Mobile-assisted learning in the second language classroom

Denis Liakin
Department of French Studies
Concordia University, Montreal, Canada
denis.liakin@concordia.ca

Walcir Cardoso, Natallia Liakina

Department of Education
Concordia University, Montreal, Canada
walcir.cardoso@concordia.ca

French Language Center
McGill University Montreal, Canada
natallia.liakina@mcgill.ca

Abstract:

Second language (L2) learners have difficulties in acquiring phonemic contrasts that are not found in theirnative languages. For instance, Anglophones have difficulty in producing and perceiving the French vowel /y/ (e.g., "u" in "tu" /ty/ 'you') due to its absence in the English vowel inventory. In this study, we examine the impact of the pedagogical use of mobile automatic speech recognition software (ASR) on the phonological acquisition of the French vowel /y/. The study took place in two universities in Montreal, Canada. The participants were beginner French students with no previous experience with pronunciation instruction or speech recognition software. They were divided into three experimental groups: (1) the ASR Group used a commercial (but free) ASR application in their mobile devices (iPhone or iPod Touch) to complete weekly pronunciation tasks, with immediate written visual feedback provided by the software; (2) the Non-ASR Group completed the same weekly pronunciation tasks in individual weekly sessions with a teacher, who providedimmediate oral feedback using recast and repetitions; finally, (3) the Control Group participated in weeklyindividual meetings "to practice their conversation skills" with a teacher, who provided no pronunciation feedback. The study followed a pretest/posttest design and lasted five weeks. ANOVA results indicated that the ASR Group outperformed the other groups in /y/ production. At the same time, ASR was evaluated very positively by the majority of the participants in our study, particularly because it is perceived as having a positive effect on pronunciation due to the immediate visual feedback that it provides.

Keywords: *mobile learning, automatic speech recognition (ASR), pronunciation teaching, e-learning*

I. INTRODUCTION

Automatic Speech Recognition (ASR) has been mainly used for business dictation, but recent developments in voice-to-text abilities have encouraged its implementation in computer-assisted language learning [1], [9], [16], [23], [31], [35]. In the context of pronunciation teaching, [31] suggest two possible applications for ASR: to teach pronunciation of a foreign language [20], and to assess students' oral production

[11], [36], [50]. The use of these applications have been confirmed in a variety of studies that demonstrate that computer-assisted pronunciation teaching via ASR can be effective in the acquisition of second or foreign language (L2) features such as speech rate, fluency, phonemes (vowels and consonants), vowel lengthening, pitch accents, stress timing, and global pronunciation skills [3], [6], [7], [8], [15], [17], [18], [19], [20], [23], [25], [26], [27], [33], [34], [35], [40], [49].

Despite the benefits for phonological acquisition described above, ASR technology fulfills most of the criteriaproposed by [5] for selecting pronunciation software and activities to develop speaking skills; namely they shouldallow for: (1) learner fit (ASR is interesting, useful for learners and allows them to identify features that they need); (2) explicit teaching (focus on particular pronunciation features; in this case, French /y/); (3) opportunities for interactions with the computer (L2 learners interact with the application via their mobile devices); (4) comprehensible feedback (ASR provides feedback that learners can understand); (5) strategy development (ASR can guide learners in developing strategies for use on their own, considering other L2 features).

The main goal of this study is to explore the use of ASR as a pedagogical tool to improve pronunciationteaching and learning of French as a second language. In the investigation, we focus on the acquisition of the French phoneme /y/ (as in “tu” /ty/ ‘you (2nd person singular) for two main reasons: it is highly difficult to acquire in both production and perception [2], [28], [39]; and it has a high functional load in the target language, as defined by [24] (e.g., [29] state that French /u-y/ is used to distinguish many French minimal pairs such as ‘audessous’/odsu/ ‘below’ and “au-dessus” /odsy/ ‘above’). To our knowledge, there are no studies that investigatemobile ASR for pronunciation teaching (see also [13] for a similar observation).

II. BACKGROUND

The majority of the studies that investigate the effects of ASR on the acquisition of L2 pronunciation skills have shown that this technology can be effective [8], [23], [26], [33], [35]. In this context, our study will contribute with more data and analyses on the effects of ASR on pronunciation, from a French L2 perspective. More importantly, it will address a gap in the literature via the implementation of a specific type of ASR, one that is easily accessible via smartphones and media players. We thus hypothesize that learners will also benefit from the technology if it is offered in a portable format.

The use of mobile devices for language learning has sparked the interest of an increasing number ofresearchers over the last decade, particularly for vocabulary acquisition [21], [22], [46], [48]. For instance, [48] found that the participants who received lessons via their phones learned more vocabulary than those who used the web or paper as means of instruction. In a more recent study, [21] investigated the use of SMS to support vocabulary learning among a group of beginning learners of Italian. The authors concluded that the students appreciated the experience overall and found it useful and enjoyable. For similar positive results on the pedagogical benefits of mobile phones, see [22] and [46].

Despite these encouraging results, [46] observe that Mobile Assisted Language Learning has not yet beenembraced on a large scale and has not received sufficient research so that we can recognize its full potential for learning. Consistent with [13]’s observation, as indicated earlier, we are not aware of any study that investigates the use of ASR and its effects on L2 phonological acquisition. To test the viability of using mobile ASR technology and its effects on learning, we will focus on the acquisition of L2 French /y/.

The target feature examined in this study is the L2 French vowel /y/. This is an ideal target phoneme forpronunciation instruction because, as mentioned earlier, /y/ is highly problematic for speakers of a variety of first languages (L1s; e.g., English, Mandarin, Spanish) in both production and perception [2], [29]. According to [10]'s Speech Learning Model, during acquisition, speech perception becomes attuned to the contrastive phonic elements of the L1 and thus learners may fail to discern the phonetic differences between sounds in the L2. In the case of French /y/, this phonemically distinct sound is thus "assimilated" into /u/, the only available option in the phonemic inventory of L1 English. In addition to its difficulty in production and perception, French /y/ has a high functional load because of the degree that it contrasts between words. The concept of functional load, attributed to [24], is customarily used to describe the extent and degree of contrast between linguistic units, usually phonemes. In phonology, it is a measure of the work that two phonemes do to maintain phonemic contrast in all possible environments. Consequently, certain phonemes in a language have higher functional loads than others depending on the degree in which they contrast meaning. For instance, French /u-y/ is used to distinguish French minimal pairs such as au-dessous /odsu/ 'below' from au-dessus /odsy/ 'above', an alternation that may change considerably the intended meaning of the speaker. Because Anglophones do not have this segment in their L1, it is essential that it be mastered early not to compromise meaning in the target language. This is one of the arguments that Jenkins used in her rationale for her English as a Lingua Franca approach, particularly in deciding priorities for pronunciation teaching. According to the author, priority should be given to sounds that have a high functional load. We believe that /y/ fulfills this requirement.

This study set out to answer the following research question: Does ASR-based pronunciation practice improve /y/ production? We hypothesize that ASR will have a positive effect on /y/ production.

III. METHODOLOGY

A. Participants and design of the study

Forty-two L2 French students participated in this study (average age: 22; 30 female, 12 male). All participants were recruited from two French courses at two Anglophone universities in Montreal; they were either native English speakers or had native-like proficiency in English. In addition, all participants had a beginner level of proficiency in French and, accordingly, had not yet acquired the target phoneme /y/. The study followed a pretest/posttest design and lasted five weeks. The participants were assigned to one of three distinct groups. The ''ASR group'' corresponded to the group that practiced French pronunciation with mobile ASR on a iPod, iPador iPhone, using a commercial (but free) ASR application (Nuance's *Dragon Dictation*). The students completed weekly pronunciation activities (total of 5) with immediate written visual feedback (speech-to-text) provided by the software, and worked on the activities for approximately 20 minutes per week. The "Non-ASR Group", on the other hand, did not have access to the mobile ASR. However, the participants in this group completed the same activities in individual, weekly 20-minute sessions with a French teacher, who provided immediate oral feedback using recast and repetitions. Finally, the "Control Group" participated in weekly individual 20-minute meetings to simply practice their conversation skills with a French teacher, who provided no pronunciation feedback.

B. Procedures

This study employed a mixed-methods approach, using a pre/post research design followed by surveys andqualitative interviews with the participants. For the production task, we used CAN 8 VirtuaLab, "an interactive, multimedia tool used for the instruction of modern languages" with which the participants were familiar. The production task consisted of the reading aloud of words and phrases, which were recorded using CAN 8. Wetargeted 20 occurrences of /y/ in the following 19 words: assume, azur, chute, culture, défendu, fumes, lune, musique, numéro, particule, perdu, plu, pulvérise, surtout, tu, ultime, unanime, une, vu.

C. Analysis

To assess the students' production, two bilingual francophone RAs listened to student's recordings anddetermined whether the pronunciation of /y/ was correct or incorrect. For the statistical analysis of the results and to test for differences among the three groups at the pretest and posttest, a one-way ANOVA was performed at each time for production and perception. To test for differences within each group over time, dependent samples t-test were carried out comparing pretest to posttest performances for each group.

D. Results

The general descriptive statistics of the analysis for /y/ production and perception appear in Table 1. Itpresents the mean scores (M) of accurate production and perception as well as standard deviations (SD) across the two tests (Pre and Post) and the three groups under consideration (ASR, NASR and Control). Because there were ten tests performed, the alpha level had to be adjusted and set at .005 (.05/10 tests). Overall, the results of the oneway ANOVA indicate that there are no group differences among the three groups either at the pretest or the posttest ($F(2, 39) = .95$, $p = .392$ and $F(2, 39) = .90$, $p = .413$ in pre and posttest respectively).

	Production ($n = 20$)					
	ASR		NASR		Control	
Test	*M*	*SD*	*M*	*SD*	*M*	*SD*
Pre	7.09	5.51	9.79	3.98	8.43	5.86
Post	10.71	4.23	11.86	3.98	9.50	5.53

Table 1. Descriptive statistics for /y/ production and /y/ perception over time, across three groups (Mean scores)

To test for differences within each group over time, dependent samples t-test were carried out comparingpretest performance to posttest performance for each group. According to the results from the dependent samples t-tests, only the ASR group improved significantly from pretest to posttest ($p < .001$). This indicates that learners who received instruction via mobile ASR learned how to produce French /y/ in a more target-like manner than those who received teacher-based input and feedback (NASR) or no input or feedback whatsoever (Control). Forillustrative purposes, the results for production are plotted in Figure 2, where the values represent the mean scores for accurate /y/ production.

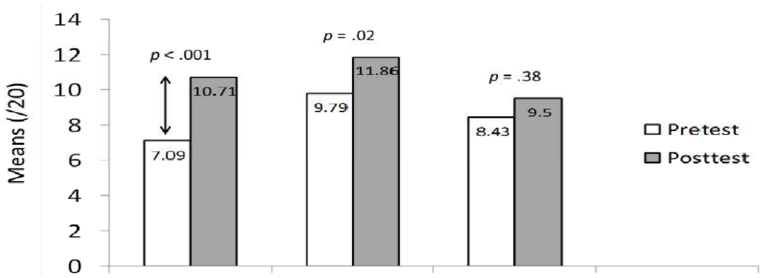

Figure 2. /y/ production results.

IV. CONCLUDING REMARKS

The main goal of this study was to explore the use of ASR as a pedagogical tool to improve pronunciationteaching and learning of French as a second language. More specifically, our study investigated the effects of ASR-based instruction on the acquisition of French /y/ in production.

The results obtained indicate that, similar to what is observed in the ASR literature, the use of speechrecognition has a positive effect on the acquisition of phonology [8], [33], [35]. We attribute these learning gains to a variety of factors that include insights from the general SLA/CALL literature, including [5]'s ideas about input enhancement and computer-aided interaction (e.g., /y/ pronunciation is reinforced via orthography, input manipulation and repetition among ASR users), the effects of an explicit focus on the target form, immediate feedback, multiple opportunities for learning, and the game-like approach to teaching afforded by mobile technologies. Lastly, mobile ASR technology ascribes to [5]'s suggestions for selecting pronunciation software to develop speaking skills: learner fit, potential for explicit teaching, opportunities for interactions with the computer, comprehensible feedback, and strategy development to guide students to start learning new L2 features on their own. Evidently, we are aware that the observed gains could also be caused by the effect of the adoption of a new technology.

Is there anything for the L2 language teacher and learner? Based on the findings of our study, we believe that ASR can be used to complement the activities conducted in the language classroom. For instance, the teacher could emphasize meaningful communicative tasks in the classroom as recommended by L2 pedagogues while she or he may assign the repetitive and lengthy tasks that characterize the activities used in our study as personalized homework assignments. We believe that ASR can and should be used in the language learning environmentbecause: (1) It has the potential to improve L2 learners' pronunciation; (2) It can relocate resources so that classroom time is used exclusively for communicative activities; (3) It accommodates a wider variety of learners; and, finally, (4) It was evaluated very positively by the participants, as they believed ASR helped them improve their pronunciation due to the immediate visual feedback that it provides as well as its portability and usability.

References :

1. Aist, G. (1999). Speech recognition in computer-assisted language learning. In Cameron, K. (Ed.), *CALL: media, design & applications*, Germany: Swets & Zeitlinger, 165-181.
2. Baker, W. & Smith, L. C. (2010). The Impact of L2 Dialect on Learning French Vowels: Native English Speakers Learning Québécois and European French. *Canadian Modern Language Review*, 66, 711-738.

3. Cauldwell, R. (2002). Streaming Speech: Listening and advanced pronunciation for advanced learners of English. Talking Computers, *Proceedings of the IATEFL Pronunciation and Computer Special Interest Groups*, pp. 18–22.
4. Chapelle, C. (2012). Using mixed-methods research in technology-based innovation for language learning. Plenary talk. Innovative Practices in Computer Assisted Language Learning, Ottawa, April 26th.
5. Chapelle, C. & Jamieson, J. (2008). Tips for teachers: Computer-assisted language learning. New York: Pearson Longman.
6. Coniam, D. (2002). Technology as an awareness raising tool for sensitising teachers to features of stress and rhythm in English. *Language Awareness*, 11(1), 30–42.
7. Chun, D. (1998). Signal analysis software for teaching discourse intonation. *Language Learning and Technology*, 2(1), 74–93.
8. Dalby. J., & Kewley-Port, D.(1999). Explicit pronunciation training using automatic speech recognition. *CALICO*, 16 (3), 425-445.
9. Eskenazi, M. (1999). Using Automatic Speech Processing for foreign language pronunciation tutoring: Some issues and a prototype. *Language Learning and Technology*, 2(2), 62–76.
10. Flege, J. E., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /¢/ and /l/. *Journal of Acoustical Society of America*, 99, 1161-1173.
11. Franco, H., Neumeyer, L., Digalakis, V., & Ronen, O. (2000). Combination of machine scores for automatic grading of pronunciation quality. *Speech Communication*, 30, 121–130.
12. Gardner, H. (1983). Frames of Mind: The Theory of Multiple Intelligences. New York: Basic Books.
13. Godwin-Jones, R. (2009). Emerging technologies: personal learning environments. Language Learning and Technology, 13(2), 3-9.
14. Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds 'l' and 'r'. *Neuropsychologia*, 9, 317–23.
15. Hardison, D. (2004). Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *Calico Journal*, 22(2), 175–190.
16. Hincks, R. (2003). Speech technologies for pronunciation feedback and evaluation. *ReCALL*, 15 (1), 3-20.
17. Hincks, R. (2005). Measures and perceptions of liveliness in student oral presentation speech: A proposal for an automatic feedback mechanism. *System*, 33, 575–591.
18. Hirata, Y. (2004). Computer assisted pronunciation training for native English speakers learning Japanese pitch and durational contrasts. *Computer Assisted Language Learning*, 17, 357-376.
19. Kaltenboeck, G. (2002). Computer-based intonation teaching: Problems and potential. Talking Computers, *Proceedings of the IATEFL Pronunciation and Computer Special Interest Groups*, 11–17.
20. Kawai, G., & Hirose, K. (2000). Teaching the pronunciation of Japanese double-mora phonemes using speech recognition technology. *Speech Communication*, 30, 131–143.
21. Kennedy, C., & Levy, M. (2008). L'italiano al telefonino: Using SMS to support beginners' language learning. *ReCALL Journal* 20(3), 315-330.
22. Kiernan, P. & Aizawa, K. (2004) Cell phones in task based learning. Are cell phones useful language learning tools? *ReCALL*, 16(1): 71-84.

23. Kim, I. S. (2006). Automatic speech recognition: Reliability and pedagogical implications for teaching pronunciation. *Educational Technology and Society*, 9(1), 322–344.
24. King, R. D. (1967). Functional load and sound change. *Language*, 43, 831-852.
25. Lambacher, S. (1999). A CALL tool for improving second language acquisition of English consonants by Japanese learners. Computer *Assisted Language Learning*, 12(2), 137–156.
26. LaRocca, S. T., Moagan J. J., & Bellinger S. M. (1999). On the path to 2X learning: Exploring the possibilities of advanced speech recognition, *CALICO*, 16 (3), 295-310.
27. Levis, J., & Pickering, L. (2004). Teaching intonation in discourse using speech visualization technology. *System*, 32(4), 505–524.
28. Levy, E. S. & Law II, F. F. (2010). Production of French vowels by American-English learners of French: Language experience, consonantal context, and the perception-production relationship. *Journal of the Acoustical Society of America*, 128, 1290-1305.
29. Levy, E. S. & Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, 36, 141-157.
30. Littlewood, W. (2004). The task-based approach: some questions and suggestions. *ELT Journal*, *58*(4), 319- 326.
31. Mak, B., Siu, Ng, Tam, Chan, Y-C., Chan, K-W. (2003). PLASER: Pronunciation Learning via Automatic Speech Recognition, *Proc. HLT-NAACL 2003 Workshop on Building Educational Applications using Natural Language Processing*. Edmonton, Canada, 23-29.
32. Meers, J. (2009). The acquisition of front rounded and nasalized vowels of French by native speakers of English. Undergraduate Honors Theses, University of Calgary.
33. Mostow, J., & Aist, G. (1999). Giving help and praise in a reading tutor with imperfect listening because automated speech recognition means never being able to say you're certain. *CALICO*, 16(3), 407-424.
34. Neri, A., Cucchiarini, C., & Strik, H. (2006). Selecting segmental errors in L2 Dutch for optimal pronunciation training. *International Review of Applied Linguistics*, 44, 357–404.
35. Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008). The effectiveness of computer assisted pronunciation training for foreign language learning by children. *Computer Assisted Language Learning*, 21(5), 393-408.
36. Neumeyer,L., H. Franco, V. Digalakis, and M. Weintraub (2000) Automatic Scoring of Pronunciation Quality. *Speech Communication*, 30, 83-93.
37. Nunan, D. (2004). *Task-Based Language Teaching.* Cambridge: Cambridge UniversityPress.
38. Piske, Thorsten, MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29, 191-215.
39. Rochet, B. 1995. Perception and Production of Second-Language Speech Sounds by Adults. In Speech Perception and Linguistic Experience: Theoretical and Methodological Issues, ed. by Winifred Strange. York Press. Timonium, MD.
40. Seferoglu, G. (2005). Improving students' pronunciation through accent reduction software. *British Journal of Educational Technology*, 36(2), 303–316.
41. Sheldon, A. (1985). The relationship between production and perception of the /r/–/l/ contrast in Korean adults learning English: A reply to Borden, Gerber, and Milsark. *Language Learning*, 35, 107–13.

42. Sheldon A and Strange W (1982) The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applies Psycholinguistics* 3: 243–61.

43. Sheldon, A. & Henly, E. (1986). Duration and context effects on the perception of English /r/ and /l/: A comparison of Cantonese and Japanese speakers. *Language Learning*, Vol. 6, No. 4, 505-521.

44. Sheldon, A. & Strange W. (1982). The Acquisition of /r/ and /l/ by Japanese Learners of English: Evidence that Speech Production Can Precede Speech Perception. *Applied Psycholinguistics*, 3.3., 243-261.

45. Song, Y. & Fox, R. (2008). Using PDA for undergraduate student incidental vocabulary testing. *ReCALL* 20 (3), 290-314.

46. Stockwell, G. (2008). Investigating learner preparedness for and usage patterns of mobile learning. *ReCALL*, 20(3), 253-270.

47. Stuckless, R. (1994). Developments in real-time speech-to-text communication for people with impaired hearing. In M. Ross (Ed.), *Communication access for people with hearing loss* (pp.197-226). Baltimore, MD: York Press.

48. Thornton, P. and Houser, C. (2005) Using mobile phones in English education in Japan. *Journal of Computer Assisted Learning*, 21: 217-228.

49. Wang, X., & Munro, M. (2004). Computer-based training for learning English vowel contrasts. *System*, 32, 539–552.

50. Witt, S., & Young, S. (2000). Phone-level pronunciation scoring and assessment for interactive language learning. *Speech Communication*, 30, 95–108.