

---

# Investigating Social Presence in Audio and Text Online Interaction for Language Learning

---

Ligao WU  
Beijing Foreign Studies University

## *Abstract*

This study aimed at comparing the level of social presence generated in a voice-based chat room and a text-based forum when learners tried to build personal relationships and form an online community for learning on an online language course in China. A mixed-method approach was taken for the study, drawing on data from questionnaires to find out about student perception of social presence, and postings of text messages and audio messages in the communication of the student learning process to search for students' projected social presence in terms of affective, interactive and cohesive features. Interviews were also conducted to supplement additional information with the hope of forming a complete picture of social presence in the reality of an online learning environment. The text-based forum and the voice-based chat room were found to have a different impact on student social presence. In terms of student perception, most of them were more likely to get to know peers in the text-based forum and thus developed a sense of community in their learning process of the online course. Yet they believed that the voice-based chat room had the advantage of helping them with course learning. In the actual interaction, the voice-based chat room was more interactive although the text-based forum was more affective and cohesive. But in terms of the affective category, the problem with the existing framework in literature was that there were no prosodic features included. Therefore, in future more research is needed to probe for the relationship between prosodic sound features and social presence, and the present theoretic framework must be extended. In interviews, students explained that in the voice-based chat room prosodic features led to higher peer awareness, which further reinforced this need.

*Keywords: social presence, learning communities, text-based forum, voice-based chat room*

## 1. Introduction

The evolution of online learning centers on educational technology. At the present time, information technology employed in online learning has been constantly changing and thus transformed the ecology of online learning. Two-way communication is now dominating, and the underpinned learning theory for online education has shifted its focus from student independence to their interdependence (White, 2009). Personal relationship building and online interaction for community forming become key issues in the new educational setting (Whiteside, 2015; Wu, Gao, & Zhang, 2014).

The particular physical detachment of students and teachers in the online learning environment motivates them to have online socialization and presence through computer-mediated communication (Dow, 2008). They try hard to be sociable, sensitive, personal, and warm to develop a sense of relating and caring in order to form an online learning community through the online media (Tu, 2002b). For online language learning programs, learners are connected through online interaction so that they practice their language skills and social skills (Terhune, 2016).

The concept of social presence is an umbrella term to include signs of interactive enhancement, expressions of affection and group cohesive devices employed in the computer-mediated communication during the online learning process (King & Ellis, 2009) and it has long been considered to be a key factor in recognizing personal relationship building in the online learning setting, and a high level of social presence benefits student interaction and learning (Garrison, Anderson, & Archer, 2000; Gunawardena & Zittle, 1997; Kehrwald, 2010; Ko, 2012; Tu, 2002b).

Social presence has been much researched in text-based communication; however, new modes of communication, such as audio or video require researchers to engage in constant re-evaluation of the concept of social presence with regard to how media affects the computer-mediated communication for interpersonal relationships (Satar, 2015; Satar & Akcan, 2018). This study set out to investigate computer-mediated communication in a voice-based chat room and the text-based forum in the Institute of Online Education, Beijing Foreign Studies University. The main aims were to compare the level of social presence generated using these two technological tools, to explore how learners utilize the tools to build personal relationships and form an online community, how they practice what they have learned supporting each other, and a further aim was to contribute to social presence theory in this area concerning technological factors as well as contextual factors.

## 2. The Concept of Social Presence

The concept of social presence was defined in earlier studies as the “degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationships...” (Short, Williams, & Christie, 1976, p. 65). This means that in a communication process with technological media, how much those participants perceive each other is

important, and this will be decisive in the building of participant relationships and the communication quality that will result from this effect. In other words, besides language, visual cues are essential in establishing social presence and thus greatly improve the quality of the communication. When we utilize a technological medium for communication, those showing more visual cues provide more opportunities to establish a higher social presence and thus the communication quality is thought to be raised to a higher level, suitable for a communicative task that involves a complexity of relationship building (Short et al., 1976).

Applying the principle, we may conclude that when a conference room with audio only and the audio plus video are used for communication, participants using the latter tool see more of the other participants in the process, and the room with audio plus video is thus considered to have more social presence than that with audio only, and therefore the tool with audio plus video is a more efficient medium for communication. Moreover, an audio tool is likely to enable participants to perceive more of the other participants than a text-based tool because the prosodic features in the audio messages reveal much of the speakers, and thus the former tool is more effective than the latter in connecting people in the communication.

Social presence becomes a critical issue in online learning settings because learners and teachers are separate in physical space, and text, audio, and video media are employed for learner-learner and learner-teacher communication. Researchers have begun to employ social presence theory to assess the quality of multimodal interaction (Ko, 2016; Satar, 2015). In the existing literature, conflicting evidence has been found with some studies claiming that voice-based communication is less interactive and effective than text-based communication and thus unlikely to increase social presence (King & Ellis, 2009), while other studies state that voice, with its capacity to carry nonverbal cues, has the potential to enhance social presence (LaPointe, Greysen, & Barrett, 2004; McIntosh, Braul, & Chao, 2003). Studies of video-based communication have encountered a similar situation with different conclusions about what video can do for social presence (Borup, West, & Graham, 2012; Homer, Plass, & Blake, 2008; Traphagan et al., 2010). In Borup et al.'s (2012) study, social presence was enhanced by the video clips produced by the teacher and fellow students for communication in the online learning process. On the contrary, in Homer et al.'s (2008) study, the video image of the teacher in the lecture did not improve social presence. If there is no consensus of ideas, the research is not particularly helpful in providing guidance for future research or for the practical use of technology in computer-mediated communication. A study is needed to explore the reasons behind the conflicting views, or perhaps suggestions for the modification of the social presence framework that should be provided in order to elicit more harmonious guidance.

### **3. Practical Situation and Motivation for the Study**

In the Chinese online learning context, the text-based forum and the voice-based chat room are the most common tools for the two-way communication between learners and their teachers. Besides the two institutional tools, social media devices like QQ and WeChat are also popular channels for online learners in the country. MSN used to be a well-known medium

for online communication in China, but it stopped its service in 2014 and thus disappeared in the Chinese online learning context. Table 1 shows the tools that Chinese online learners have frequently used in recent years and their functions.

**Table 1.** Technological tools of computer-mediated communication for Chinese online learners

Communication form	Electronic hardware that supports it	Functions enabled
Text-based forum	Computers	Broadcast text messaging, sharing of photos and videos
Voice-based chat room	Computers	Instant talk, text messaging
MSN	Computers	Instant text messaging,
QQ	Computers, cell phones, personal digital assistants (PDAs)	Text messaging, hold-to-talk voice messaging, broadcast messaging, sharing of photos and videos
WeChat	Computers, cell phones, personal digital assistants (PDAs)	Text messaging, hold-to-talk voice messaging, broadcast messaging, sharing of photos and videos

A text-based forum is usually located in an enclosed platform of a learning institution. Learners have to log onto the forum with their user names and passwords. The text-based forum is a convenient channel of communication because participants can access it in their own time. When a message is posted on the forum, it reaches every co-learner once they are connected to the internet. It is also a great place to share digital information in an audio and video form. The text-based forum provides a communication channel that allows participants to take time to respond, and this is particularly useful for part-time students. When their schedules are tight, the forum allows them to produce a delayed response to fellow learners. It enables them to make use of bits and pieces of their spare time. This asynchronous nature of communication is also recommended as beneficial for reflection and deep thoughts for the online interaction (Paulus, 2007). The voice-based chat room is a particularly useful tool for language students. The voice facility provides a chance for them to practice listening and speaking, which is difficult in the text-based environment. The synchronous nature of the communication with the voice-based chat room is also a great tool, which is likely to be more effective for talking about sensitive issues online.

This study involves students who study English through Literature, an online course in the second term of their second year in a bachelor's degree program of English language learning. The course has eight units which are designed to include literary works that represent different genres and historical periods. Each of the eight units adopts a particular perspective to examine literature with some sections of literary work included as explanatory samples. The purpose of the course is to learn English language as well as to acquire knowledge about English literature.

The assessment of the course consists of three parts: self-assessment exercises, assignments, and the final examination. The final examination accounts for 70 percent of the total score. The other two parts make up 30 percent. The final exam takes place on campus, which requires students to be physically present. Students submit their self-assessment exercises and

assignments online. The former comprises multiple-choice questions and is scored on the computer, and the latter have short answer questions and essay questions and are marked by the instructor.

The instructional model of the online courses in the program is similar. For example, English through Literature provides four 4-hour face-to-face tutorials in Beijing with each 4-hour tutorial occurring once every two weeks. The course lasts for eight weeks. In an 8-week period, the instructor of the course provides four tutorials. In every face-to-face tutorial, the instructor covers two units of the course textbook. Because students have full-time jobs and they may live some way away, a quarter of the students may attend the tutorial. Students mostly study their textbooks on their own. During their learning process, students are expected to ask questions and discuss them with peer learners and their instructors in a text-based forum and in the voice-based chat room.

Concerned with the practical context, this study focuses on investigating social presence in the text and online audio interaction. It aims to conduct a comparison of social presence in a voice-based chat room with a text-based forum where learners are using them for learning an online course in a foreign language degree program. It explores how students employ the two technological tools for online interaction and how the two tools make an impact on student social presence and thus influence their interaction and language learning experiences.

Guided by the theory of social presence in the existing literature, the present study adopted a mixed-method approach to measuring learner perception of social presence and the level of their projected social presence in the voice-based chat room as well as in the text-based forum. Previous research has mostly been conducted into text-based interaction in which the theoretical frameworks have been formed. The results of this research will contribute to the theory of social presence by broadening the investigation into voice-based interaction, which is emerging as an important type of communication in the field of online education, particularly with foreign language learning, but which has rarely been studied in the existing research literature. The significance of the study may also lie in the fact that it provides some insights for social presence research as well as practical guidance for the application of voice-based interaction in online learning institutions.

## 4. Research Questions

In this study, the key research questions are:

- 1) Is student social presence different in the text-based forum and in the voice-based chat room?
- 2) If there is a difference, to what extent is it different?
- 3) How does the difference impact student interaction and their language learning experience?

The dearth of literature on social presence related to voice-based interaction presents a problem as computer technology is integrating sound technology rapidly in online learning, particularly in the area of foreign language learning. Thus it is essential to explore this new

area to provide some insights as well as a practical guide for the understanding of social presence with mediated interaction using voice-based tools in the field of online learning. The researcher intends to use the social presence categories (Rourke, Anderson, Archer, & Garrison, 2001) as a framework from the existing literature because there is a lack of theory on researching social presence with voice-based interaction. The dual purpose of the research is to validate the theory as well as providing suggestions to modify it for future research of a similar kind. Moreover, it is possible to provide a larger picture of social presence with mixed-method research in order to collect data to explain the conflicting evidence against voice-based interaction in the existing literature. If this is true, this research will be useful for the practical application of voice-based interaction for studying in an online course.

## 5. Research Design

The researcher adopts a mixed-method for the study by employing three methods: survey, text analysis, and interview (see Table 2).

**Table 2.** Methods employed in the study

Methods	Survey	Text analysis	Interview
Data collection	questionnaire	online interaction	recording and transcription
Data analysis	SPSS calculation	content analysis	thematic analysis

A questionnaire was used to discover learners' perceptions of social presence when they used the text-based forum and the voice-based chat room for communication during their learning process of the online course English through Literature. Text analysis was completed to search for the projected social presence in the online interaction of the students. Finally, interviews with some students went beyond the projected social presence to explore some of the possible reasons for causing the differences in social presence. The three methods were combined to make a complete study to show an all-round picture of social presence in the teaching context of the researcher.

A questionnaire was designed based on the survey questions proposed by Swan and Shih (2005). Questions were modified to suit the specific purpose of measuring social presence in the voice-based chat room and the text-based forum and translated into Chinese. The questions focused on how learners felt about using the text-based forum and the voice-based chat room for online interaction, their feelings regarding the interaction with fellow learners and the course instructor using the two technological tools and their perceived learning in the practical context. For example, "I felt comfortable conversing through the text-based forum/the voice-based chat room with course mates." "I was able to make friends with other course mates through the text-based forum/the voice-based chat room." "The text-based forum/The voice-based chat room helped me with my course learning. Students were asked to rate their

agreement on a 5-point Likert scale with 1 being “strongly disagree” and 5 “strongly agree”.

The modified questionnaire was tested in a pilot study to ensure the reliability and validity. When the results from the pilot study were entered into SPSS, the Cronbach’s Alpha was 0.869, indicating that the consistency of the questions was acceptable. After analyzing the results, the researcher believed that the questions had measured what he intended to measure. Finally, a main study was conducted with the questionnaire.

With text analysis, the researcher employed the social presence categories and indicators (Rourke et al., 2001) to analyze the recordings from the text-based forum for the English through Literature course and the voice-based chat room where students initiated activities for learning the course in the online interaction. In the recorded data, three categories were identified: affective, interactive, cohesive. Expressions of emotions, use of humor, and self-disclosure were the indicators of affective category. Continuing a thread, asking questions, and expressing agreement and similar interaction were indicators of interactive category. The cohesive category is made of indicators such as vocatives and addresses or refers to the group using inclusive pronouns and phatic salutations.

The interaction recordings from the text-based forum were taken from the threads dated August 31 to November 18, 2011, during the school term when the online course was studied. In the term, 17 threads were posted. Among them, two threads were discarded because there was no interaction. The text-based forum was the place for students to ask questions about their course learning and discuss learning issues. Students often accessed it, and in the survey, it was found that 52 percent of the students reported attending more than once a week. However, as there was no requirement for posting in the forum as part of course assessment, participation was not high, and most of the students only read the postings. Generally, in the text-based forum, students posted more than the teacher, and consequently, most of the interaction was student-student.

The 15 threads of messages were made of 96 postings. The response rate was reasonably good. Each thread had an average of 6.4 postings. When students posted messages, they used both English and Chinese. Altogether, 1,264 English words and 3,509 Chinese words appeared in the recorded data. Each thread had an average of 83 English words and 234 Chinese words. It may probably mean that students found it easier to communicate with Chinese, especially when they expected quick answers for their questions, and in the meantime, they also wanted to practice English in the text-based forum.

The interaction data recorded from the voice-based chat room came from three learning activities that students initiated for reviewing what they had learned from the course textbook and for practicing English language skills. On May 25, 2011, a group of students scheduled a time to meet in a voice-based chat room to talk about important knowledge points they have learned from their course textbooks in the term. The three activities lasted for one hour and a half with each activity for a different focus. After each activity, they took a break and then continued. The group leader chaired the three activities. She sent an outline beforehand for all the students who agreed to participate in the learning activities so that they may prepare in advance.

The recordings of the activities were transcribed. A draft transcription was produced, and recordings were checked against this initial transcription a second time to refine it. Thus, a more accurate version of what was actually said was obtained. Occasionally the recording had to be listened to repeatedly to catch the real interaction. At times it was almost impossible to capture what was said, and the researcher was forced to ignore it. Generally, the transcription was considered to be a sound reflection of the real interaction. The three sessions involved seven students actively in the conversations. Other students were passively involved in the activities and were silent. They did not speak at all, and that made it difficult to analyze their performance. Therefore, the researcher chose the data that had interaction among learners. For the convenience of analysis, the transcription of the activities was divided into 15 files. Altogether 311 turns of conversation were caught with 6,412 English words and 7,785 Chinese words spoken in the recording data.

The two coders coded the files independently and exchanged coding results. After reviewing each other's coding, any differences were examined, and an online discussion followed using Skype. An agreement was reached and considered the final decision. The numbers of their coding categories were entered into SPSS where inter-rater reliability was achieved. For the files with the text-based interaction, the Cronbach's Alpha was 0.995, which showed that there was a high level of agreement between the two codings. For the voice message transcription, the inter-rater reliability was 90.4 percent.

The interview questions intended to cover the three categories of social presence: affective, interactive and cohesive, as described by Rourke et al. (2001). The affective category concerns how the students express their feelings and how they care for each other in order to draw closer to their fellow students and make friends with them. For example, "When you use the text-based forum/the voice-based chat room, what strategies do you use to get closer to your course mates?" The interactive category concerns how they communicate their ideas and respond to what fellow students think. For example, "When you use the text-based forum/the voice-based chat room, how do you make your ideas understood by your course mates?" The cohesive category covers learners' strategies when using the voice-based chat room and the text-based forum for building a learning community in the course learning process. All the interview questions were conducted in Chinese as the students' native language is Chinese.

Interviewees could have been chosen by random sampling of the student name list at our online school. However, if the students chosen had never used the text-based forum and the voice-based chat room, the interview would be invalid. To avoid this problem, therefore, the interviewees were chosen out of a convenience principle from the face-to-face tutorial classes. They were active students with stronger motivation to learn than the average students.

From October 23 to December 1, 2011, eight face-to-face interviews were conducted with a digital sound recorder. The interviews were semi-structured with the interviewer following the main themes of the questions as guidelines, although questions might be asked with slight variations. In the eight interviews, five interviews were conducted with individual students and three with a pair of students. The reason why the three pairs of students were grouped together was that it was difficult to meet online students face-to-face. When two students met for an



interview, the interviewer decided to interview them at the same time; otherwise, one student had to wait for about an hour. As the students knew each other well, they were relaxed in a group, and the interviews were useful.

## 6. Results

### 6.1 Student uses of the text-based forum and the voice-based chat room

Altogether, 152 questionnaires were distributed in the classroom, and 124 questionnaires were collected. 16 were invalid for various reasons. Of the valid, submitted questionnaires, 84 were from females, and 24 were from males. They were all Chinese students living all over the country. There were three age groups: 20 students were 25 years old or less, making 19% of the cohort in the year of 2011, 70 were 26-35, making 65%, and 18 were 36-45, making 16%. The age range was wider in an online school than in a normal regular university. When comparing the frequency of student access to the voice-based chat room and the text-based forum, the researcher found that students accessed the latter more than they did the former (Table 3).

**Table 3.** Student access to the voice-based chat room and the text-based forum

Item	often	percentage	seldom	percentage
text	57	52.8%	51	47.2%
voice	21	19.4%	87	80.6%

As can be seen in Table 3, more than half of the students often accessed the forum, whereas only 19.4% of students often used the voice-based chat room. That is, 80.6% of students seldom used the voice-based chat room, and this was very high. There was also a high percentage of students who seldom used the text-based forum. This may well explain why, in the text-based forum and in the voice-based chat room, only a few students were active, with most students being passive bystanders. In the text-based forum, they were many lingerers who merely browsed the forum and may never have posted. In the voice-based forum, active students not only attended teacher-led sessions but also initiated their own sessions; however, most students chose to watch and never spoke. In the interviews later, students explained their difficulties in accessing the chat room in terms of the technology and the need for a higher computer specification, and they also explained why most of the students were passive in the text-based forum.

### 6.2 Student perceptions of social presence in the text-based forum and the voice-based chat room

When comparing the mean score of student perception in using the two technological

tools for interaction with peers, it was found that they had a slightly higher opinion of the voice-based chat room. However, when a paired t-test was completed, it was found that there was no significant difference between students' perceptions of the two technological forms as media to communicate with peers ( $M = -0.1111$ ,  $SD = 0.702$ ,  $t = -0.646$ ,  $df = 107$ ,  $p = 0.103$ ). With regard to the perception of fellow students there was a significant difference between the two technological forms ( $M = 0.194$ ,  $SD = 0.803$ ,  $t = 2.518$ ,  $df = 107$ ,  $p = 0.013 < 0.001$ , the mean score for the text-based forum was 4.22 and the mean score was 4.03 for the voice-based chat room). Although both tools were considered positive in discerning the thoughts of the fellow participants, the text-based forum obtained a higher rating in this aspect. Concerning the making of friends in the two technological forms there was no significant difference ( $M = 0.148$ ,  $SD = 0.955$ ,  $t = 1.612$ ,  $df = 107$ ,  $p = 0.110$ ). The mean score revealed that the text-based forum was slightly more effective as perceived by students.

As for helping develop a sense of community, the two technological tools were both very useful (mean score: 4.28 and 4.14). The text-based forum seemed to have a higher rating, but not at a significant level ( $M = 0.139$ ,  $SD = 0.779$ ,  $t = 1.853$ ,  $df = 107$ ,  $p = 0.067$ ). Both the text-based forum and the voice chat room functioned well in helping students develop their sense of community. Most of the students studied on their own and did not even attend the face-to-face tutorials on campus, although it was part of a service for them from the online school. The problem was that their first priority was their work, and they did not have time for the tutorials. This increased the need for online contact, and the two technological tools could serve as a link to connect them. They felt that with the text-based forum and the voice-based chat room, they were included in a group and had a sense of community.

It was interesting to find that students believed that they learned more with the voice-based chat room ( $M = -0.167$ ,  $SD = 0.755$ ,  $t = -2.294$ ,  $df = 107$ ,  $p = 0.024 < 0.001$ ). The mean score revealed that students were quite sure they had learned more effectively in the voice-based chat room, and the difference was statistically significant.

### **6.3 Projected social presence in the text-based forum and the voice-based chat room**

In general, the study showed that the voice-based interaction had a slightly higher projected social presence, but not at a statistically significant level, than the text-based interaction. In order to show a clear picture of what the statistics actually present, a Mann-Whitney U Test, a nonparametric test, was used to compare text and audio data in terms of the separate categories of social presence.

When the two data sets were compared in terms of categories of social presence, they showed a detailed picture of each specific aspect of social presence. The density of affective and cohesive indicators was higher in the text data, whereas interactive indicators had a higher density in the audio data.

**Table 4.** Affective, interactive and cohesive categories in the text-based forum and in the voice-based chat room

Item	Categories	Affective	Interactive	Cohesive
Text	No. of categories	57	123	45
	Density	12.23	26.41	9.66
Audio	No. of categories	77	508	114
	Density	5.42	35.78	8.02

When the interactive category was compared using SPSS, the social presence density scores in the audio data were shown to be significantly higher than in the text data (Mann-Whitney  $U = 12$ ,  $Z = -4.182$  and  $p = 0.000$ ). In terms of the cohesive category, the text data were significantly higher than the audio data (Mann-Whitney  $U = 55.5$ ,  $Z = -2.415$  and  $p = 0.016$ ). When affective categories were compared, it was found that the text and audio data were not significantly different (Mann-Whitney  $U = 91.5$ ,  $Z = -.910$  and  $p$  value is above 0.05). Interestingly, the density of affective categories appeared to be two times higher in the text data than in the voice data (see Table 4), but when all the files were compared, these categories were found not to be significantly different.

#### 6.4 Language practice differences in the text-based forum and in the voice-based chat room

As the affective and cohesive features are strong in the text-based forum, it is natural to see that students feel warm there and find it a comfortable place for the online community. More students access the text-based forum for interaction. But because of the strong interactive features of the voice-based chat room, it is a better place for language practice. The language output is obviously much larger in the chat room. From the text-based forum, 96 postings with 4,754 words were produced in the whole term. In contrast, the data from the chat room demonstrated much more language practice. Three sessions in less than three hours came up with 14,197 words. Moreover, in language practice, students speak more English (41.2%) in the voice-based chat room. One of the interviewees shared this view, saying, “In the voice-based chat room, you can hear the voice of the teacher and classmates, you feel closer to them. We use them to practice our spoken English.”

## 7. Discussion and Implications

### 7.1 Observing social presence from various angles

From the questionnaire survey, as it was found that the general situation was that most of the students were not active either in the text-based forum or the voice-based chat room. But the

text messages were more easily accessed, and the voice messages were “hidden”. This might explain that students felt that it was easier for them to make friends and learn about other students’ ideas. Therefore, most students felt that the text-based forum provided a place to bring them a sense of belonging. One student in the interview told an interesting story that when she studied the course, she opened the text-based forum, not because she wanted any information from there, but just because she had a feeling that she had her fellow students around in the online community.

But those students who had voice-based chat room activities claimed that they had in-depth interaction with their peers. From the text analysis of their conversations in the voice-based chat room, it was true that they talked about personal, emotional and sensitive topics and when they started to talk about these topics they had many turns of conversation so that they could go in depth to what they had talked about.

#### Example 1

- 3 A: So how’s your homework? Hehe.  
 4 B: Hehe. Well done. Well done.  
 5 A: Oh.  
 6 B: 亲爱的, 你报名了吗?  
     “Darling, have you got registered?”  
 7 A: 什么东西?  
     “What is it?”  
 8 B: 考试, 预约考试。  
     “The exam. The registration for the final exam.”  
 9 A: Yeah, I did that.  
 10 B: 你预约了吗?  
     “And you?”  
 11 A: Yeah.  
 12 B: 我还没有报呢?  
     “No, I haven’t.”

In this voice-based conversation, learners were talking about their final exam and assignment, which were part of the course assessment. This was quite a sensitive topic, and for sensitive questions, participants needed an immediate answer. In the asynchronous text chat, it was difficult to talk about such issues. This has been studied in other studies using synchronous interaction for emotional support (Motteram, 2001; Xie, 2008). When participants are emotionally involved in the mediated interaction, they need immediate responses. In such a situation, synchronous voice-based interaction was a suitable choice to meet their needs. In contrast, asynchronous text-based interaction was slower paced and did not satisfy such a need.

When the text-based interaction is compared, because it is asynchronous, the advantage is that it gives students time to think, and thus, the complexity of information may be

increased in the interaction at one posting (Swan & Shih, 2005). However, the tempo is slowed down, and two-way communication reduced. It is also true that when a reply is delayed, the participant may lose interest or patience to continue. When all the text files were combined, the average number of postings for a thread was 6.4. If the longest thread with 35 postings was not included (that one was especially long because it was an old posting lasting for almost a year), the average number of postings was even less. However, in the voice-based chat room, when one topic was initiated, or a question asked, five or six turns of interaction was normal. It means that discussion is likely to go into depth.

### **Example 2**

- 1 A: By the way, I really appreciate the material you offer us. Really, it's fantastic.
- 2 B: Is it funny? Hehe. Do you think it's funny that the marshal talks with his twin daughters?  
Hehe.
- 3 A: Scene Two.
- 4 B: Yeah, Scene Two.
- 5 A: Yeah.
- 6 B: Scene One...
- 7 A: Your imagination is magnificent.
- 8 B: Without boundary.

This conversation was taken from a practice activity after the students learned an American story, "Hearts and Hands," written by O. Henry. In the story, a police officer was taking a criminal from Denver to Boston. When they were on a train, they met a lady who knew the criminal. The police officer pretended that he was the criminal before the lady and gave the criminal a chance to have a happy talk. Our assignment for students was:

Instructions: Here is a complete short story, "Hearts and Hands," written by O. Henry. Read it and answer the following questions. Write your answers on the Answer Sheet.

1. Imagine that you are the Marshal and retell the story from his point of view.
2. You should keep the basic content of the story.
3. Your writing should be approximately 200 words.

Student B had twin daughters and wrote an imagined story that the police officer had twin daughters, and when they heard the story that their father had pretended to be a criminal on the train, they asked questions of their father, trying to understand what was in the father's mind for the action. Student B was deeply involved in the story and shared her imagined story with Student A. They started to talk about it in this voice-based activity. Here from this activity, it was found that when students participated in the activity, they could have many turns of conversation about some specific questions or topics. They may have talked in depth about them. They had prepared in advance and talked at length in discussion. This may be connected with the issue in the questionnaire that students believed that voice-based chat room helped

them learn better.

In contrast, messages from the text-based forum showed that topics and questions were more general than specific. In Example 3, one student posted a question to those who had learned the course for some advice. The student was wondering if he should watch films or read literary works in order to understand the selected literary readings in their textbook. For this question, there was an answer from one of the previous students. She seemed to care about the final exam of this course, and she encouraged this student to read more in order to achieve a better score later in the exam, which was practical.

### Example 3

III. 学本课程，是看电影呢还是读原著呢？

“To take this course, what should I do, to watch related films or read original literary works?”

1 有哪位学哥学姐在百忙中指导一二呢？先谢了。

“Can any elder brother or sister from this course spare me some time to give suggestions for the above question? Thank you in advance.”

2 我觉得我在期末考试小说改写那道题失分可能很多，总结一下，可能是平时阅读量和写小说的量都太少。

“I think I lost many points in the story rewriting in the final exam. To sum it up, I haven't read enough and written little in term time.”

In the interviews, one student talked about how the text messages had lost authenticity and spontaneity of conversations in the mediated interaction, whereas the voice-based interaction seemed to be more engaging for the parties involved. She said:

[In the text-based interaction] you seem to avoid something and you don't speak out frankly. ...Of course, speaking is more direct than writing. You have your way of speaking and prosodic features. You have more freedom to express yourself. I personally think that speaking can be more real than writing.

## 7.2 The need to expand social presence and implications for language learning

From other interviews, students have described that prosodic features of the voice-based interaction provided them with rich information about their fellow learners, and thus, it led to higher peer awareness of individual students in the voice-based chat room, which encouraged more two-way interaction than in the text-based forum. The high peer awareness encouraged learners in the interaction to concentrate more on communication and spend more effort trying to meet the communication needs. In contrast, in the text-based forum, peer awareness was geared to a group of learners, and thus sharing information was common, but sharing emotion was less common.

When the text-based forum and the voice-based chat room are both used in the online language learning course, the former is more frequently accessed by learners and perceived

to be a more effective medium to understand peer learners and a better place to make friends and to form a learning community. As the text-based forum is one-to-many interaction and more easily accessed, it provided the efficiency of reaching fellow learners in a short time. But learners had a slightly higher view of the voice-based chat room as an effective medium for communication, and they believed that it helped them more with their learning. As the voice-based communication was synchronous, and participants spoke instead of writing, the tempo of interaction was faster. In such a situation, they tended to talk about emotional issues, which required immediate responses.

## 8. Conclusion and Limitation

The text-based forum and the voice-based chat room had a different impact on student social presence. In terms of student perception, most of them were more likely to get to know peers in the text-based forum and thus developed a sense of community in their learning process of the online course. Yet they believed that the voice-based chat room had the advantage of helping them with language learning. In the actual interaction, it was found that the voice-based chat room was more interactive, although the text-based forum was more affective and cohesive. But the problem with the existing framework in literature was that there were no prosodic features included in the affective indicators in the social presence category. Therefore in future more research is needed to search for the relationship between prosodic sound features and social presence, and the present theoretic framework requires to be extended. In interviews, students explained that in the voice-based chat room prosodic features led to higher peer awareness, which further reinforces this need.

The limitation of the study is the comparison of the text-based forum and the voice-based chat room with regard to social presence in the live language learning context, which involves numerous elements at any point in the interaction process. It may raise doubt about the issue of comparing like with like. The text-based forum is asynchronous while the voice-based chat room is synchronous, and students used them for different tasks with different purposes. The research results may need further evidence from studies in different contexts, but the two popular media for online interaction are worth exploring for their affordances and constraints for practical use.

## References

- Borup, J., West, R., & Graham, C. (2012). Improving online social presence through asynchronous video. *Internet and Higher Education, 15*, 195-203.
- Dow, M. J. (2008). Implications of social presence for online learning: A case study of MLS students. *Journal of Education for Library and Information Science, 49*(4), 231-242.
- Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2*(2/3), 87-105.
- Gunawardena, C. N., & Zittle, F. (1997). Social presence as a predictor of satisfaction within a computer

- mediated conferencing environment. *American Journal of Distance Education*, 11(3), 8-26.
- Homer, B. D., Plass, J. L., & Blake, L. (2008). The effects of video on cognitive load and social presence in multimedia-learning. *Computers in Human Behavior*, 24, 786-797.
- Kehrwald, B. (2010). Being online: Social presence as subjectivity in online learning. *London Review of Education*, 8(1), 39-50.
- King, K., & Ellis, T. J. (2009). Comparison of social presence in voice-based and text-based asynchronous computer conference. *Proceedings of the 42nd Hawaii International Conference on System Sciences – 2009*, pp 1-10.
- Ko, C. (2012). Can synchronous computer-mediated communication (CMC) help beginning-level foreign language learners speak? *Computer Assisted Language Learning*, 25(3), 217-236.
- Ko, C. (2016). The effect of task types on foreign language learners' social presence in synchronous computer mediated communication (SCMC). *The JALT CALL Journal*, 12(2), 103-122.
- LaPointe, D., Greysen, K., & Barrett, K. (2004). Using synchronous audio for ESL teaching in Taiwan. *International Review of Research in Open and Distance Learning*, 5(1), 1-6.
- McIntosh, S., Braul, B., & Chao, T. (2003). A case study in asynchronous voice conference for language instruction. *Education Media International*, 40(1/2), 63-75.
- Motteram, G. (2001). The role of synchronous communication in fully distance education. *Australian Journal of Educational Technology*, 17(2), 131-149.
- Paulus, T. (2007). CMC modes for learning tasks at a distance. *Journal of Computer-Mediated Communication*, 12(4), 1322-1345.
- Rourke, L., Anderson, T., Archer, W., & Garrison, D. R. (2001). Assessing social presence in asynchronous, text-based computer conferencing. *Journal of Distance Education*, 14(3), 51-70.
- Satar, H. M. (2015). Sustaining multimodal language learner interactions online. *Calico Journal*, 32(3), 480-507.
- Satar, H. M., & Akcan, S. (2018). Pre-service EFL teachers' online participation, interaction, and social presence. *Language Learning & Technology*, 22(1), 157-183.
- Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. London: John Wiley & Sons.
- Swan, K., & Shih, L. F. (2005). On the nature and development of social presence in online course discussions. *Journal of Asynchronous Learning Networks*, 9(3), 115-136.
- Terhune, N. M. (2016). Language learning going global: Linking teachers and learners via commercial Skype-based CMC. *Computer-Assisted Language Learning*, 29(6), 1071-1089.
- Traphagan, T., Chiang, Y., Chang, H., Wattanawaha, B., Lee, H., Mayrath, M., ... & Resta, P. (2010). Cognitive, social and teaching presence in a virtual world and a text chat. *Computer & Education*, 55, 923-936.
- Tu, C. H. (2002b). The impact of text-based CMC on online social presence. *The Journal of Interactive Online Learning*, 1(2), 1-24.
- White, C. (2009). Towards a learner-based theory of distance language learning: The concept of learner-context interface. In P. Hubbard (Ed.) *Computer assisted language learning: Critical concepts in linguistics, vol. IV: Present trends and future directions in CALL* (pp. 97-112). London: Routledge.
- Whiteside, A. L. (2015). Introducing the social presence model to explore online and blended learning



experiences. *Online Learning Journal*, 19(2). Retrieved from <http://olj.onlinelearningconsortium.org/index.php/olj/article/view/453/137>.

Wu, H., Gao, J., & Zhang, W. (2014). Chinese EFL teachers' social interaction and socio-cognitive presence in synchronous computer-mediated communication. *Language Learning & Technology*, 18(3), 228-254. <https://dx.doi.org/10125/44392>.

Xie, B. (2008). Multimodal computer-mediated communication and social support among older Chinese internet users. *Journal of Computer-Mediated Communication*, 13, 728-750.

## About the author

Ligao WU is an associate professor at the Institute of Online Education, Beijing Foreign Studies University. His research efforts focus on computer-assisted language learning. Email: [wuligao@bfsu.edu.cn](mailto:wuligao@bfsu.edu.cn)