms such as vocabulary and grammar, communicative language teaching (CLT) emphasizes engaging in real and meaningful communication in a given social context (Richards, 2006). Learners are expected to acquire a foreign language implicitly during meaningful and spontaneous conversation, just like how a child acquires his or her native language (Krashen & Terrell, 1983). Consequently, the instructor intentionally avoids correcting FL learners’ errors in language forms in the conversation to help FL learners better focus on the meaning and fluency of their conversation (Richards, 2006).

However, most traditional CLT occurs in the classroom (e.g., Khadka, 2017; Yu, 2001). Although scholars (e.g., Richards, 2006; Siberstein, 1977) have stated that the CLT activities in the classroom should be “intended as a preparation for survival in the real world” (Richards, 2006, p. 20) or “parallel the “real world” as close as possible” (Siberstein, 1977, p. 51), most
CLT activities such as information-gathering activities, opinion-sharing activities, and role playing (Richards, 2006) have limitations in imitating the conversational situations in real-life settings, since the real conversational scenarios involve native English speakers, authentic environments, and real social events.

Based on the findings of the three articles in this dissertation, telepresence robots provide a way to make the TPFL activities as close to the “real world” as possible. For example, the participants in the first article described that they felt “the interaction and the use of technology to actually experience real things” (Liao & Lu, 2018, p. 25) and “The robot makes me feel like manipulating a virtual person walking in a virtual world, … Meanwhile, I know the world where the robot is walking is real world, … This makes me excited” (Liao & Lu, 2019, p. 26). The survey results in the second article (Liao et al., 2019) show that most participants agree that the TPFL activities situated in the remote real-life setting are helpful for improving their listening, speaking, and cultural awareness. The third article further points out that the authenticity of language use is highly related to the remote place, as well as the people and the social actions at the place from a lens of geosemiotics.

Furthermore, TPFL can address the teaching resource issue in traditional CLT activities. Studies (e.g., Marcellino, 2005; Rao, 2013) have found that many English teachers, especially in middle schools in Asia, lack both the confidence and the English proficiency to conduct CLT due to their insufficient English training experience. Also, the class sizes in these Asian countries usually reach between 40 and 60 people. In such environments, FL learners barely have the chance to communicate with people who have advanced English proficiency. Telepresence robots provide a way for these FL learners to gain access to native English speakers in English-speaking countries.
In addition, as mentioned in the last section, the instructor needs to scaffold the learners during the learning process in one-on-one sessions, instead of only facilitating the conversation among multiple learners. This method provides a new perspective for understanding the social interactions in TPFLL and the role of the instructors. More studies can be conducted to explore the specific techniques of scaffolding and their effectiveness.

In general, TPFLL extended the location of CLT from classrooms to real-life settings in the target countries. The places, the native speakers, and the social activities at the places provided rich language and cultural contexts for FL learning. FL learners can then be exposed to and engaged in authentic language use for improving their communicative competence.

*Implications for Designs of Educational Telepresence Robots*

So far, few studies have investigated the design of telepresence robots from the perspective of language learning or teaching. The findings of the three articles in this dissertation provide insight into the design of the appearance and functions of the telepresence robots. The quality of audio and video streaming provided by the robot basically met the needs of TPFLL activities. The learner could see clear video streaming of the instructor and the remote environment at most times. The video quality of Keebot was better than Romo, as Keebot is a more advanced telepresence robot, and the network speed in the second and third study was improved compared to the first pilot study. Interestingly, one of the learner participants finished the learning activities with her mobile phone while she was on a subway in China. It is assumable that the audio/video quality of the telepresence robots and any network delay will be further improved with the support of 5G information technologies that will be available in the near future.
However, five limitations were found in the current design of telepresence robots for language learning use. First, the heights of both robots, Romo (about 8 inches) and Keebot (35 inches), are too short. Although Keebot is much taller than Romo, the instructors generally had to crouch down or sit on the ground to maintain eye contact with the learner on the robot’s head screen (See Figure 3 and Figure 5 in the second article). A close analysis on the learner view in the third article (Liao et al., 2019) shows that the learner could only see both the instructor and the objects on a screen simultaneously when the instructor was crouching and the object being referred to was on the ground (See Figure 3 and Table 3 in the third article). In other words, if an instructor stands close to the robot, the learner has to tilt the robot’s head up to make eye contact with the instructor. Meanwhile, the learner is not able to see the objects on the ground due to sight limitation. It would be easier for the instructor to talk while standing if the height of the robot was closer to the height of an adult.

The second issue is the size of the screen on the telepresence robot (Liao & Lu, 2018, p. 28). Romo uses a phone such as an iPhone to show the remote learner’s face, while Keebot uses an iPad. The instructor reported that the size of the screens was too small to see the gestures and body language of the remote learner, which is important to convey information during interaction from the perspective of multimodality of interaction (O’Halloran, 2011). It would be better to have a bigger screen that can cover not just the learner’s face, but their upper body and arms as well. In addition, the learner shown on the screen should be observable even under strong sunshine since an instructor reported it was difficult to see the learner’s face on the screen in strong sunlight (Liao & Lu, 2008).

The third issue is relevant to the robot’s audio quality. The first article (Liao & Lu, 2018) shows that the volume of Romo was so low that the instructor had to position himself very close
to the robot in order to hear. However, the instructor participants didn’t report this issue in the later studies (Liao et al., 2019) since the volume of Keebot is much higher than that of Romo. The instructor can hear the voices of learners at a distance of about six feet (See Figure 7 in the second article as well as Table 6 and Table 8 in the third article). Also, it is reported in the third article that the noise of the moving robot was rather loud, which may hinder the FL learners and the instructor from hearing the conversation clearly while the robot is in motion.

The fourth issue is the mobility of the telepresence robots. The speed of both Romo and Keebot was slow, so some instructors chose to carry the robot to save time if the pathway between two talking spots was too long. Also, the motion of the robots was not steady when the ground was not flat enough. Although both Romo and Keebot can walk on concrete ground, the robots may fall down if there are small gaps in the pathway. The robot cannot walk on other terrains such as grass, so the route of the learning activities has to be along the concrete ground, which may limit the learner from engaging in social activities in those places. It would be better to have faster and steadier driving designs for the robots.

Finally, both Romo and Keebot don’t have mechanic arms, so the FL learners are limited from directly engaging in some social activities such as playing baseball (see emic engagement in the third article). Consequently, FL learners have few chances to be exposed to the conversations that will be used by the social roles in these social activities. Fortunately, some telepresence robots have been designed with mechanical arms (e.g., Tanaka et al., 2013). It would be interesting to investigate how mechanical arms can help the FL learner engage in social activities as well as conversations in the activities.

In general, current telepresence robots meet the needs of TPFLL activities especially for display, exhibit, etic engagement, and passage purposes (see seven patterns at three types of
spaces in the third article) with the video streaming and moving functions of the robots. In order to improve the learning experience for emic engagement into the social activities and conversations at special use spaces, the telepresence robot needs to be more like a real human being, such as moving faster and steadier, showing the FL learner’s upper body, and having mechanical arms to make gestures or engage in social actions such as playing baseball. These demands are challenging to implement but might be realized in the future with the improvement of robotics.

**Pathway to Applying TPFLL in Educational Practice**

Although this study has shown the potential to use telepresence robots at a remote place for FL learning purposes, the current research is still at the beginning stage. More research is needed to disseminate the innovative way of using telepresence robots in the educational practice of foreign language learning. As far as the author’s concerns, there are four possible ways to apply TPFLL in educational practice at a large scale.

One scenario is for local foreign language training organizations, in which foreign language learners would have one-on-one training sessions with local language tutors. The studies in this dissertation have demonstrated the affordance telepresence robots have on one-on-one TPFLL activity sessions. Instead of having local foreign language instructors tutor the learners, the training organization can hire remote instructors from the target-language country. A downside of this method is the cost of hiring the remote instructors since each instructor can only tutor a few learners one-on-one. However, it is possible to recruit volunteers as instructors. For example, Granny Cloud is an organization for recruiting volunteers in Western countries to help poor children in other countries learn through interaction in English (Cadwalladr, 2015). The volunteers are labeled as ‘grannies,’ since most of them are retired or aged people, who have
idle time and may feel lonely in their lives. The grannies can use computers and Skype to communicate with the children in the remote countries. In the case of TPFLL, the volunteers can use telepresence robots to communicate with the FL learners at a place near the volunteer.

The second scenario would be intensive language programs in the target-language countries. As mentioned in the Chapter 1, currently the cost of those intensive language programs is extremely expensive for most FL learners globally, especially in the developing countries. However, those programs will not need the FL learners to pay the fees for boarding and travelling if they use telepresence robots to conduct TPFLL activities, which will save considerable amounts of money and make it more affordable for more FL learners.

The third scenario could be building an informal bond between FL learners and native English speakers. This case is more similar to traditional “pen pal” activities, in which people from different countries can make friends through letter-writing. Instead of mailing letters or using digital correspondence like email, Facebook, Twitter, or Skype, it is possible for people to make friends and increase their bonding by using telepresence robots. Language learning would naturally take place while a youth guides his or her friend from another country to virtually visit him or her at home, a park, or any place that they can have conversations.

Finally, the last scenario would be in formal education. In this case, the main concern is that the number of students in a class, especially in Asian countries, is between 40 and 60, which makes it impossible to have one-on-one TPFLL sessions. In this case, it is important to explore the way to share a telepresence robot with more students in one session and build a curriculum that integrates traditional classroom teaching and the TPFLL learning activities. More studies still need to be conducted to examine the effectiveness of TPFLL in group settings and to explore the ways to build an integrated curriculum.
No matter which scenario, one concern that needs to be emphasized is ethical dilemmas because telepresence robots capture videos and videotape what the remote user can see. For the purpose of this research, it is possible to have every participant consent to the use of their voices or videos. However, it would be inevitable to involve other, possibly non-consenting people at a public space in the screen if the TPFLC is eventually used at a large scale. Faklaris et al. (2016) reviewed the relevant policies for using mobile live-streaming video apps, which are similar to using live-streaming videos in TPFLC. Generally speaking, individuals or organizations have the right to publicity, which prevent the unauthorized use of their image or identifiable information. However, there are some exceptions depending on the specific policies in different countries. For example, recordings of some kind of employees such as nannies via hidden cameras on private property are generally allowed under U.S. federal and state laws. But still, it is important to make sure that the use of telepresence robots is legal and ethical in specific situations. Also, it is possible to provide more designs in the telepresence robot system to protect other individuals’ privacy, such as the features in the app Snapchat to prevent archiving the videos.

In brief, telepresence robots are one kind of relatively-new technology. With the development of the ICT and mobile devices, the cost of the telepresence robot itself has decreased, and the quality of audio-video streaming and chatting has improved significantly in the last decade. It is predictable that telepresence robots and the resulting pedagogical methods will be applied in FL teaching in the coming years. This section lists possible application scenarios and the practical concerns for the application at a large scale.

**Boundaries of the Study and Future Research**

As an exploratory study, this dissertation focuses on the perceived benefits and challenges of using telepresence robots in foreign language learning in real-life settings. Also,
the design of learning activities and the robots are further examined based on the analysis of the learning process. However, this study doesn’t measure the learning outcomes at this stage, since it is not clear what factors will influence the learning outcomes. Moreover, the learning sites of this study are limited to two sites including one of the main buildings and an arboretum at an American university campus. It is not clear if the proposed design principles can be applied to other learning sites such as a hotel or a parking lot with more participants.

Accordingly, more studies can be conducted in the future to measure the learning outcomes systematically and to examine the design principles at diverse learning sites with more participants. More research questions for future studies include: (1) to what degree can the TPFLL activities improve FL learners’ communicative competence; (2) compared to other technologies such as Skype, can telepresence robots improve learners’ communicative competence more; (3) what are valid and reliable instruments to measure FL learners’ communicative competence; and (4) what other learning outcomes can TPFLL improve?

In addition, more research topics relevant to TPFLL are interesting and haven’t been investigated yet. For example, (1) the design of TPFLL curriculum, (2) how TPFLL can be integrated with FL activities in classroom, and (3) what approaches or techniques can be used to better scaffold FL learners before, during, and after the TPFLL activities.

In general, the studies on the TPFLL are still at an early stage. The telepresence robots have much potential to support foreign language learning, so many topics can be investigated further in the future.

**Conclusion**

The three articles in this dissertation examine the perceived benefits and challenges of using telepresence robots in real-life settings for foreign language learning purposes. The results
have shown that telepresence robots can provide FL learners with more authentic language learning environments than classroom settings or even than virtual reality settings. The instructor and the FL learners can naturally develop their conversational topics based on the observed objects, buildings, people, and social activities at the place. The learning process was learner-centered since the robots allow FL learners to control what they want to see and then respond based on what they have seen. However, some learners with low English proficiency felt it was difficult to deal with the unscripted interaction.

Accordingly, three design principles were proposed to guide the design of TPFL activities including situating the learning in the real-life settings, scaffolding the learning process, and enhancing learner agency. Most participants held positive attitudes towards these design principles. Then, the design of learning activities and robots was further investigated by analyzing the learning process through the lens of geosemiotics. The results show seven patterns at three space types including display/exhibit space, special use space, and passage space. Consequently, the design concerns on each space type were proposed.

In summary, the study shows the potential of telepresence robots to support foreign language learning, especially for improving FL learners’ communicative competence in authentic real-life settings. More research topics are suggested for guiding future studies.

**References**


APPENDICES

Appendix A . Interview Questions for Article 1 and Article 2

• Interview Questions (Remote Participant / Language Learner)

1. How do you feel about the procedure of the activity?

2. How do you feel about the interaction between you and the native speaker?

3. How do you feel about the site of the activity?

4. What are the differences between talking to others via a telepresence robot outdoors and face-to-face interaction?

5. What are the differences between talking to others via a telepresence robot outdoors and talking to others via other technology, such as Skype?

6. Have you imagined yourself as an international student at an American university before? How about now?

7. Did you feel any cultural differences between China and the US during the activities?

8. Do you have any suggestions for improving the system?

• Interview Questions (Local Participant / Native Speaker)

1. Do you have any experience in teaching English as a second language?

2. Have you used any technology to improve your English before?

3. How do you feel about the process of communication with others via this robot?

4. What are the differences between talking to others via a telepresence robot outdoors and talking to others via Skype at home?

5. Do you feel the telepresence robot is helpful for teaching a foreign language or not?

6. Do you have any suggestions for improving the system?
Appendix B. Observation Guide for Article 1 and Article 2

This observation guide describes what will be taken in the field notes and captured in video clips by researchers during the activity in the study. These observed data would be used to triangulate the findings in the interview and survey about the experience of remote participants and local participants on the learning environment.

<table>
<thead>
<tr>
<th>Data</th>
<th>Description</th>
<th>Documentation</th>
<th>Typical Research Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setting</td>
<td>How the scene and people look</td>
<td>Observation notes, video</td>
<td>Introducing setting and people in final report</td>
</tr>
<tr>
<td>Appearance</td>
<td></td>
<td></td>
<td>Raises questions for interviews; supports or challenges interview data; thick description; pattern analysis; generates hypotheses</td>
</tr>
<tr>
<td>Acts</td>
<td>What the people do during the activity</td>
<td>Observation notes, video</td>
<td>Raises questions for interviews; supports or challenges interview data; thick description; pattern analysis; generates hypotheses</td>
</tr>
<tr>
<td>Events</td>
<td>What are the plans and steps of the activity</td>
<td>Observation notes, video</td>
<td>Raises questions for interviews; supports or challenges interview data; thick description; pattern analysis; generates hypotheses</td>
</tr>
<tr>
<td>Processes</td>
<td>Rules about how the telepresence communication activity works</td>
<td>Observation notes, video</td>
<td>Raises questions for interviews; supports or challenges interview data; thick description; pattern analysis; generates hypotheses</td>
</tr>
<tr>
<td>Talk</td>
<td>What people say to each other</td>
<td>Observation notes, video</td>
<td>Raises questions for interviews; supports or challenges interview data; thick description;</td>
</tr>
<tr>
<td>Documents</td>
<td>What would the remote participants prepare for the activity</td>
<td>Web address of online materials or the name of books,</td>
<td>Raises questions for interviews; supports or challenges interview data; thick description; pattern analysis; generates hypotheses</td>
</tr>
<tr>
<td>Artifacts</td>
<td>Videotaping</td>
<td>Observation notes, video</td>
<td>Raises questions for interviews; supports or challenges interview data; thick description; pattern analysis; generates hypotheses</td>
</tr>
</tbody>
</table>
Appendix C. Demographic Information Survey for Article 2

Please answer these questions:

1. Gender (Male/Female)
2. Age
3. Are you students?
4. What is the highest degree you have now?
5. How many years have you being learning English?
6. What is your English level now?
7. Are you still having English class or training now? If so, what kind of class or training?
8. Which part of language is hardest for you? (Speaking/Listening/Reading/Writing)
9. Do you have any experience living or travelling in a foreign country?
10. What are the challenges in your language learning?
11. What mobile device will you use in the learning activity?
12. Have you used any technologies in your language learning? If so, how do you use them for what?
**Appendix D. IRB Exemption Determination for Article 1**

**Date:** April 8, 2015  
**From:** Jodi Mathieu, IRB Analyst  
**To:** Jian Liao

<table>
<thead>
<tr>
<th>Type of Submission:</th>
<th>Initial Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title of Study:</strong></td>
<td>Penn State Life – An Immersive Second Language Learning Environment in Real World</td>
</tr>
<tr>
<td><strong>Principal Investigator:</strong></td>
<td>Jian Liao</td>
</tr>
<tr>
<td><strong>Study ID:</strong></td>
<td>STUDY00002111</td>
</tr>
<tr>
<td><strong>Submission ID:</strong></td>
<td>STUDY00002111</td>
</tr>
<tr>
<td><strong>Funding:</strong></td>
<td>Not Applicable</td>
</tr>
<tr>
<td><strong>Documents Approved:</strong></td>
<td></td>
</tr>
</tbody>
</table>
- HRP-591_Leo(revised) (0405(0.02)), Category: IRB Protocol  
- InterviewQuestions(Revised).docx (0405(0.02)), Category: Data Collection Instrument  
- Survey(Revised) (0405(0.03)), Category: Data Collection Instrument  
- Observation-Guide.docx (0405(0.02)), Category: Data Collection Instrument |

The Office for Research Protections determined that the proposed activity, as described in the above-referenced submission, does not require formal IRB review because the research met the criteria for exempt research according to the policies of this institution and the provisions of applicable federal regulations.

Continuing Progress Reports are **not** required for exempt research. Record of this research determined to be exempt will be maintained for five years from the date of this notification. If your research will continue beyond five years, please contact the Office for Research Protections closer to the determination end date.

Changes to exempt research only need to be submitted to the Office for Research Protections in limited circumstances described in the below-referenced Investigator Manual. If changes are being considered and there are questions about whether IRB review is needed, please contact the Office for Research Protections.

Penn State researchers are required to follow the requirements listed in the Investigator Manual (HRP-103), which can be found by navigating to the IRB Library within CATS IRB (http://irb.psu.edu).

This correspondence should be maintained with your records.
Appendix E. IRB Exemption Determination for Article 2 and 3

EXEMPTION DETERMINATION

Date: July 11, 2016
From: Jodi Mathieu, IRB Analyst
To: Jian Liao

<table>
<thead>
<tr>
<th>Type of Submission</th>
<th>Initial Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title of Study</td>
<td>Penn State Life – An Foreign Language Learning Environment based on Telepresence Robots</td>
</tr>
<tr>
<td>Principal Investigator</td>
<td>Jian Liao</td>
</tr>
<tr>
<td>Study ID</td>
<td>STUDY00005104</td>
</tr>
<tr>
<td>Submission ID</td>
<td>STUDY00005104</td>
</tr>
<tr>
<td>Funding</td>
<td>Not Applicable</td>
</tr>
</tbody>
</table>

Documents Approved:
- HRP-591(Revision2).pdf (0722016), Category: IRB Protocol
- InterviewQuestions.docx (0.01), Category: Data Collection Instrument
- Observation-Guide.docx (0.01), Category: Data Collection Instrument
- Pre_Post_Follow-up Speaking Test.docx (0.01), Category: Data Collection Instrument
- Survey.docx (0.01), Category: Data Collection Instrument
- Word Quiz.docx (0.01), Category: Data Collection Instrument

The Office for Research Protections determined that the proposed activity, as described in the above-referenced submission, does not require formal IRB review because the research met the criteria for exempt research according to the policies of this institution and the provisions of applicable federal regulations.

Continuing Progress Reports are **not** required for exempt research. Record of this research determined to be exempt will be maintained for five years from the date of this notification. If your research will continue beyond five years, please contact the Office for Research Protections closer to the determination end date.

Changes to exempt research only need to be submitted to the Office for Research Protections in limited circumstances described in the below-referenced Investigator Manual. If changes are being considered and there are questions about whether IRB review is needed, please contact the Office for Research Protections.

Penn State researchers are required to follow the requirements listed in the Investigator Manual (HRP-103), which can be found by navigating to the IRB Library within CATS IRB (http://irb.psu.edu).

This correspondence should be maintained with your records.
VITA

Jian Liao is a Ph.D. candidate in the program of Learning, Design, and Technology at the Pennsylvania State University. He has a strong desire to facilitate online learning with various learning technologies. His research interests include learning sciences, informal learning, online-learning, telepresence-robots, and gamification. So far, he has published nine academic journal articles and 28 conference papers. He is also a designer and developer working on various projects relevant to learning, teaching, and assessment with a wide range of technologies including website development (PHP, Drupal, React, Python, JavaScript, CSS, GitHub), database management (MySQL, Postgres), mobile app development (Objective C), and server management (Linux, Apache, AWS).

Education
- The Pennsylvania State University, University Park, PA. Ph.D. in Learning, Design and Technology Aug. 2013 - Present
- Beijing Normal University, China, M.S. in Educational Technology, 2005 - 2008
- Southwest University, China, B.E. in Computer Sciences, 1996 - 2000

Selected Projects
- AvePM (Progress monitoring with students who are deaf or hard of hearing) (https://avenuepm.org)
- CEAPP (Corpus of English for academic and professional purpose) (https://ceapp.la.psu.edu)

Selected Journal Articles and Conference Proceedings

Awards
- Ralph T. Heimer Award, College of Education at Pennsylvania State University, 2019-2020
- Dean’s Graduate Assistantship, College of Education at Pennsylvania State University, 2013-2015