Mobile learning and speech technology for language teachers' professional development: A design-based study

Tran Le Nghi Tran
University of Queensland

This study aims to investigate the use of mobile learning to provide pronunciation training for lecturers of English as a Foreign Language (EFL) from Vietnamese provincial universities. Mobile learning offers a potential solution for the delivery of professional development to lecturers based outside major cities thanks to its capacity to enable learning anytime, anywhere. Mobile learning and speech technology are expected to facilitate lecturers' self-direct learning to fulfil their professional development needs using their own devices. This paper reports results from a pilot study which serves as the first phase of an on-going design-based research project. The pilot study was carried out to explore the feasibility of an online pronunciation course and identify potential problems for future course iterations in the context of participants living outside major cities in a developing country. The objectives of the project are to establish and test a set of fundamental principles for mobile learning to be an effective way of providing online professional development for lecturers based outside major cities and to shed light on the necessary adjustments in course design to make it a scalable model for future education planning. In this study, both qualitative and quantitative data were collected during two iterations of an online pronunciation professional development course for EFL lecturers from Vietnamese provincial universities.

Introduction

In 2008, the Vietnamese Ministry of Education and Training (MOET) initiated the National Foreign Language Project 2020, and invested approximately five billion USD in reforming language teaching methods, enhancing teachers' language competence, and purchasing learning resources (Hoang, 2011). However, there was widespread doubt about the successful achievement of the project (Parks, 2011). Critics argued that there were too few EFL teachers – about 80,000 for approximately 17.5 million students, and most were unqualified (Hoang, 2011). By the end of 2015, the percentage of teachers who met MOET's proficiency requirements increased from under 10% to 32% (Yen Anh, 2016a). However, the greatest challenge remains the improvement of teachers' oral skills, especially pronunciation (Quynh Trang, 2014).

For nearly 47,000 Vietnamese EFL teachers, improving their language competence to meet the language proficiency requirements set by Project 2020 is now a must (Yen Anh, 2016b). For those living outside major cities like, this means regular long distance travel to attend training. Sending teacher trainers to small towns to deliver on-site training, which MOET did between 2011 and 2015, has proved to be of limited efficacy, therefore alternative forms of providing English training are needed. At the end of 2016, MOET proposed to focus on online learning, and learning technologies for language learning for professional teacher training.

Mobile learning offers a potential solution to providing PD language training to EFL teachers for several reasons. Firstly, it reduces the traveling time for both educators and trainees to deliver or acquire training. Secondly, it is cost-effective since "technology can reduce training costs if there are a large number of learners, if the learners are geographically dispersed, and if the course will be repeated several times" (Welsh, Wanberg, Brown, & Simmering, 2003, p. 255). It can also enable teachers to sustain professional development using their own mobile devices.

Literature review

This study adopts the definition of professional development (PD) as "the development of a person in his or her professional role" proposed by Villegas-Reimers (2003, p. 11). However, it restricts the scope to EFL teachers, whose PD is "a lifelong process which begins with the initial preparation that teachers receive (whether at an institute of teacher education or actually on the job)

and continues until retirement" (Villegas-Reimers, 2003, p. 8) in the context of Vietnam. Recent studies of Vietnamese tertiary EFL teachers' PD (Nguyen, Fehring, & Warren, 2015) indicates an increasing amount of attention in PD for teachers of English at higher education institutions. However, the limited number of studies is not sufficient to provide an overview of PD among EFL teachers in the country. Moreover, the current approaches to providing language PD for teachers of English in Vietnam have also been criticized for the inequitable selection of only experienced lecturers to attend external PD activities, especially overseas (Tran, 2016).

Mobile learning is defined as "both learning with portable technology, and also learning in an era characterized by mobility of people and knowledge" (Sharples, Taylor, & Vavoula, 2006). In this study, mobile learning involves the use of both mobile and stationary devices (i.e. desktops) that facilitate learning on the move, i.e. when learners use a hotel desktop to study while on a business trip. This conceptualization suits the context of EFL teachers from Vietnamese provincial universities, where long distances and low income may hinder their access to high-end devices or state-of-the-art technologies. In Vietnam, there have been few studies into the use of mobile learning for EFL teachers. A rare exception is the work of Murphy, Midgley, and Farley (2014) on mobile learning trends among 44 EFL teachers who took an MA in TESOL course held in Ho Chi Minh City. The findings revealed that all participants owned or had access to up to four mobile devices, with acceptable to moderate Internet quality. However, it may not be appropriate to generalize these findings for all Vietnamese teachers.

Computer-Assisted Pronunciation Training (CAPT) is defined as the employment of digitized speech for developing language pronunciation (Rostron & Kinsell, 1995). Of all the CAPT technologies, automatic speech recognition (ASR) was recognized as the most valuable for instantaneous, individualized feedback (Hansen, 2006). ASR is a technology which enables a computer or a hand-held device to transcribe words that are read aloud or spoken into any sound-recorder (Myers, 2000). ASR was found to help improve learners' pronunciation (Golonka, Bowles, Frank, Richardson, & Freynik, 2014), overall intelligibility, learners' confidence and autonomy (Geertsema & Campbell, 2014). However, it often fails to recognize accented speech, and is unable to provide meaningful pronunciation evaluation (Neri, Cucchiarini, & Strik, 2003). The concern was addressed by the development of ASR systems that can recognize non-native speech with acceptable performance. ELSA Speak and USpell, the apps adopted in this study, are such systems.

Pronunciation was not a popular research topic in Vietnam until the 21st century, when there was a strong emphasis on learning English. Research suggests that there is an observed lack of pronunciation pedagogical training for teachers (Tweedy, 2012). Although Vietnamese highly value native-like pronunciation (Cunningham, 2009), acquiring intelligible pronunciation is a real challenge for them (Vu, 2016). Fortunately, research shows that Vietnamese learners' pronunciation problems can be successfully addressed thanks to explicit training and practice (Ngo & Setter, 2011), and technology is a promising solution (Dang, 2011).

Research questions: The research questions of this study are:

1. What conditions need to be met for an online pronunciation course to be held for EFL lecturers from provincial universities in Vietnam?
2. What adjustments need to be done in course design and implementation to make such an online course feasible and scalable?

Methodology

Design-based research is adopted as the research paradigm of the study because its characteristics align well with those of mobile learning. Design-based research is defined as a systematic but flexible methodology aimed to improve educational practices through iterative analysis, design, development, and implementation, based on collaboration among researchers and practitioners in real-world settings, and leading to contextually-sensitive design principles and theories" (Wang & Hannafin, 2005). This research paradigm is characterized by being (a) pragmatic, with the aim to solve real-world problems; (b) grounded, in both theory and real-world context; (c) interactive, with collaboration between researchers and practitioners, iterative, with theories and interventions continuously developed and refined from analysis to design then evaluation and redesign, and flexible, ongoing recursive; (d) integrative, with the utilization of a variety of research methods and approaches from both qualitative and quantitative research paradigms; and (e) contextual, connected with both the design process through which results are generated and the setting where the research is conducted (Wang & Hannafin, 2005).

In the study, three theoretical frameworks were used: The framework for analysing mobile learning proposed by Sharples et al (2005) reflects the dialectical relationship between technology and learning; the seamless language learning framework mediated by ubiquitous technology by Wong (2012) frames the course design and implementation for data collection; and the technical quality model by Sarab et al (2016) is employed for the selection of mobile technologies.

This pilot study was carried out between September and December 2016, with four participants completing the

online pronunciation course and participating in the study. They were all female with between five and eleven years of teaching experience. Two work in mountainous areas in the North and Central Vietnam, the third one is based in a small town in the Mekong Delta in the South, and the fourth is from a coastal Central province. They all have laptops, one uses an iPhone, and the others have Android smartphones. Before the course started, the participants completed a questionnaire about their background, previous pronunciation training and training needs, then took a pre-test using automatic voice recognition technology. The participants recorded their own voice reading out loud a short paragraph (107 words) and sent their recordings to the researcher who used Dragon Dictation, a voice recognition app on her phone to transcribe the recordings into texts, then the transcription were compared against the original text to identify pronunciation mistakes. A training syllabus was built upon the questionnaire responses and the pre-test results.

The three-month online pronunciation course had two phases: In the first eight weeks (October to November 2016), the participants met with the researcher for two hours every week using online conference tools like Skype to discuss their learning problems and practice their pronunciation. There was no direct instruction since participants had been provided with learning videos. Outside class meeting times, participants practiced their pronunciation using pronunciation apps and mobile technologies on their own devices. The two apps used by participants were ELSA Speak, a voice recognition app focusing on problematic pronunciation features for Vietnamese speakers, and USpell, a pronunciation app with video lessons and practice session for each sound in English. Other technologies include Rachel's English, a well-known YouTube channel for pronunciation, and online text-to-speech tools such as Speechnotes or Dictation.io using Google voice recognition technology. The participants uploaded screenshots of their in-app practice, i.e. level completion reports to Edmodo, the learning management used in the course. They also provided consent for all the online meetings to be audio-recorded, and receive links to download these recordings. In the second phase of the course (December 2016), there were no weekly meetings, and participants self-directed their pronunciation practice.

After the course, the participants completed a post-test in the same way as the pre-test, using the same reading passage. The transcribed and original texts were compared, as were the pre and post test results. Three months after that, a questionnaire was sent to the participants to collection their feedback on the course and their pronunciation practice. During the course, the researcher also documented observations of technological issues and participants' behaviours.

Results and discussion

In terms of device ownership and accessibility, all the participants had no difficulty taking the course. They often used either a laptop, tablet or their phone to take part in the Skype meetings, and switched between them for more flexibility or convenience while on the move. They were asked to install Skype on all their devices so they can have backup devices in case of battery exhaustion. However, the compatibility of apps and websites seems to be challenging to address. Some apps, like Dragon Dictation, are only available on iOS. USpell, while works on both iOS and Android devices, is not available on Windows phones. Therefore, a participant borrowed an iPad from a family member to practice. The proposed solution was to use web-based dictation tools such as Speechnotes and Dictation.io to practice dictating English texts. Again, there was an issue: Most of the free online dictation websites employ Google voice recognition technology and therefore can only work on Chrome browser. One participant had problems with Chrome and did not know how to fix them. All the participants were busy lecturers, so an ideal tool for them must be not only compatible to as many operating systems and browsers as possible but also have different versions (i.e. web-based, mobile phone apps, etc.) to allow for flexibility and choice of devices. Towards the end of the course, a new solution was discovered: The Voice Typing function in Google Docs. This can be used on Chrome browser in a laptop or desktop, and have apps for iOS and Android. The availability of mobile apps is convenient for participants for their anywhere, anytime learning.

At the beginning of the course, participants struggled with the technological issues. At first, Adobe Connect was used to make an online video conference call between the researcher and four participants, but two participants could not join the call. Either the link did not work, or they could join and had no sounds and videos. And the participants also found the call interface not very user-friendly – they did not know what to do with the buttons and did not seem to read the pdf instructions sent to them before the meeting. The researcher then had to switch to Skype, and Zoom online conferencing tools to communicate with the participants. However, the video was often lagged, the audio was distorted and there were unwanted noises when there were four or five people in a real-time video call. When videos are turned off, the sound quality of the call improved, but occasionally a participant lost connection. After a few weeks, it was discovered that some participants were using unstable Wi-Fi or 3G connection, so they used a wired connection. It was also recognized that Zoom is far more effective than Skype in multi-participant calls, and it has a Mute-all button which is great for reducing or eliminating background noises. These suggest that in future iterations of the course, participants should use wired Internet

connection, and the call group size should be small in online real-time meetings using conferencing tools. Moreover, video calls should be minimalized and replaced by audio calls when possible, and participants should be provided with initial technical training, probably in form of video tutorials before the course starts.

During the first phase of the course with regular weekly meetings with the researchers, all the participants showed a high level of commitment and engagement in both the class meetings and individual self-practice of pronunciation. During the online meetings, they proactively and eagerly took part in learning activities and were excited in addressing their pronunciation mistakes. For individual practice, they were asked to spend up to 15 minutes per day to complete one level or lesson in a pronunciation app, but only one of the participants strictly followed the instructions. All other participants spent between 30 minutes and an hour every day on pronunciation practice. They often practiced during breaks, at lunchtime or whenever they had some free time. Most of them uploaded screenshots of their in-app practice to Edmodo every day.

In the second self-directed learning phase of the course, in weeks 9 – 12, there were no longer weekly online meetings with the researcher, and the participants' level of commitment and engagement went downhill. They kept doing individual practice and uploaded their practice to Edmodo for one or two more weeks, then stopped asking questions and practicing although the researcher encouraged them to keep learning. When asked for the reasons, they admitted they were too busy, or lazy, and promised to go back to practice soon, but then did not.

There were some possible explanations for this sudden decrease in the participants' levels of commitment and engagement. Firstly, three months was a long time, and the participants lost their eagerness and excitement after two thirds of the duration. Secondly, the participants enjoyed having personal feedback and discussion during the online meetings in the first phase, and were motivated by the improvement in their pronunciation, but then felt lost and unsupported when there were no class meetings, while the apps and websites could not give them the personalized feedback they wanted. Thirdly, the participants might prefer guided learning to self-directed learning. Finally, the second phase of the course was in December, and with the semester-end examinations and the holiday season drawing closer, the participants were too busy and distracted to self-study, especially without the pressure of an upcoming meeting with the researcher. Therefore, it was proposed that in the next iterations, the course duration should be reduced to six weeks, and online meetings should be maintained during the whole duration.

The pre and post test results indicated that the participants' pronunciation accuracy improved significantly after the course. It seemed that the participants had better awareness of their pronunciation mistakes and made efforts to address them, especially in pronouncing vowels, ending sounds and consonant clusters. However, there was also an observed reluctance among participants to provide comments or suggestions or ask questions regarding the course syllabus and implementation. It seemed that most of the times, the participants just simply agreed with whatever suggested by the researcher. Therefore, when designing the next course iterations, participants' passive learning style and dependence on the researcher need to be taken into consideration.

Conclusion

This small-scaled pilot study serves as the first exploratory cycle of an ongoing design-based research project on the use of mobile learning for providing language PD for lecturers from Vietnamese provincial universities, and was followed by two iterations of course design, implementation and evaluation. Results from this study suggested that for mobile learning to be an effective method of PD provision in the context of remote participants in a developing country, special attention should be paid to initial technological training, Internet connection quality, and participants' learning style and culture. These will be incorporated into a new iterative cycle of the pronunciation course, and findings from these will be reported in future papers.

References

Cunningham, U. (2009). Models and targets for the pronunciation of English in Vietnam and Sweden. *Research in Language, 7*, 113-128. doi:10.2478/v10015-009-0008-3

Dang, N. T. (2011). *Exploring CALL options for teaching EFL in Vietnam.* (Master thesis), Minnesota State University, Mankato, Minnesota. Retrieved from http://search.proquest.com.ezproxy.library.uq.edu.au/docview/879631946#

Geertsema, M., & Campbell, C. (2014). Teachers using mobile devices to receive immediate feedback and improve learners' language learning. *Progressio, 36*(2), 38-52.

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning, 27*(1), 70-105. doi:10.1080/09588221.2012.700315

Hansen, T. K. (2006, 22-25 Nov). *Computer assisted pronunciation training: The four 'K's of feedback.* Paper presented at the m-ICTE2006 - Conference on Multimedia and Information and Communication Technologies in Education, Seville, Spain.

Hoang, V. V. (2011). The current situation and issues of the teaching of English in Vietnam. *Ritsumeikan Language and Culture Studies, 22*(1), 7-18.

Murphy, A., Midgley, W., & Farley, H. (2014, November 3-5). *Mobile learning trends among students in Vietnam.* Paper presented at the 13th World Conference on Mobile and Contextual Learning, mLearn 2014, Istanbul, Turkey.

Myers, M. J. (2000). Voice recognition software and a hand-held translation machine for second-language learning. *Computer-Assisted-Language-Learning, 13*(1), 29-41. doi:http://dx.doi.org/10.1076/0958-8221(200002)13:1;1-K;FT029

Neri, A., Cucchiarini, C., & Strik, W. (2003). *Automatic speech recognition for second language learning: How and why it actually works.* Paper presented at the Proc. ICPhS, Barcelona.

Ngo, A. P., & Setter, J. (2011, 17-21 August). *Evaluating the intonation perception ability of Vietnamese learners of English.* Paper presented at the 17th International Congress of Phonetic Sciences, Hong Kong

Nguyen, H. T., Fehring, H., & Warren, W. (2015). EFL teaching and learning at a Vietnamese university: What do teachers say? *English Language Teaching, 8*(1), 31-43. doi:10.5539/elt.v8n1p31

Parks, E. (2011, 8 Nov). Vietnam demands English language teaching "miracle". *The Guardian*. Retrieved from http://www.guardian.co.uk/education/2011/nov/08/vietnam-unrealistic-english-teaching-goals

Quynh Trang. (2014, 15 Oct). Students complain about teachers' pronunciation (Phat am sai, giao vien tieng Anh bi hoc tro phan nan). *VNExpress*. Retrieved from http://vnexpress.net/tin-tuc/giao-duc/phat-am-sai-giao-vien-tieng-anh-bi-hoc-tro-phan-nan-3093325.html

Rostron, A., & Kinsell, P. (1995). Learning pronunciation using CALL: Some experimental evidence. *ReCALL Newsletter, 5*(1).

Sharples, M., Taylor, J., & Vavoula, G. (2006). A theory of learning for the mobile age. In R. Andrews & C.

Haythornthwaite (Eds.), *The Sage Handbook of Elearning Research* (pp. 221-247). London: Sage.

Tran, H. (2016). *Professional learning and development (PLD) in higher education: The experiences of teacher educators in Vietnam.* (PhD thesis), University of Wellington, Wellington, New Zealand.

Tweedy, L. (2012). *Discovering factors that influence English pronunciation of native Vietnamese speakers.* (Master thesis), Hamline University, Minnesota, USA.

Villegas-Reimers, E. (2003). *Teacher professional development: an international review of the literature*. Paris: UNESCO: International Institute for Educational Planning

Vu, H. Y. (2016). *Exploring English pronunciation teaching in Vietnam: Time for a new approach?* (MPhil thesis), Macquarie University.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational technology research and development, 53*(4), 5-23.

Welsh, E. T., Wanberg, C. R., Brown, K. G., & Simmering, M. J. (2003). E-learning: Emerging uses, empirical results and future directions. *International Journal of Training and Development, 7*(4), 245-258.

Yen Anh. (2016a, 28/01/2016). Hiem giao vien tieng Anh dat chuan (Few teachers of English met the proficiency requirements). *Nguoi lao dong (The labourer)*. Retrieved from http://nld.com.vn/giao-duc-khoa-hoc/hiem-giao-vien-tieng-anh-dat-chuan-20160128214338106.htm

Yen Anh. (2016b, 14 Apr 2016). Thieu tram trong giao vien tieng Anh (Serious shortage of teachers of English). *Nguoi lao dong (The labourer)*. Retrieved from http://nld.com.vn/giao-duc-khoa-hoc/thieu-tram-trong-giao-vien-tieng-anh-20160414221658642.htm

Note: All published papers are refereed, having undergone a double-blind peer-review process.