
Mobile speech recognition software: A tool for teaching second language pronunciation

Denis Liakin

denis.liakin@concordia.ca

CONCORDIA UNIVERSITY

Walcir Cardoso

walcir@education.concordia.ca

CONCORDIA UNIVERSITY

Natallia Liakina

natallia.liakina@mcgill.ca

MCGILL UNIVERSITY

Abstract

This study examines the impact of the pedagogical use of mobile automatic speech recognition software (ASR) on the acquisition of the French vowel /y/ in production and perception. The participants were 42 beginner French students with no previous training in French phonetics and exposure to speech recognition software. They were divided into three experimental groups: (1) the ASR Group used an ASR application installed on their mobile devices to complete weekly pronunciation activities, with immediate written visual (textual) feedback provided by the software; (2) the Non-ASR Group completed the same weekly pronunciation activities in individual weekly sessions with a teacher, who provided immediate oral feedback using recast and repetitions; finally, (3) the Control Group participated in weekly individual meetings “to practice their conversation skills” with a teacher, who provided no pronunciation feedback. Following a pre-test/post-test design, our findings indicate that the ASR Group outperformed the other groups in French /y/ production, but not in perception.

Key words: automatic speech recognition, pronunciation teaching, French phonology, production, perception

Résumé

Cette recherche étudie l'impact de l'utilisation pédagogique des logiciels mobiles de reconnaissance automatique de la parole (RAP) sur l'acquisition de la voyelle française /y/ dans la production et dans la perception.

Quarante-deux participants étaient des débutants en français sans aucune expérience préalable dans la phonétique et sans exposition à des logiciels de reconnaissance de la parole. Ils ont été divisés en trois groupes expérimentaux : (1) le groupe RAP a utilisé une application de RAP installée sur des appareils mobiles pour compléter des activités hebdomadaires de prononciation, avec une rétroaction immédiate écrite visuelle (textuelle) fournie par le logiciel, (2) le groupe non-RAP a complété les mêmes activités hebdomadaires de prononciation lors des rencontres individuelles avec un professeur qui a fourni une rétroaction immédiate orale, et enfin, (3) le groupe de contrôle a participé à des rencontres hebdomadaires individuelles avec un professeur pour pratiquer l'oral sans aucune rétroaction. En suivant le design de pré-test/post-test, nos résultats indiquent que le groupe RAP a dépassé les autres groupes dans la production de la voyelle /y/, mais pas dans sa perception.

Mots-clés : reconnaissance automatique de la parole, enseignement de la prononciation, phonologie du français, production, perception

Introduction

Automatic Speech Recognition (ASR) is a computer-based technology that transcribes speech into readable text in real time. While ASR has been mainly used for business dictation, recent developments in voice-to-text abilities have encouraged its implementation in computer-assisted language learning (CALL—e.g. Aist, 1999; Eskenazi, 1999; Hincks, 2003; Kim, 2006; Neri, Mich, Gerosa and Giuliani, 2008). In the context of teaching pronunciation, Mak et al. (2003) suggest two possible applications for ASR: (1) to teach pronunciation of a foreign language (Kawai and Hirose, 2000); and (2) to assess students' oral production (Franco, Neumeyer, Digalakis and Ronen, 2000; Witt and Young, 2000). These applications have been investigated in a variety of studies that demonstrate that computer-assisted pronunciation teaching (CAPT) via ASR can be effective in the acquisition of second (L2) or foreign language features such as phonemes, stress, and general pronunciation skills (e.g. Chun, 1998; Hardison, 2004; Hincks, 2005; Kim, 2006; Levis, 2007; Neri et al., 2008). In addition to these benefits, ASR technology fulfills most of the criteria proposed by Chapelle and Jamieson (2008) for selecting pronunciation software and activities to develop speaking skills. Specifically, ASR allows for: (1) learner fit (ASR is useful for learners, allowing them to identify needed features); (2) explicit teaching (ASR can be used to focus on particular pronunciation features); (3) opportunities for interactions with the computer; (4) comprehensible feedback; and (5) the development of strategies to learn new features on their own.

The main goal of this study is to explore the use of mobile ASR as a pedagogical tool to improve pronunciation teaching and learning of French as a sec-

ond language. In the investigation, we focus on the acquisition of the French phoneme /y/ (as in *tu* /ty/ ‘you’) for two main reasons: (1) the sound is highly difficult to acquire in both production and perception (e.g. Baker and Smith, 2010; Levy and Law II, 2010; Rochet, 1995); and (2) it has a high functional load in the target language (Jenkins, 2002; King, 1967), as it is used to distinguish many French minimal pairs such as *au-dessous* /od.su/ ‘below’ and *au-dessus* /od.syl/ ‘above’. To our knowledge, there are no studies that have investigated the use of ASR on mobile devices for pronunciation teaching (see also Godwin-Jones, 2009, for a similar observation).

Background

ASR in the acquisition of second language pronunciation

The majority of the studies that have investigated the effects of ASR on the acquisition of L2 pronunciation skills have shown that this technology can be effective. For instance, Neri et al., (2008) investigated whether a CAPT system can help learners improve word-level pronunciation skills in English at a level comparable to that achieved through traditional teacher-led training. Their results showed that the pronunciation quality of isolated words improved significantly after ASR-based treatment. Kim (2006) examined the reliability of ASR software used to teach English pronunciation. The oral production of 36 students was compared to pronunciation scores determined by native English speaking instructors. Although the results indicated that ASR technology is still not as accurate as human analysis, the author concluded that the software may be useful for student practice with certain aspects of pronunciation (see also Dalby and Kewley-Port (1999), LaRocca, Morgan and Bellinger, (1999) and Mostow and Aist (1999) for other similar positive results).

In sum, the available literature suggests that ASR technology may have positive effects on the acquisition of L2 pronunciation. In this context, our study will provide 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.

Mobile technology and second language acquisition

The use of mobile devices for language learning has sparked the interest of an increasing number of researchers over the last decade, particularly for *vocabulary* acquisition (e.g. Kiernan and Aizawa, 2004; Kennedy and Levy, 2008; Thornton and Houser, 2005; Stockwell, 2008). For instance, Thornton and Houser (2005) found that the participants who received lessons via their phones

learned more vocabulary than those who accessed vocabulary through the web or print resources. In a more recent study, Kennedy and Levy (2008) investigated the use of *SMS (short message service) to support vocabulary learning in L2 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 Kiernan and Aizawa (2004) and Stockwell (2008).

Despite these encouraging results, Kukulska-Hulme and Shield (2008) observe that Mobile Assisted Language Learning has not yet been embraced on a large scale and has not yet received sufficient research for its full potential as a pedagogic practice to be recognized. Consistent with Godwin-Jones' (2009) observation, as indicated earlier, we are not aware of any study that investigates the use of ASR on mobile devices and its effects on L2 phonological acquisition. To assess the viability of using mobile ASR technology and test its effects on learning, we focused on the acquisition of L2 French /y/.

French /y/ and its acquisition

The target French pronunciation feature examined in this study is the vowel /y/. This is an ideal target phoneme for pronunciation instruction because, as mentioned earlier, /y/ is highly problematic for speakers of a variety of first languages (e.g. English, Mandarin, Spanish) in both production and perception (Baker and Smith, 2010; Levy and Strange, 2008). According to Flege's (1995) 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 a similar phoneme in the learner's L1 (e.g. /u/ or /i/ for English and Farsi speakers, respectively).

In addition to its difficulty in production and perception, /y/ has a high functional load (King, 1967), a concept 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 many languages (L1s) lack this phoneme, it is essential that it be mastered early on in order to not compromise meaning in the target language. This is one of the arguments that Jenkins (2002) used in her rationale for her English as a Lingua Franca approach, particularly in deciding priorities for pronuncia-

tion teaching. According to the author, priority should be given to sounds that have a high functional load. We believe that /y/ fulfills this requirement.

Research questions

This study set out to answer the following two general research questions:

1. Does ASR-based pronunciation practice using a mobile device improve French L2 /y/ *production*?
2. Does ASR-based pronunciation practice using a mobile device improve French L2 /y/ *perception*?

We hypothesized that the use of ASR would have a positive effect on /y/ production, and that learners would be able to extend the newly acquired productive skill into perception, based on the assumption that perceptual learning may transfer to L2 speech production (e.g. Rochet, 1995; Bradlow, Pisoni, Yamada and Tohkura, 1997). We define perception as the participant's ability to discriminate between a set of options, namely /y/, /u/ and /i/ embedded in words, phrases and sentences.

Methodology

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/. Figure 1 illustrates the design of this study.

The study followed a pretest/posttest design and lasted five weeks. The participants were randomly assigned to one of three distinct groups (Table 1).

The "ASR group" corresponded to the group that practiced French pronunciation with mobile ASR on an iPod or iPhone, using a commercial (but free) ASR application (Nuance's Dragon Dictation). The students completed at home, on a weekly basis, five 20-minute pronunciation activities that consisted of reading aloud of target words and phrases, using the ASR software installed on their mobile devices. After each reading attempt, students were provided with immediate written visual feedback via an orthographic representation of their attempt (speech-to-text analysis). To illustrate, if students attempted to pronounce the word *pure* but they read *pour* or *pire* as the written (visual) result, this should indicate that their pronunciation was incorrect, thus requiring another attempt (in some exceptional case, slow connection or background noise would affect it, but students were aware of it). Students were asked to

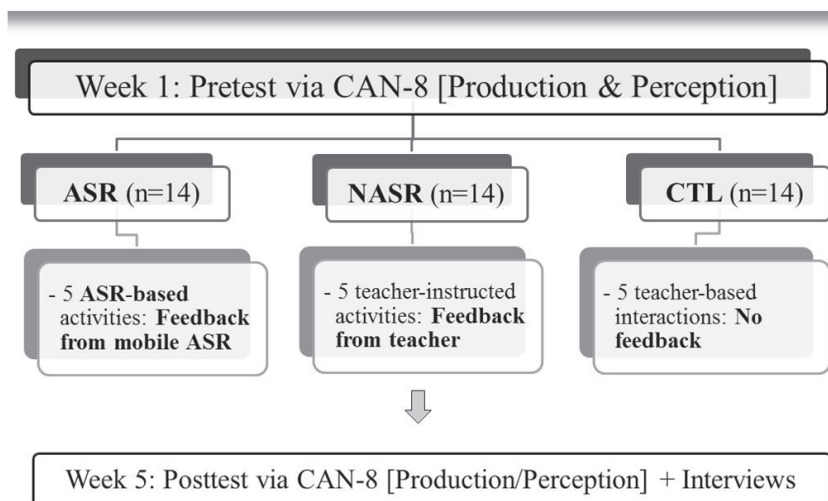


FIGURE 1

Design of the study

spend one minute per word/phrase, for a total of 20 minutes. They were also asked to indicate, in a form, the number of times they repeated each form until they were able to produce it accurately, or until their 1-minute limit had expired. The “Non-ASR Group”, on the other hand, did not have access to mobile ASR. However, they completed the same activities (i.e. reading aloud the same words and phrases) in individual, weekly 20-minute sessions with a French teacher, who provided immediate oral feedback on their pronunciation using recast and repetitions. Finally, the “Control Group” participated in weekly individual 20-minute meetings with the goal of practicing their conversation skills with a French teacher, who provided no feedback on /y/ pronunciation.

Procedures: Tasks

The study employed a mixed-methods approach, using a pre/post-test research design followed by surveys and interviews with the participants. For the pre/post-test’s production and perception tasks used to measure students’ pronunciation capabilities, we employed 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. We targeted 20 instances of /y/ (plus 15 distractors) in 19 words (Table 2), carefully selected so that the target /y/ occurred

TABLE 1
Summary of groups and related tasks

	ASR Group	Non-ASR Group	Control Group
Tasks	Pronunciation activities using mobile ASR app (reading words and expressions)	Reading the same words and expressions as ASR group + conversation with teacher	Conversations with teacher
Focus	Phonemes /y/ + distractors	Phonemes /y/ + distractors	None (oral proficiency)
Feedback	Mobile ASR (written visual feedback)	Teacher (oral feedback)	None
Duration	5 weekly mobile ASR-based activities (20 mins. per week)	5 weekly individual sessions (20 mins. per week)	5 weekly individual sessions (20 mins. per week)

in open (CV; n=10) and closed (CVC; n=10) syllabic environments. Figure 2 shows the CAN-8 interface illustrating a production task.

TABLE 2
Production task: Words and phrases

assume	musique	surtout
azur	numéro	tu
chute	particule	ultime
culture	perdu	unanime
défendu	plu	une
fumes	pulvérise	vu
lune		

In the perception task, the participants listened to 45 monosyllabic “French” pseudowords containing the vowels /y/, /u/ and /i/ (15 instances of each vowel; e.g. foupe, fuppe, fippe). Pseudowords were used to avoid frequency and familiarity effects (e.g. some participants could select *tu* as containing /y/ simply because of their familiarity with this word). The task followed a 4-item multiple-choice format, with each alternative representing one of the relevant three vowels and “I don’t know” to avoid random selection. After listening to a word, participants were asked to choose the alternative that corresponded to what they heard. Figure 3 illustrates the interface of the perception task.

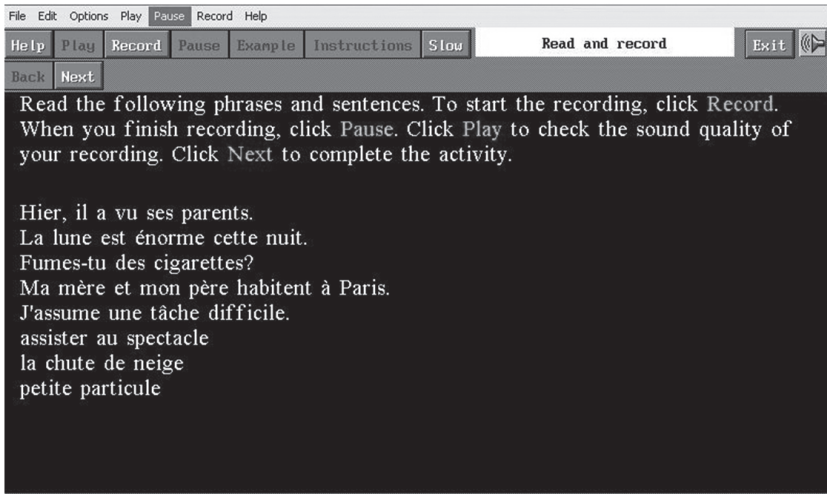


FIGURE 2

Example of production task

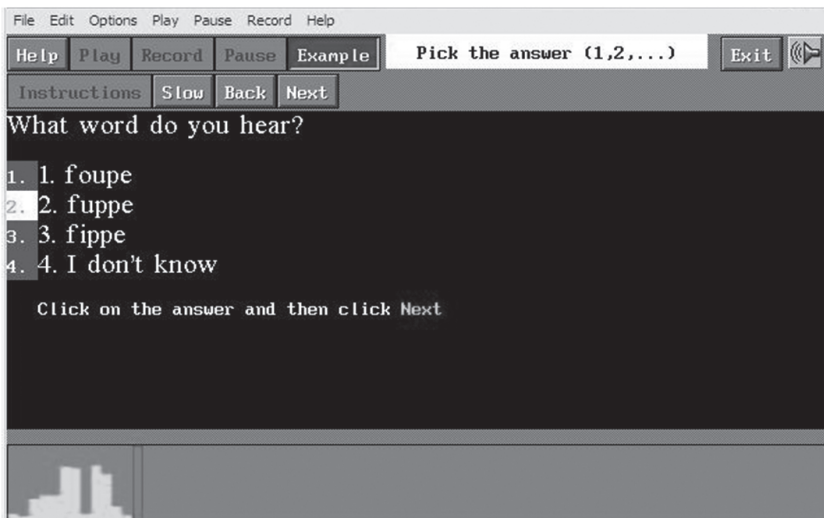


FIGURE 3

Example of perception task

Analysis

To assess the students' production, two bilingual francophone RAs (students in applied linguistics with strong knowledge of phonetics and phonology) listened to each student's recordings and determined whether the pronunciation of /y/ was correct or incorrect. In the case of divergence, a member of our team listened to those occurrences and made the decision. In total, there were 1,680 occurrences of /y/ and the inter-rater reliability was 88.7 (1,490/1,680). As for the assessment of the students' perception, this was done automatically by CAN-8, which was programmed to assess each response as correct or incorrect. For the statistical analysis of the data and to test for differences among the three groups in 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.

Results

The general descriptive statistics of the analysis for /y/ production and perception appear in Table 1. It presents 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, Non-ASR 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 one-way ANOVA indicate that there are no differences among the three groups either in the pretest or the posttest in both /y/ production ($F(2, 39) = .95, p = .392$ and $F(2, 39) = .90, p = .413$ in pre and posttest respectively) and /y/ perception ($F(2, 39) = 1.57, p = .221$ and $F(2, 39) = .32, p = .731$ in pre- and posttest, respectively).

To test for differences within each group over time, dependent samples t-test were carried out comparing pretest performance to posttest performance for each group. In this analysis, only the ASR Group improved significantly from pretest to posttest in /y/ production ($p < .001$) and no group improved in /y/ perception. We will now discuss each of these sets of results.

/y/ production

The first research question asked: Does ASR-based pronunciation practice using a mobile device improve French L2 /y/ production? 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 the mobile ASR application learned how to produce French /y/ in a more target-like manner than those who received teacher-based input and feedback (Non-ASR) or no input or feedback whatsoever (Control).

TABLE 3

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

	Production <i>n</i> = 20						Perception <i>n</i> = 15					
	ASR		Non-ASR		Control		ASR		Non-ASR		Control	
Test	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Pre	7.07	5.51	9.79	3.98	8.43	5.86	8.07	3.64	9.64	3.67	10.29	2.81
Post	10.71	4.23	11.86	3.98	9.50	5.53	9.93	2.89	10.29	3.77	10.93	3.40

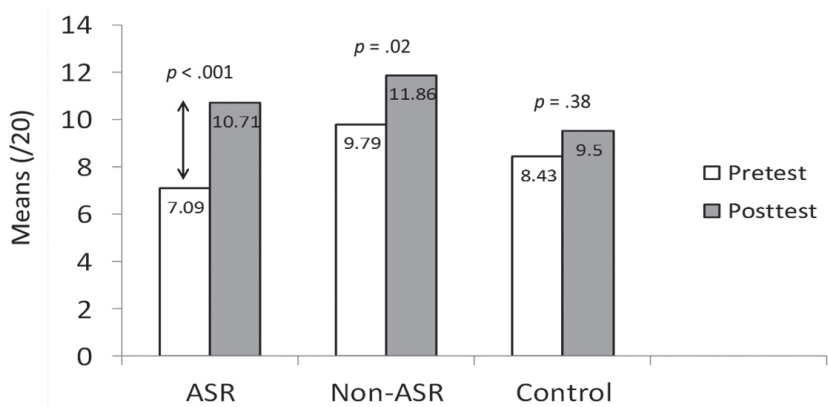


FIGURE 4

/y/ production results

For illustrative purposes, the results for production are presented in Figure 4, where the values illustrate the mean scores for accurate /y/ production.

/y/ perception

The second research question asked: Does ASR-based pronunciation practice using a mobile device improve French L2 /y/ perception? The results of the dependent samples t-tests indicate that despite slightly greater gains for the ASR group, the three groups behaved in a similar way (pre/posttest differences: ASR: $p > .05$; Non-ASR: $p > .38$; Control: $p > .37$). This indicates that the group that received ASR-based treatment was not able to extend the newly acquired knowledge detected in production to perception.

Discussion and concluding remarks

The main goal of this study was to explore the use of ASR software on mobile devices as a pedagogical tool to improve the learning of L2 French pronunciation. More specifically, our study investigated the effects of ASR-based practice on the acquisition of French /y/ in production and perception.

With regards to production, the results indicate that, similar to what is observed in the ASR literature, the use of speech recognition appears to have a positive effect on the acquisition of phonology (e.g. Dalby and Kewley-Port, 1999; Mostow and Aist, 1999; Neri et al., 2008). We attribute these learning gains to a variety of factors that include insights from the general SLA/CALL literature, including Chapelle's (2001) 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 (Dabaghi, 2010; Dekeyser, 1993), immediate feedback (Rosa and Leow, 2004), multiple opportunities for learning (Christison, 1999; Chun and Plass, 1996), and the game-like approach to teaching afforded by mobile technologies (Bruff, 2009). Lastly, mobile ASR technology ascribes to Chapelle and Jamieson's (2008) suggestions for selecting pronunciation software to develop speaking skills (based on research by Hardison, 2004, 2005; Derwing, Munro and Wiebe, 1998; MacDonald, Yule and Powers, 1994): 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, increasing interest and motivation of the students (Clark, 1983; Strambi, 2001; Warschauer, 1996).

Regarding perception, our results indicate that L2 learners were not able to transfer the acquired knowledge about /y/ into perception. We attribute the results to at least two main factors. Firstly, it is possible that the total of 1.5 hours of instruction were not sufficient for learners to acquire /y/ in perception. Secondly, we admit that we were originally optimistic to conjecture that a focus on production could translate into gains in perception. Along the lines of Goto (1971), Henly and Sheldon (1986), Sheldon (1985), our findings seem to suggest that speech production can sometimes precede its perception, as the participants in the ASR group improved only in the former. We are aware that it is premature to arrive at generalizable conclusions due to some of the limitations of the study (e.g. small number of participants, short duration of training sessions and treatment, heterogeneous groups of participants whose first languages differed, and the focus on one single phoneme). Despite these limitations, and based on the general trends observed (e.g. the ASR group did outperform the other two groups, but not significantly), we are optimistic about

the potential of ASR for the development of speech perception.

In sum, based on the findings of our study, we believe that ASR software on mobile technology should be further explored as a potential complement for pronunciation activities conducted in the language classroom. For instance, a teacher could emphasize meaningful communicative tasks in the classroom, as recommended by L2 pedagogues (e.g. Littlewood, 2004; Nunan, 2004), and assign the repetitive ASR-based activities for personalized homework assignments. Accordingly, we believe that ASR can and should be used in the language learning environment because: (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, as mentioned above; (3) It accommodates a wider variety of learners (e.g. *spatial* or visual learners—those who could benefit the most from the visual interactions afforded by speech recognition software; Gardener, 1983); and, finally, according to the questionnaires and the interviews that we conducted with participants, (4) It was evaluated very positively, as the participants believed ASR helped them improve their pronunciation due to the immediate visual feedback that it provides as well as its portability and usability.

References

- Aist, G. 1999. Speech recognition in computer-assisted language learning. In K. Cameron (ed.), *CALL: Media, design and applications*. Lisse: Swets and Zeitlinger, pp. 165–181.
- Baker, W. and L. Smith. 2010. The impact of L2 dialect on learning French vowels: Native English speakers learning Québécois and European French. *Canadian Modern Language Review*, 66, pp. 711–738.
- Bradlow, A.R., D.B. Pisoni, R.A. Yamada and Y. Tohkura. 1997. Training Japanese listeners to identify English /r/ and /l/: II. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101, pp. 2299–2310.
- Bruff, D. 2009. *Teaching with classroom response systems: Creating active learning environments*. San Francisco: Jossey-Bass.
- Chapelle, C. 2001. *Computer applications in second language acquisition: Foundations for teaching, testing, and research*. Cambridge: Cambridge University Press.
- Chapelle, C. and J. Jamieson. 2008. *Tips for teachers: Computer-assisted language learning*. New York: Pearson Longman.
- Christison, M.A. 1999. *A guidebook for applying multiple intelligences theory in the ESL/EFL classroom*. Burlingame, CA: Alta Book Center Publishers.
- Chun, D. 1998. Signal analysis software for teaching discourse intonation. *Language Learning and Technology*, 2, pp. 74–93.

- Chun, D.M. and J.L. Plass. 1996. Effects of multimedia annotations on vocabulary acquisition. *Modern Language Journal*, 80, pp. 183–198.
- Clark, R. 1983. Reconsidering research on learning from media. *Review of Educational Research*, 53, pp. 445–459.
- Dalby, J. and D. Kewley-Port. 1999. Explicit pronunciation training using automatic speech recognition. *CALICO*, 16, pp. 425–445.
- Dabaghi, A. 2010. *Corrective feedback in second language acquisition: Theory, research and practice*. Saarbrücken: LAP Lambert Academic Publishing.
- Dekeyser, R.M. 1993. The effect of error correction on L2 grammar knowledge and oral proficiency. *Modern Language Journal*, 77, pp. 501–514.
- Derwing, T., M. Munro and G. Wiebe. 1998. Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48, pp. 393–410.
- Eskenazi, M. 1999. Using automatic speech processing for foreign language pronunciation tutoring: Some issues and a prototype. *Language Learning and Technology*, 22, pp. 62–76.
- Flege, J. 1995. Second language speech learning: Theory, findings and problems. In W. Strange (ed.), *Speech perception and linguistic experience: Theoretical and methodological issues*. Baltimore: York Press, pp. 233–277.
- Franco, H., L. Neumeyer, V. Digalakis and O. Ronen. 2000. Combination of machine scores for automatic grading of pronunciation quality. *Speech Communication*, 30, pp. 121–130.
- Gardner, H. 1983. *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Godwin-Jones, R. 2009. Emerging technologies: personal learning environments. *Language Learning and Technology*, 132, pp. 3–9.
- Goto, H. 1971. Auditory perception by normal Japanese adults of the sounds ‘l’ and ‘r’. *Neuropsychologia*, 9, pp. 317–323.
- Hardison, D. 2004. Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *CALICO Journal*, 222, pp. 175–190.
- Hardison, D. 2005. Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *CALICO*, 222, pp. 175–190.
- Henly, E. and A. Sheldon. 1986. Duration and context effects on the perception of English /r/ and /l/: A comparison of Cantonese and Japanese speakers. *Language Learning*, 6, pp. 505–521.
- Hincks, R. 2003. Speech technologies for pronunciation feedback and evaluation. *ReCALL*, 15, pp. 3–20.
- Hincks, R. 2005. Measures and perceptions of liveliness in student oral presentation speech: A proposal for an automatic feedback mechanism. *System*, 33, 575–591.

- Kawai, G. and K. Hirose. 2000. Teaching the pronunciation of Japanese double-mora phonemes using speech recognition technology. *Speech Communication*, 30, pp. 131–143.
- Jenkins, J. 2000. *The phonology of English as an international language: New models, new norms, new goals*. Oxford: Oxford University Press.
- Kennedy, C. and M. Levy. 2008. L'italiano al telefonino: Using SMS to support beginners' language learning. *ReCALL*, 20, pp. 315–330.
- Kiernan, P. and K. Aizawa. 2004. Cell phones in task based learning. Are cell phones useful language learning tools? *ReCALL*, 16, pp. 71–84.
- Kim, I.S. 2006. Automatic speech recognition: Reliability and pedagogical implications for teaching pronunciation. *Educational Technology and Society*, 9, pp. 322–344.
- King, R.D. 1967. Functional load and sound change. *Language*, 43, pp. 831–852.
- Kukulska-Hulme, A. and L. Shield. 2008. An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, 20, pp. 271–289.
- LaRocca, S.T., J.J. Morgan and S.M. Bellinger. 1999. On the path to 2X learning: Exploring the possibilities of advanced speech recognition. *CALICO*, 16, pp. 295–310.
- Levis, J. 2007. Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics*, 27, pp. 1–19.
- Levy, E.S. and F.F. Law II. 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, pp. 1290–1305.
- Levy, E.S. and W. Strange. 2008. Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, 36, pp. 141–157.
- Littlewood, W. 2004. The task-based approach: Some questions and suggestions. *English Language Teaching Journal*, 58, pp. 319–326.
- MacDonald, D., G. Yule and M. Powers. 1994. Attempts to improve English L2 pronunciation: The variable effects of different types of instruction. *Language Learning*, 44, pp. 75–100.
- Mak, B., M. Siu, M. Ng, Y.-C. Tam, Y.-C. Chan, K.-W. Chan, K.-Y. Leung, S. Ho, F.-H. Chong J. Wong and J.Lo. 2003. PLASER: Pronunciation Learning via Automatic Speech Recognition. In *Proceedings of the Workshop on Building Educational Applications using Natural Language Processing*. Edmonton, Canada, pp. 23–29.

- Mostow, J. and G. Aist. 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, pp. 407–424.
- Neri, A., O. Mich, M. Gerosa and D. Giuliani. 2008. The effectiveness of computer assisted pronunciation training for foreign language learning by children. *Computer Assisted Language Learning*, 21, pp. 393–408.
- Nunan, D. 2004. *Task-based language teaching*. Cambridge: Cambridge University Press.
- Rochet, B. 1995. Perception and production of second-language speech sounds by adults. In W. Strange (ed.), *Speech perception and linguistic experience: Theoretical and methodological issues*. York Press. Timonium, pp. 379–410.
- Rosa, E. and R. Leow. 2004. Computerized task-based exposure, explicitness, type of feedback, and Spanish L2 development. *Modern Language Journal*, 88, pp. 192–216.
- 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, pp. 107–113.
- Stockwell, G. 2008. Investigating learner preparedness for and usage patterns of mobile learning. *ReCALL*, 20, pp. 253–270.
- Strambi, A. 2001. The interaction of web-based interaction and collaboration on the language learner. Unpublished doctoral thesis, University of Sydney.
- Thornton, P. and C. Houser. 2005. Using mobile phones in English education in Japan. *Journal of Computer Assisted Learning*, 21, pp. 217–228.
- Warschauer, M. 1996. Comparing face-to-face and electronic communication in the second language classroom. *CALICO*, 13, pp. 7–26.
- Witt, S. and S. Young. 2000. Phone-level pronunciation scoring and assessment for interactive language learning. *Speech Communication*, 30, pp. 95–108.