Online Chatting as a Tool for Language Learning: the Case of Milao, for Beginner Spanish

Tina Chan, Jhotisha Mugon and Camelia Nunez[*]
University of Waterloo
tk3chan@uwaterloo.ca
jmugon@uwaterloo.ca
*cnunez@uwaterloo.ca

Abstract: Online chatting has become a primary tool for communication in our society: Facebook chat, google chat and others alike are fast and comfortable ways for sharing information and have become preferred ways of communication for most. Post-secondary students are no exception to this new societal trend. The current paper reports the preliminary results of an ongoing study assessing online chatting as a tool for learning a foreign language among a group of beginner Spanish students. The students were using Milao, an innovative online tool that allows learners of Spanish to chat back and forth with an artificially intelligent avatar that acts as a native chat buddy. It can engage in conversations that reflect real-life situations just as a real person would. A group of beginner Spanish students at a Canadian post-secondary institution engaged with the tool and the data collected offers valuable preliminary insight into the language learning process. The results presented in this paper are a first look into the "learning-through-chatting-process" and allow for interesting preliminary conclusions and further research directions. Primarily, we have learned that in chat-based interactions, comprehension is not an issue. Rather, learners struggle with formulating grammatically accurate responses.

Keywords: language learning, Spanish, chatting, natural language processing, technology

1. Introduction

Anyone who has tried it, knows that learning a foreign language is not an easy task and that many challenges need to be overcome in the process. As a result, most often many of those attempting to learn, give up before achieving any success.

Before going into further detail, however, it is important to establish what the goal of learning a foreign language is considered to be. Despite the great deal of debate in the field of language teaching and research, most people would agree that the aim of teaching and learning a second language should be communicative competence (Canale & Swain 1980). That is, the learner should acquire not only linguistic competence and knowledge of the grammar and lexicon, but also the ability to use language in real life settings in order to communicate. However, in

---

[*] Principal investigator and corresponding author.

the foreign language context, where the language being learned is not the language of the community, it is often very difficult to achieve anything close to communicative competence. Contact hours with the foreign language are reduced to a few hours a week, often the only contact with a native speaker is the teacher, classes are relatively large, and there are few opportunities to use the language. Many textbook publishers try to complement this poverty with additional work on-line. However, this kind of interaction is mechanical, repetitive and most importantly, it is as static as the traditional textbook.

The question is then: how to provide learners with opportunities to develop these communicative skills they are aiming for in a foreign language? This question has been tackled in many different ways and many attempts have been made – few of them successful, however. Long (1996) found that learners become uninvolved and very soon give up on the task of learning a foreign language. Thus, a new question arises: how to maintain learners' interest in order to help them become fluent communicators? The obvious response would be: by creating interaction opportunities through which they can develop the communicative skills they are seeking. Creating interaction opportunities, however, has proven in itself to be quite a challenge since most often the time spent in a language classroom at the post-secondary level is limited to three or four hours a week. Outside of the classroom, interaction opportunities are even scarcer as learners do not often know other speakers of the language; and even if they do, scheduling time to practice can become quite an inconvenience.

Subsequently, many have tried to create these opportunities through the use of technology. Looking at today's young generation, it is evident that technology is highly engrained in their daily routines. An even closer look at post-secondary learners in particular, will immediately further reveal that their preferred and most natural way of communicating amongst each other is in fact through technology (Thayer & Ray 2006, Quam- Haase 2008). Students in every university classroom keep in touch through email, Facebook chat, texting, Gmail chat and others of the sort; instead of talking on their smart phones, they in fact use these gadgets to communicate in writing through the above mentioned

methods. Previous research has also shown that there is a positive change in second language learners' performance when they incorporated the use of mobile application for learning new vocabulary in their course (Rezaei et al. 2014). Given the circumstances described in this section, the natural solution to this problem would be to adapt such communication methods to the process of language learning and create interaction opportunities for language learning. Milao, the technology described in the current paper aims to do just that.

2. Literature Review

Previous research has shown that the use of different kinds of technology for language learning can facilitate communication, instill more confidence by reducing anxiety and finally, it can develop the writing-thinking connection (Abrams 2002, Salaberry 2001, Weininger & Shield 2003 among others). As for writing skills in a foreign language, research has shown they can be significantly improved through the use of technology and computers. For example, studies done by Hertel (2003), LeLoup (1997) and Warschauer (1995) demonstrate that foreign language teachers who integrated email activities into their curriculum, noticed significant advancements in learners' written communication skills compared to those who did not integrate email.

Along the same lines, Lee (2002) completed a study using synchronous chats combined with task-based instructions in order to help develop learners' communication skills. Lee concluded that chat-based communication using more open-ended exchanges had a positive impact on the language learning process. Furthermore, Lunde (1990) studied email as a medium for language learning among Japanese students and concluded that said method significantly improved learners' reading comprehension skills. A follow-up study by Beauvois (1994) reported that because of these initially written communication activities, the students had developed an increased confidence in speaking. It was concluded that students had routinized certain expressions and therefore developed automated structures that were very helpful in oral production.

In their 2008 paper, Ruan & Wang described a language learning system that incorporated three applications for L2 English learners. The first, web-enabled mobile phones, provided speaking practice by having students record responses to question prompts. The second application promoted vocabulary learning by sending a word or expression via SMS during class time. Students then used these in class conversations or a web-based text chat application, which they accessed on laptops in class. The authors concluded this method to have been highly beneficial in the language learning process.

Another field of technologies for language learning that is related to the current study is that of Human Language Technologies (HLT) for Computer Assisted Language Learning (CALL). Matthews (1993) points out a close relationship between this field, Linguistic Theory and Second Language Acquisition Theory. Gupta & Schultze (2010) show that progress in HLT was made possible in great part by the insights drawn from linguistics and language acquisition research. Skeptics of Human Language Technologies, however, argue that "*Natural Language Processing (NLP) Programs cannot account for the full complexity of natural human languages*" (Salaberry 1996: 54). It should thus be noted that the claim of Milao, the technology described in this paper, is not to be able to replicate something as complex as human language in its entirety. Rather, the objective of computer applications such as this one is to focus on specific aspects of a given language. With the help of human language technologies such as dialogue managers, the computer can now offer natural human-like interactions and thus provide learners with the necessary input and communicative opportunities. Such technologies are known as Natural Language Processing (NPL), a field that sits at a crossroads between linguistics and computer science, exploring the interaction between human language and computers.

Milao is powered by NLP technologies. This online platform, which allows language learners to engage in text-based interactions with a virtual language partner/tutor and gain feedback on their performance, was created by a team of researchers (including one of the authors of this paper) in the field of second language acquisition and natural language processing. It is meant to act as

an online chat buddy, available anywhere and anytime. Learners can access it simply by creating an online account using their email to sign up. The tool will also be available on mobile phones and tablets in the form of a downloadable application.

3. Current Study

The very first time post-secondary students used Milao as part of their language class was labeled as a beta-test. The purpose of this beta-test was to introduce students to a new learning interface that gives them the opportunity to put into practice vocabulary and grammatical rules learned in class. At the same time, it was an opportunity for the researchers conducting the test to observe the interaction and learn more about the shortfalls of the technology. Students engaged in chats with an avatar that – just as a 'real' Spanish-speaking friend – chatted with them on familiar topics covered in class such as the family, at a restaurant etc. The avatar asked and answered questions just as a normal person would and by the same token, the students had to ensure they were able to write sentences that contributed meaning to the conversation and were understood by the avatar. The study provided valuable information on three main areas: (1) learner comprehension of the question they were asked, (2) learner ability to answer the question within context and (3) grammatical accuracy of learner answers. Our working hypothesis was that learners would mostly comprehend what they were being asked, but would struggle in giving a grammatically accurate answer.

4. Methodology

4.1 Participants

Participants included 16 undergraduate students who were taking introductory Spanish in a post-secondary institution for the first time. None of these learners had been exposed to formal training in the Spanish language before. They were completing the chats during the scheduled weekly one-hour language labs. Each chat topic was aligned with the topics students were covering in class, meaning

that all of them would have already been familiar with the grammar and vocabulary before engaging in the chatting with the avatar.

4.2   Measures

During the lab sessions, students chatted on a variety of topics from basic introductions to discussing their studies, classes they were taking at the university, to family members. All of these topics were familiar to the students, as they had covered the pertinent vocabulary and grammar concepts during class time with their instructors. The sole purpose of the chats was to simply put into practice what they were learning through interactions.

As mentioned earlier, three measures were looked at during these chats: (1) comprehension, (2) ability to answer in context and (3) grammatical accuracy of the answer. Comprehension and ability to answer in context was measured by looking at whether the students understood the question asked and whether he or she was able to provide an answer within the context of the question. Example (1) below clearly shows that the student in question did not understand the question and therefore the provided answer is out of context as well; instead of answering with the number of classes the student was taking, the answer was about what classes this student was enrolled in.

(1)   **Avatar:**   ¿Cuántas clases tomas este semestre?
                    (How many courses are you taking this term?)

         Student: Tomo arte y español.
                    (I am taking Art and Spanish.)

        Grammatical accuracy was measured by looking at the grammatical errors made by each student to each question within each chat. For the purpose of this beta-test, it was assumed that grammatical mistakes were made when students used the wrong verb conjugation or word order or the wrong prepositions. Therefore, if students gave the correct verb conjugation, preposition and/or used the correct word order, their answer was considered to be grammatically correct.

In (2) we give an example of a grammatically inaccurate answer provided by the student to one of the questions asked by the avatar.

(2)   **Avatar:** ¿Caminas a clase?
            (Do you walk to class?)

    Student: *Sí, caminas a clase.
            (Yes, (you) walk to class.)

In the case of an outside context answer or a grammatically inaccurate answer provided by the student, the system would repeat the question by rephrasing it and therefore give learners the opportunity to correct themselves. Each question was programmed to be repeated for up to 3 times, after which the avatar would move onto the next question even if the student continued to provide a grammatically incorrect answer. Consequently, the system produced variations of the same questions simply to give learners the chance to observe and learn from their mistakes. The repetition of the questions acted as a clue to students to let them know that there is something wrong in their answer. Thus, they would be able to correct their answer.

4.3   Procedure

As previously mentioned, students were tested during their Spanish laboratory sessions and after having created an account on the Milao website. Students completed each chat after having covered the material in class with their professors. All students already had an overview of the topic prior to starting their chats and each conversation built on previously learned material.

Each chapter covered in class had a related conversation topic on Milao and within each topic there were three scenarios in which students could take part. Topics were made available to students one at a time and in the order they were being covered in class. This ensured that students would not be able to engage in chats beyond their level.

4.4   Data Analysis

Firstly, each scenario was analyzed to see whether students gave a within-context answer to each question. For example, if a specific scenario was programmed with 6 questions and a student provided an outside context answer to one of the question, then the total number of questions asked in that chat would be 7 and the total number of in-context answers would be 6.

Secondly, each scenario was analyzed for grammatically incorrect answers. This included incorrect verb conjugations and prepositions and the wrong word order. For example, if a student wrote a within context but grammatically wrong answer in a chat programmed with 6 questions, then the student would receive a score of 6 out of 7 for grammatical accuracy for that specific chat (the online avatar would have repeated the questions for which the student gave the grammatically incorrect answer which would increase the total number of actual questions asked to 7).

One-Way ANOVA was performed for each chat to determine if there was a significant difference between the total number of questions asked, the number of within context answers and the number of grammatically correct answers. The number of actual questions asked included the number of programmed questions and the number of repeated questions. The numbers of context appropriate answers and grammatically accurate answers were derived from the total number of questions asked.

5.   Results

In total we had 6 different chats to be tested by students. Chat 1 was programmed with very basic informal greetings, chat 2 was about formal greetings and introductions, chat 3 asked the students about who they are and their country of origin, chat 4 inquired about students' university and their general study habit, chat 5 targeted the types of courses students were taking this term and their favourite class and teacher and finally, chat 6 was related to their actual Spanish class and the logistics of their class (i.e. how many students were in their class). Due to technical difficulties/ programming error/ network error, all the answers to

chat 2 were not recorded properly. To be consistent in our analysis of comprehension, in-context answers and grammatical accuracy, chat 2 was excluded from further analyses. Figure 1 shows the number of programmed questions versus the number of actual questions asked. As it can be seen, more questions were asked per chat than programmed. This made us wonder what was causing that to happen. Why was the system repeating or rephrasing the questions so often? Were students not understanding the questions, or were they providing ungrammatical answers?

Figure 1
Number of programmed questions vs. number of actual questions asked for each chat
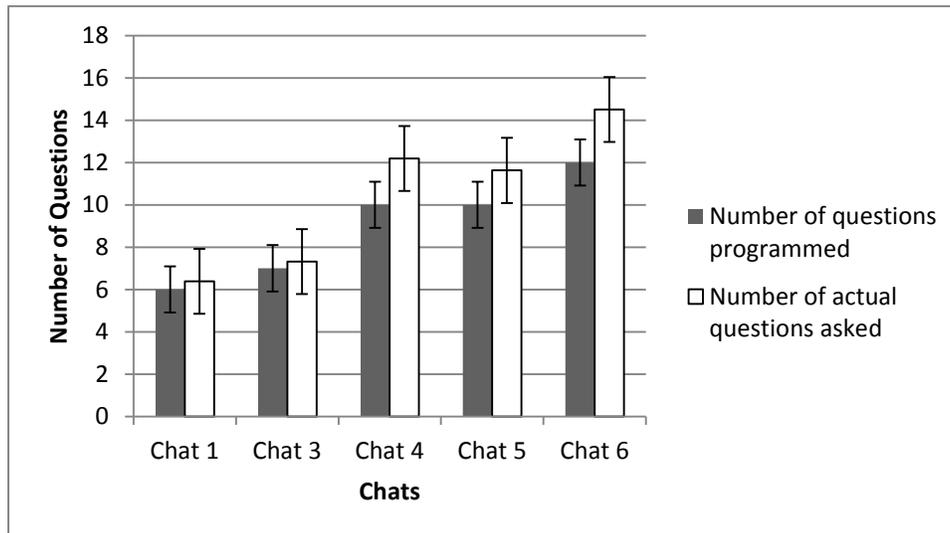

An individual analysis of each chat revealed that students were able to provide context appropriate answers more than they were able to provide grammatically correct answers. This means that while they did understand the questions they were being asked and were attempting to provide answers within context, their answers were not grammatically correct. Figure 2 shows a comparison between the number of actual questions asked vs. the number of context appropriate answers vs. the number of grammatically correct answers.

It is clear from Figure 2 that while most answers provided were in context, many of them were not grammatically correct. The only exception is for chat 1.

No statistical significant differences were found between the number of questions asked, the number of within- context answers and the number of grammatically correct answers in chat 1. This may be due to the fact that chat 1 consisted of very basic informal greetings which students might have already been exposed to previously. Nevertheless, using One-Way Anova, statistical differences were found between context appropriate answers and grammatically appropriate answers, as well as total questions asked and grammatically appropriate answers within chats 3-6.

Figure 2
Comprehension, in-context answers vs. grammatical accuracy within each chat
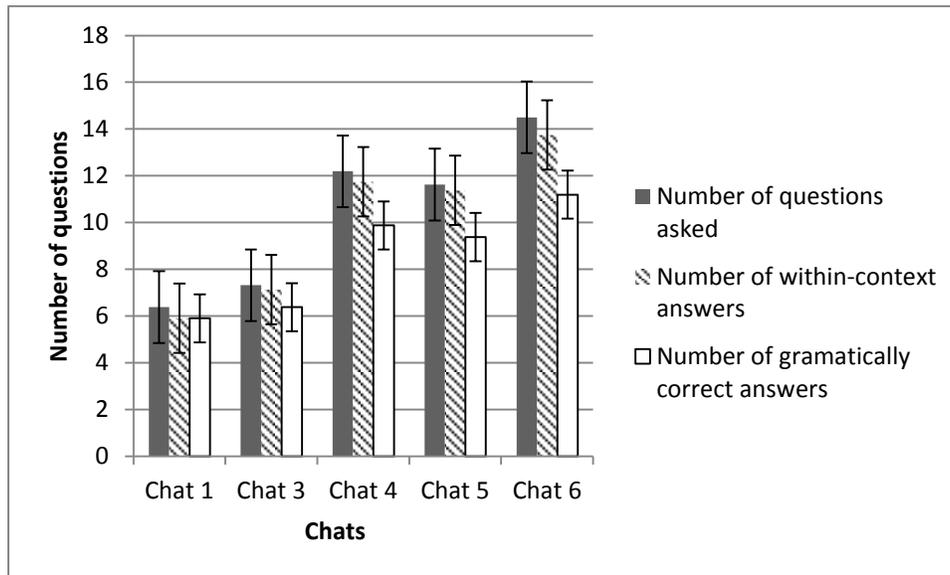
For chats 3-6, there was statistical difference between the number of questions asked, the number of within-context answers and the number of grammatically correct answers [respectively: $F(2, 45)= 6.58$, p= .003; $F(2, 45)= 5.62$, p= .007; $F(2, 45)= 9.18$, p= .00; $F(2, 45)= 14.16$, $p = .00$]. Multiple comparisons revealed that more questions were asked than grammatically appropriate answers were given for chats 3-6 ($p \leq .008$). Furthermore, for chats 3-6 more within context answers were provided than grammatically appropriate answers ($p < .05$).

Moreover, Tukey post-hoc test from chat 3-6 revealed that there was no significant difference between total questions asked and context appropriate answers in each of the chats respectively. This indicates that while students understood the questions they were being asked and while they tried to provide a within context answer, their answers were often grammatically incorrect. Figure 3 depicts the trend between programmed questions, actual number of questions asked, context appropriate answers and, grammatically incorrect answers for all chats in order to summarize all the results.

Figure 3
Mean responses from each chat at each categorical level
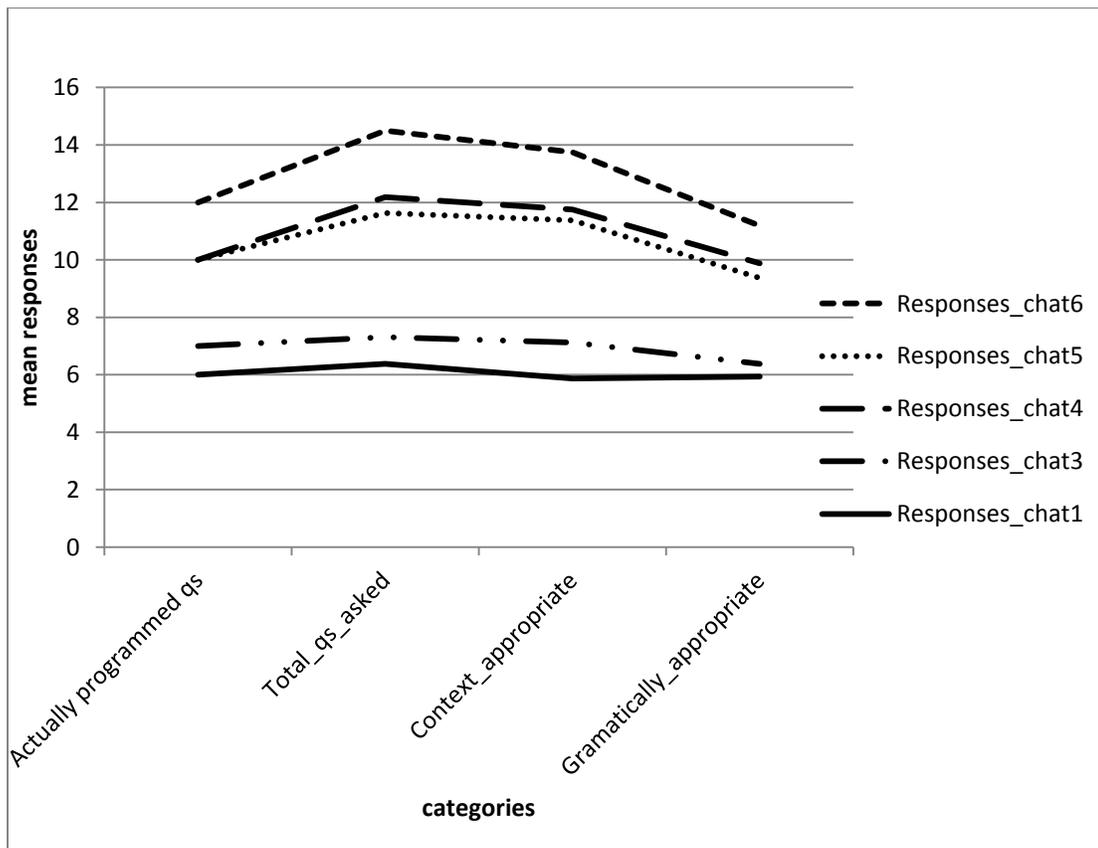
6. Discussion and Conclusions

In the introductory section of the current paper, we discussed the goal of learning a foreign language to be achieving communicative competence or the capability

to communicate in real life situations in the target language. Nonetheless, based on research on the topic up to this point, most studies concluded that communicative competence is rarely attained in second language learning. Lack of interaction or opportunities are in great part at fault for such low success rates in second language acquisition. Gass & Mackey (2007) conclude that "the interaction approach attempts to account for learning through the learner's exposure to language, production of language and feedback on that production". Long (1996) explains the relationship between interaction and learning as follows:

> […] negotiation for meaning, and especially negotiation work that triggers interactional adjustments by the native speaker (NS) or more competent interlocutor, facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive way. (pp. 451-452).

In other words, what Long implies is that through interaction, learners' attention will be directed to elements of the language that are problematic, be it in knowledge or in production. By engaging in a communicative task with a more competent speaker, the learner will discover elements in their interlocutor's speech that they are not aware of and that he or she may not understand; such elements may be new lexical items or grammatical constructions.

Nonetheless, the reality is that interaction opportunities with native speakers or other speakers of a foreign language (that is not that of the community) are scarce and therefore the majority of learners do not have the opportunity to develop any communicative skills.

As shown in this paper however, by using NLP technologies, Milao was created as an online chatting tool where learners can chat with an artificially intelligent chatting buddy about everyday real life scenarios. To remedy the lack of interaction with native speakers, we created a tool that would help learners to put into practice grammar and vocabulary rules through meaningful conversation. After all, both oral and written communication play an important role in the development of communication skills. Moreover, most student testing continues to take place in a written form rather than oral. A tool such as Milao, is an

opportunity for them to improve on this testing as well. Previous studies such as Beauvois (1994) showed that initial written practice led to increased confidence that better prepared learners for oral production. Thus, the Beauvois study has already demonstrated the benefits of written interaction.

The results of the current study supplement existing work and offer further evidence that writing can facilitate the development of oral skills. The data collected has allowed us to gain a better understanding on not only communicative competence, but also grammatical accuracy. We have shown that even though learners may have developed comprehension skills, this does not necessarily mean they demonstrate grammatical competence in production.

The goal of the Milao technology is to provide meaningful interaction opportunities in situations that closely mimic real life situations. The results presented in this paper give us a great initial insight on the learning process through chatting technologies, however further data analysis is needed in order to make any conclusive claims. A further feature of the Milao chat will be an evaluation component where learners can get immediate feedback on their performance throughout a chat. This feature will allow learners to consistently improve on their interactions and at the same time, will allow us to get a better understanding of the exact grammatical mistakes learners make. Subsequently, we will be able to provide more specific conclusions on the learner grammatical errors.

References

Abrams, Zsuzsanna. (2002). Surfing to cross-cultural awareness: using Internet-mediated projects to explore cultural stereotypes. *Foreign Language Annals* 35(2): 141-160.

Beauvois, Margaret Haley. (1994). E-talk: attitudes and motivation in computer-assisted classroom discussion. *Computers and the Humanities* 28(3): 177-190

Canale, Michael & Merrill Swain (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1(1): 1-47.

Gass, Susan & Alison Mackey (2007). Input, Interaction and Output in Second Language Acquisition. In Vanpatten, Bill & Jessica Williams (eds.) *Theories in Second Language Acquisition: An Introduction.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc., pp. 175-200.

Gupta, Piklu & Mathias Schultze (2010). Human Language Technologies (HLT). Module 3.5. In Davies, Graham (ed.) *Information and Communications Technology for Language Teachers (ICT4LT).* Slough, Thames Valley University [Online]. Available from: http://www.ict4lt.org/en/en_mod3-5.htm [Accessed 10 December, 2010].

Hertel, Tammy (2003). Using an e-mail exchange to promote cultural learning. *Foreign Language Annals* 36(3): 386–396.

Lee, Lina (2002). Enhancing learners' communication skills through synchronous electronic interaction and task-based instruction. *Foreign Language Annals* 35(1): 16-24.

LeLoup, Jean (1997). But I only have e-mail – what can I do? *Learning Languages* 2: 10–15

Long, Michael M. (1996). The Role of the Linguistic Environment in Second Language Acquisition. In Ritchie, William C. & Tej K. Bhatia (eds.) *Handbook of Second Language Acquisition*. San Diego: Academic Press, pp. 413-68.

Lunde, Ken (1990). Using electronic mail as a medium for foreign language study and instruction. *CALICO Journal* 7(3): 68–78.

Matthews, Clive (1993). Grammar frameworks in Intelligent CALL. *CALICO Journal* 11(1): 5-27.

Rezaei, Azadeh, Neo Mai, & Ahmad Pesaranghader (2014). The effect of mobile applications on English vocabulary acquisition. *Jurnal Teknologi* 68(2): 73-83

Ruan, GuiLong & JiuYun Wang (2008). Design of English learning system based on mobile technology. *Proceedings of International Conference on Computer Science and Software Engineering* (pp. 1062– 1065). Los Alamitos, CA: IEEE Computer Society. Retrievable from http://ieeexplore.ieee.org

Quan-Haase, Anabel (2008) University students' local and distant social ties: Using and integrating modes of communication on campus. *Information, Communication & Society* 10(5): 671-693.

Salaberry, Rafael (1996). A Theoretical foundation for development of pedagogical tasks in computer mediated communication. *CALICO Journal* 14(1): 5-34.

Salaberry, Rafael (2001). The use of technology for second language learning and teaching: a retrospective. *The Modern Language Journal* 85(1): 39-56.

Thayer, Stacy & Sukanya Ray (2006). Online Communication Preferences across Age, Gender, and Duration of Internet Use. *CyberPsychology and Behaviour* 9(4): 432-439.

Warschauer, Mark (1995). Comparing face-to-face and electronic communication in the second language classroom. *CALICO Journal* 13(2): 7–26.

Weininger, Markus & Lesley Shield (2003). Promoting oral production in a written channel: an investigation of learner language in MOO. *Computer Assisted Language Learning* 16(4): 329–349.