

Designing a Mobile Chinese Learning System with Speech Recognition for Foreign Students

Wei-Tung TANG^{a*} & Shwu-Ching YOUNG^b

^a*Institute of Information Systems and Applications , National TsingHua University, Taiwan*

^b*Institute of Learning Sciences, National TsingHua University, Taiwan*

*s100065531@m100.nthu.edu.tw

Abstract: This study aims to design and implement a micro-learning Chinese vocabulary pronunciation practice system on mobile device for international students from a university in northern Taiwan. Learning a foreign language is difficult, yet using a foreign language to initiate social interaction with native speakers is even harder. In this study, we propose a location-based contextual Chinese learning system which aims to aid the foreign learners to learn daily life vocabulary by repeating practicing vocabulary pronunciation with speech recognition functionality. An immediate feedback will be shown to the learners, allowing them to check their correctness level of vocabulary pronunciation. To ensure the learning takes place in real context, the language learning will be enhanced with location-based service which is provided by Facebook. Location information will be available for the learners to access local tourist attraction information by selecting from nearby point-of-interests via Facebook's check-in module. An experiment will be conducted to measure students' language learning performance as well as their language learning motivation. We expect to recruit international students whose Chinese proficiency is at entry level. Participants will be selected from a northern Taiwan university.

Keywords: Mobile assisted language learning, automatic speech recognition, micro-learning, text to speech.

1. Introduction

As a modernized and globalized country, Taiwan has attracted numbers of foreign visitors here for various visiting purposes. Although English is a second language and a large proportion of Taiwanese citizens do understand it, many foreign visitors still find it difficult to blend in Taiwanese culture if they know absolute nothing about Chinese. “Learning Mandarin is considered to be difficult, and acquiring a deep understanding of Chinese culture is thought to be near to impossible.” (McKeon, 2009).

The education system in Taiwan has launched programs to support foreign students to continue their higher education for a long. National TsingHua University (NTHU) is one of the schools that accommodate nearly 3% of the total student population. One main issue that foreign students face is how they are going to adapt to a Chinese-speaking environment without assistance. We have observed that a large percentage of NTHU foreign students spend most of their time with similar culture groups. A foreign student from Africa once remarked Taiwanese students as “*smart, diligent, but less willingly to speak up*”. Another student also made such comment: “*Sometimes I have to take courses offered in Chinese. I have to sit there, listening to the Chinese lecture which I have absolutely no idea. After the class, I have only a few places to go, namely library, cafeteria or dormitory. Finally, I come to realize NTHU is just a school that gets me through four years college education.*” These learners cannot find the opportunities to interact with others, and do not have many chances to revise and adjust their Chinese speaking (Xin, 2001).

Second language learners can be divided into two different types based on their learning motivation. For people who seek to conquer the “language divide” by gaining a dominating language and improve their socioeconomic mobility, although their determination seems stronger than any others, they are likely to be frustrated by the needed time to master a foreign language, resulting in loss of learning interests. On the other hand, for those who are driven by intellectual-challenge or desire to initiate sustainable social interaction with target culture members, the learning curve for them is balanced, but it will require longer learning period. These two types of learners is the origin of “language divide”, which refers to the extra effort Chinese as Second Language (CSL) learners have to make to cross the language gap. For both groups of second language learners, the challenge is to maintain critical motivation and spend time in learning the target language in order to complete demands from work or social life (Dörnyei and Ottó, 1998).

With the sustained development of information technology (ICT), mobile phones now are equipped with downloadable applications. Integrating ICT into teaching and learning processes can make learners fully engage in learning tasks. (Agostini, Di and Loregian, 2010). Through mobile assisted language learning (MALL), now language learning is no longer limited to fixed location such as classrooms. With language learning application on mobile phone, learning can occur anytime, anywhere. Learning that occurs at increasingly spaced intervals will help learners remember what they have seen before. Such repetition will help them to recall information stored in short-term memory. Learning is enhanced when learners recall the learned detail rather than restudying it (Pashler, Rohrer, Cepeda and Carpenter, 2008). In this study, we propose an automatic speech recognition (ASR) system along with a text to speech (TTS) module, providing a way to conduct repeated practice between different time and places. We seek to integrate TTS and ASR modules with location-based service supported by Facebook. We expect the learners will be able to undergo micro-learning activity which deals with small learning units or short-term learning activities (Kovachev, Cao, Klamka and Jarke, 2011), and learn about daily life vocabulary with the real-life experience.

2. Purpose of the present study

The first aim of this study is to connect international students to the Chinese learning material, providing them a way to connect their lives to the local culture by learning how to identify vocabulary and based on a contextual learning scenario. Next, it is to identify the correctness level of Chinese vocabulary pronunciation with the aid of the proposed system design. There will be an attempt to examine learners' learning style and trend. Finally, it is to evaluate learners' willingness level of utilizing newly acquired vocabulary. Three research questions, corresponding to the purposes of this study will be addressed: (1) Would CSL learners' level of learning motivation be raised with the aid of the proposed location-based learning design? (2) Are there any preferred types of learning styles which lead to better learning outcome? (3) Does the immediate practice feedback have positive impact on CSL learners' Chinese vocabulary daily usage?

3. Literature Review

3.1 Current Chinese Language Learning Trend

According to American Council on the Teaching of Foreign Languages (ACTFL), there were only 5000 American high school students who took Mandarin Chinese as their secondary language in 2000. Yet, in 2007, there were more than 50,000 students who took Mandarin Chinese course. We understand that Mandarin Chinese has gotten more attention than it had before. Having that said, the number of students who have ambition to master Mandarin Chinese as their second language is still climbing high. Yet, it remains a mystery for most Chinese learners whether they are able to successfully cross the “language divide” or not.

3.2 Vocabulary Learning Strategy

Related researches indicated that, the ability to recognize vocabulary is the most basic and urgent one for all beginners (Ho, 2008; Sun, Huang and Liu, 2011). This applies to international students who pursue their higher education in Taiwan as well. Upon arriving in Taiwan, their life will be closely connected to this Chinese, sometimes Taiwanese speaking environment. If they do not manage to recognize some Chinese vocabulary, it is likely they will face a very difficult time during their stay. Moreover, recognizing the vocabulary is one thing, knowing how and when to use them is another, especially when most of Taiwanese feel comfortable speaking Chinese. Allowing the learners to understand how to apply the learned vocabulary to their daily life is crucial as an ideal Chinese teaching methodology. In addition to the language usage, applying sound and real-life context to the unfamiliar vocabulary will help the learners to remember it (Oxford, 1990). Learners should establish a meaningful connection between the unfamiliar vocabulary and their mother language as fast as possible, and then use multiple learning strategies and deep elaboration to make the information convert into unforgettable knowledge (Gu and Johnson, 1996). If we can associate unfamiliar vocabulary to real-life objects such as a coffee shop or a bus stop, and we have the learners listen to pronunciation of the vocabulary, language learning will be enhanced because next time the learners attempt to recall the information, the memory can be triggered by the associated pronunciation.

3.3 Adopting location-based service on mobile platform as a learning tool in situated learning

Most of established Chinese learning systems embrace the model of content delivery, which is dialog-based, story-telling and vocabulary flashcard. These types of learning systems aim to provide a scenario in which walk the user through the entire content of the course. “Let’s play Chinese Characters” is an example of using mobile games for children to learn stroke order of Chinese characters (Tian et al, 2010). For foreign Chinese learners, a context-aware mobile Chinese language learning system (CAMCLL) seeks solutions to solve real-life problem which low level Chinese learners might encounter by using context awareness technology (Al-Mekhlafi, Hu and Zheng, 2009).

An interesting cloud-based language learning system (Learn-as-you-go) uses tagging model to support Chinese learning. The authors seek to solve knowledge source management problem by offering its learners to fetch any web resource or service through cloud storage in their fingertips (Kovachev, Cao, Klamma and Jarke, 2011). Another systemsimilar to CAMCLL that uses location to support Chinese learning isMicroMandarin. The authors adopt a location-based social networking service “Foursquare” to determine learners’ location and push local information to them (Edge, Searle, Chiu, Zhao and Landay, 2011).

Although we have seen some interesting applications that assist learners to conduct ubiquitous learning outside of the classrooms, there is none offers automatic speech recognition. In our opinion, it is vital that the entry level Chinese learners verify their pronunciation of the learned vocabulary even with the help of situated learning. Without any guidance, students might mispronounce the characters until further correction made by peers or native speakers. Researches indicated that automatic speech recognition is able to lower learners’ anxiety level while attempting to learn new vocabulary, and further promote their learning motivation because they are aware of correctness level of their performance (Wang, Young and Jang, 2013). Darren et al (2011) also indicated that “Learning is best when performed through participatory experience”. Integrating location-based service with mobile assisted language learning (MALL) enhance learning experience by associating learning material to real-life objects (e.g. coffee shop, local cuisine, tourist attraction). When learners are enabled to associate abstract Chinese vocabulary to local objects, learning outcome will be amplified.

This study aim to establish a situated Chinese learning system based on location-based service supported by Facebook and automatic speech recognition technique. The proposed system will provide learners vocabulary pronunciation in a real-life context. Student will receive immediate

feedback on their pronunciation performance. The repeating vocabulary pronunciation practices which take place between different locations will help them retain real-life object Chinese information for future use.

4. Design of the mobile application

The design of this mobile application consists of 3 major functions:

(1) Location-based contextual language learning (2) Automatic speech recognition (pronunciation recognition & rating) (3) Text to speech (TTS) conversion. The implementation snapshots are shown in Figure 1.

First of all, when a learner logs onto their Facebook account, he or she will be redirected to the main user interface (see Figure 1a.). The centered map will show user's current location. The flickering blue dot indicates user's exact geography location. Upon pressing the pick-place button, the system will perform a nearby location check to determine learners' nearby location information. Learners then will be able to select the surrounding street objects to start acquiring new Chinese vocabulary (see Figure 1b).



(a) Main user interface (b) Street object selection (c) Learning content

Figure 1. Mobile Application User Interface Snapshots

Once a street object is selected, the corresponding street object will be shown to the learners. For this step, we adopt Google Text to Speechweb service to convert the street object string into a playable and repeatable mp3 file. The next step of the learning task, the learners are able to listen and record their pronunciation of the corresponding vocabulary (see Figure 1c). The corresponding street object image will also be shown. Automatic Speech Recognition (ASR) module will fetch the recorded sound file and send it back the processing server for performance evaluation. Finally, the result of learners' performance will be shown to them. Result will be graded in numbers. The application will also keep track how many times the learners has encountered the corresponding vocabulary.

The learning process is taking place in a real context where the learners can observe the objects while acquiring related vocabulary. As the authors of 3P-learning model indicate, learning is personal and self-directed. It is also driven by knowledge-pull. Learning also requires students' participation (Mohamed, Matthias and Marcus, 2010). In this proposed mobile application design, we apply 3P-learning model to the use of technology-enhanced language learning (TELL). Students can use it to acquire vocabulary according to their individual needs based on their current location. ASR module enables students to understand if they are on the track of learning by showing performance result to them. Students then will be able to use newly acquired vocabulary to initiate social interaction with native speakers. The location-based learning enables students to discover new knowledge that is crucial to their survival in a foreign country. The repeated pronunciation practices will gradually assist students to develop native-like speaking ability. Although this application do not provide regular classroom setting learning, students will have to go out there and interact with social context. We can this type of learning known as "Micro-learning". In this case, our learning process only takes place between learners' desired locations.

5. System Implementation

The mobile application will be implemented with various web development techniques and Apple iOS development SDK. It consists of Javascript Object Notation (JSON) and SQL database. Javascript will facilitate the communication between iOS and the SQL database. In addition, we will adopt Facebook Graph API as our location-based learning core module. The speech recognition is crucial for the application implementation. The ASR (Automatic Speech Recognition) toolbox is founded by Professor Jyh-Shiung Roger Jang (National Taiwan University, Taiwan). This core module is responsible for

deploying vocabulary soundtrack file playing and recording. After fetching learners' recorded file, it will be send back to the remote server for further performance analysis. Finally, the server will send back learners' performance feedback and score.

6. CONCLUSION

In this study, we have discussed our proposed mobile Chinese learning application. It is a system based on contextual learning in which learners can undergo their learning in a real-life context. With ASR module, learners will be provided with authentic Chinese pronunciation and tone which frequently troubles them. Learners will be provided with different real-life object and its information based on the location. Students' pronunciation will be corrected with ASR module, resulting in better performance. We expect our CSL learners to be less anxious when it comes to initiating social interaction with native speakers because they know they are speaking with correct pronunciation.

3. References

- Agostini, A., E. Di Biase, and M. Loregian. (2010). Stimulating cooperative and participative learning to match digital natives' needs, *8th IEEE International Conference on Pervasive Computing and Communications Workshops*, 274-279.
- Al-Mekhlafi, K, Hu, X. &Zheng, Z. (2009).An Approach toContext-Aware Mobile Chinese Language Learning forForeign Students.*Mobile Business*, 340–346
- Dörnyei, Z. &Ottó, I. (1998).Motivation in Action: A ProcessModel of L2 Motivation. Working Papers in AppliedLinguistics 4, 43–69.
- Edge, D., Searle, E., Chiu, K., Zhao, J., Landay, J.A.: MicroMandarin: Mobile Language Learning in Context. In: Proc. of CHI 2011, pp. 3169–3178. ACM (2011)
- Gu, Y., Johnson, R.K. (1996).Vocabulary Learning Strategies and Language LearningOutcomes. *Language Learning* 46(4), 643–679.

- Ho, W. C., Chang, C. Z. (2008). A Preliminary study of Chinese Learning Material for Oversea Chinese Young Children. *Annual conference of teaching Chinese as a second language*. Steering Committee for the Test of Proficiency- Huayu, Taipei.
- Hug, Theo (2005). Micro Learning and Narration. Exploring possibilities of utilization of narrations and storytelling for the designing of "micro units" and didactical micro-learning arrangements. Paper presented at the fourth Media in Transition conference, May 6–8, 2005, MIT, Cambridge (MA), USA.
- K.T. Sun, Y.M. Huang, M.C. Liu. (2011). A WordNet-based near-synonyms and similar-looking word learning system. *Educational Technology & Society*, 14 (1), pp. 121–134.
- Kovachev, D., Cao, Y., Klamma, R. and Jarke, M. (2011) 'Learn-as-you-go: New Ways of Cloud-Based Micro-Learning for the Mobile Web, H. Leung, et.al, (Eds), *ICWL 201*, Springer -Verlag, Berlin, , LNCS 7048, pp. 51-61
- M. A. Chatti, M. Jarke, and M. Specht. The 3p learning model. *Journal of Educational Technology & Society*, 13(4):74-85, 2010.
- McKeon, L (2009). Learning to speak Mandarin and understanding Chinese culture is different not difficult. Foucault: 25 years on. A conference hosted by the Centre for Post-colonial and Globalization Studies, Adelaide.
- Oxford, R. L (1990). Language learning strategies: What every teacher should know. Boston: *Heinle&Heinle Publishers*.
- Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin and Review*, 14, 187–193.
- Tian, F., Lv, F., Wang, J., Wang, H., Luo, W., Kam, M., Setlur, V., Dai, G. & Canny, J. (2010). Let's play chinese characters: mobile learning approaches via culturally inspired group games. CHI'10, 1603–1612.
- Wang, Y. S., Young, S. S. –C., Jang, J. S., –R. (2013). Using Tangible Companions for Enhancing Learning English Conversation. *Journal of Educational Technology & Society*, 16(2), 296-309.
- Xin, S. C. (2001). The Chinese language reading strategy and research. Taipei, National Taiwan Normal University, Taipei, Taiwan.