Inceoglu, S. & Gnevsheva, K. (2020). Ultrasound imaging in the foreign language classroom: Outcomes, challenges, and students' perceptions. In O. Kang, S. Staples, K. Yaw, & K. Hirschi (Eds.), *Proceedings of the 11th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, Northern Arizona University, September 2019 (pp. 115–126). Ames, IA: Iowa State University.

ULTRASOUND IMAGING IN THE FOREIGN LANGUAGE CLASSROOM: OUTCOMES, CHALLENGES, AND STUDENTS' PERCEPTIONS

Solène Inceoglu, Australian National University
Ksenia Gnevsheva, Australian National University

Recent years have witnessed an increase in the use of ultrasound imaging as a technology to provide visual biofeedback in L2 speech training. Despite showing promising effects, most of the past studies have been conducted in laboratory settings on a small group of learners (Bliss et al., 2018), and little is known about the use of ultrasound technology in real classroom settings. The current study explored the benefits as well as students' perceptions of using ultrasound biofeedback to teach the French /y/-/u/ in two intact French pronunciation classes. Thirty-one learners completed oral assessment tasks before and after ultrasound activities, and reported their impressions in reflective journals. Group 1 was first trained on the /y/-/u/ contrast while Group 2 first received feedback on the /e/-/ɛ/ contrast. Each learner received in-class ultrasound visual feedback on their articulation of the vowels while practicing a list of #CV# words in facilitating and difficult consonantal contexts. F2 mid-point measurements of /y/ and /u/ were extracted from target words. Results revealed an overall significant improvement in /u/—although not directly linked to the ultrasound practice, and native-like F2 values for /y/. The findings also highlighted learners' overwhelming positive perceptions and attitudes towards their experience.

INTRODUCTION

An important area of second language (L2) pronunciation research is the exploration of ways to help learners acquire the phonological system of their L2. In the past decades, studies have pointed to the importance of visual information in L2 learning and teaching as a way to supplement or interpret the auditory signal. Research shows that audiovisual perceptual training can enhance L2 pronunciation even in the absence of production training (e.g., Hardison, 2003; Hazan et al., 2005; Inceoglu, 2016). In addition, a plethora of studies have demonstrated the pedagogical benefits of visual displays (e.g., pitch, spectrograms, and waveforms) on segmental and suprasegmental learning (e.g., Levis & Pickering, 2004; Okuno & Hardison, 2016; Olson, 2014; Patten & Edmonds, 2015). Another tool that is getting increasing attention in L2 research is ultrasound imaging technology (Bliss et al., 2018). Ultrasound imaging is an effective and non-invasive method that is used to display internal articulatory processes in order to explain, monitor, and facilitate the understanding of how to pronounce challenging sounds. In the field of speech pathology, its effectiveness has been attested in the remediation of difficult segmental contrasts by hard-of-hearing adolescents and adults (e.g., Bacsfalvi et al., 2007; Preston et al., 2017).

Ultrasound Imaging for L2 Pronunciation Development

Recent studies support the use of ultrasound imaging technology in L2 learning (Cleland et al., 2015; Ouni, 2013; Tateishi & Winters, 2013; Tsui, 2012; Wilson, 2014; Wilson & Gick, 2006; Wu et al., 2015). An obvious advantage lies in the ability to provide direct and immediate feedback of the tongue movements in real-time speech (Kartushina et al., 2015). In addition, ultrasound visualization has been found to be easier to interpret for linguistically naïve participants than a spectrogram (Wilson, 2014), offering attractive possibilities for L2 pronunciation teaching and training. In terms of the effectiveness of the technique, studies have looked at small groups of learners—often 10 or fewer (Kocjančič Antolík et al., 2019; Tateishi & Winters, 2013; Tsui, 2012; Wu et al., 2015). For instance, Kocjančič Antolík et al. (2019) examined the effects of three individual 45-minute ultrasound pronunciation training sessions on the /y/-/u/ production of four Japanese learners of French (and two control participants) enrolled in a French pronunciation course. Their results showed that three of the learners in the experimental group improved the lingual articulation of the two vowels and the contrast between them, and that this improvement was maintained in a delayed posttest. Similarly to that study, previous research mostly focused on time-intensive ultrasound sessions—30 to 60 minutes of training—with multiple sessions of training (Tsui, 2012) and little is known about how short ultrasound imaging training can affect L2 pronunciation development. Moreover, no study has explored the use of ultrasound in intact L2 classrooms, which is now possible thanks to the increasing accessibility and portability of ultrasound technology.

Accordingly, the goal of the current study was to explore the benefits and students' perceptions of using ultrasound biofeedback to teach French L2 vowel contrasts. The research questions that guided the study were:

1. What are the effects of one in-class session of ultrasound training on the production of /y/-/u/ by learners of French?
2. How do learners of French perceive ultrasound imaging technology as a tool to improve their pronunciation?

METHOD

Participants and Setting

Participants were 31 learners of French enrolled in a 12-week French pronunciation and phonetics course. They were between 18 and 34 years old (*M* = 21.5) and had been learning French for 2.5 to 14 years (*M* = 6.3). Most of them were Australian English native speakers, one participant was bilingual in Australian English and Korean, and one was a Mandarin Chinese native speaker. The course met twice per week, for a total of three hours, and the prerequisite was five semesters of French (i.e., intermediate level).

Recording Tasks

Production of the French /y/-/u/ contrast was assessed by means of a read-aloud task. Table 1 shows the distribution of the minimal pairs in each reading passage. Because of the nature of the texts and in order not to overtly draw learners' attention to these specific words, it was not possible to get a balanced distribution of all the stimuli. However, each vowel appeared at least once in each consonantal context.

Table 1
Number and distribution of /y/ and /u/ in the reading passages

	/ty/	/tu/	/dy/	/du/	/vy/	/vu/	Total /y/	Total /u/
Oral 1 (170 words)	6	2	2	2	2	2	10	6
Oral 2 (202 words)	2	2	3	1	5	2	10	5
Oral 3 (226 words)	7	3	4	1	1	2	12	6

Reflective Journals

As part of the course assessment, students submitted reflective journals every second week (i.e., weeks when they did not submit an oral recording). The purpose of the journals was to encourage the students to reflect on the current state of their pronunciation and their perceived difficulties, and to think critically about anything new they had learned or noticed in terms of pronunciation skills. They were also asked to reflect on the classroom activities and any other task they found helpful in their pronunciation learning.

Procedures

Participants submitted their oral recordings at the end of Weeks 2, 4, and 7. As illustrated in Figure 1, the first recording was done before the first ultrasound session and served as a pretest. The second and third recordings immediately followed the first and second ultrasound sessions, respectively. Because the data were collected in real classroom environments, the extent of the experimental setting was limited, and it was deemed unethical to provide ultrasound training to one group only. Instead, the switching replications design was used: group 1 received ultrasound feedback on the /y/-/u/ contrast in Week 4 and on the mid-vowel /e/-/ɛ/ contrast in Week 7, whereas group 2 was first trained on the mid-vowels. This allowed researchers to examine the effects of ultrasound imaging instruction in both groups at different points in time. Vowel height for the /e/-/ɛ/ contrast was not analyzed in the current study and only served as additional ultrasound practice for the counterbalancing nature of the experiment.

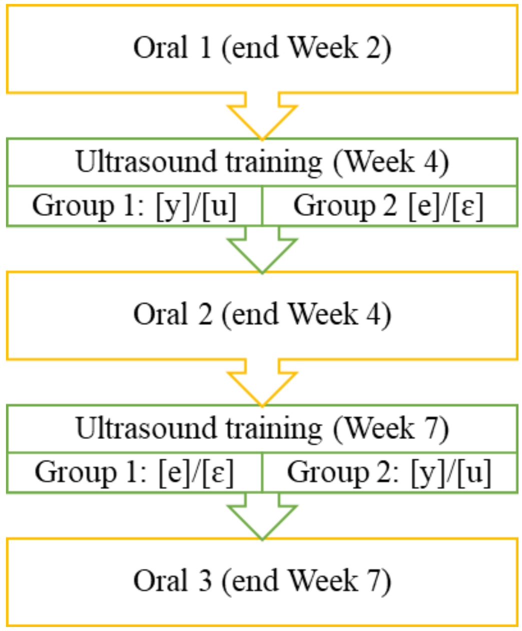

Figure 1. Overview of the procedures.

The two ultrasound imagining sessions followed similar procedures. First, the instructor—a native speaker of French and the first author—demonstrated how ultrasound technology works and how the differences between the contrastive pairs /y/-/u/ or /e/-/ɛ/ are visually salient. As seen in Figure 2, the white line on each image indicates the tongue, with the front of the mouth on the left, and the differences in backness between /y/ and /u/ and in height between /e/ and /ɛ/ are clearly visible. After making sure that learners could interpret and recognize the vowels solely based on the ultrasound images (e.g., producing silent vowels and asking students if it was an /y/ or an /u/), each learner received approximately five minutes of individualized ultrasound practice with feedback. The training focused on producing #CV# words with various initial consonants (e.g., /my/-/mu/, /ty/-/tu/) along with an examination of the ultrasound image feedback. When a learner's production of /y/ was not front enough, they were asked to produce an /i/ to see how front their tongue should reach and then round their lips. Conversely, when the /u/ was not back enough, learners practiced with the /a/ and /o/ sounds.

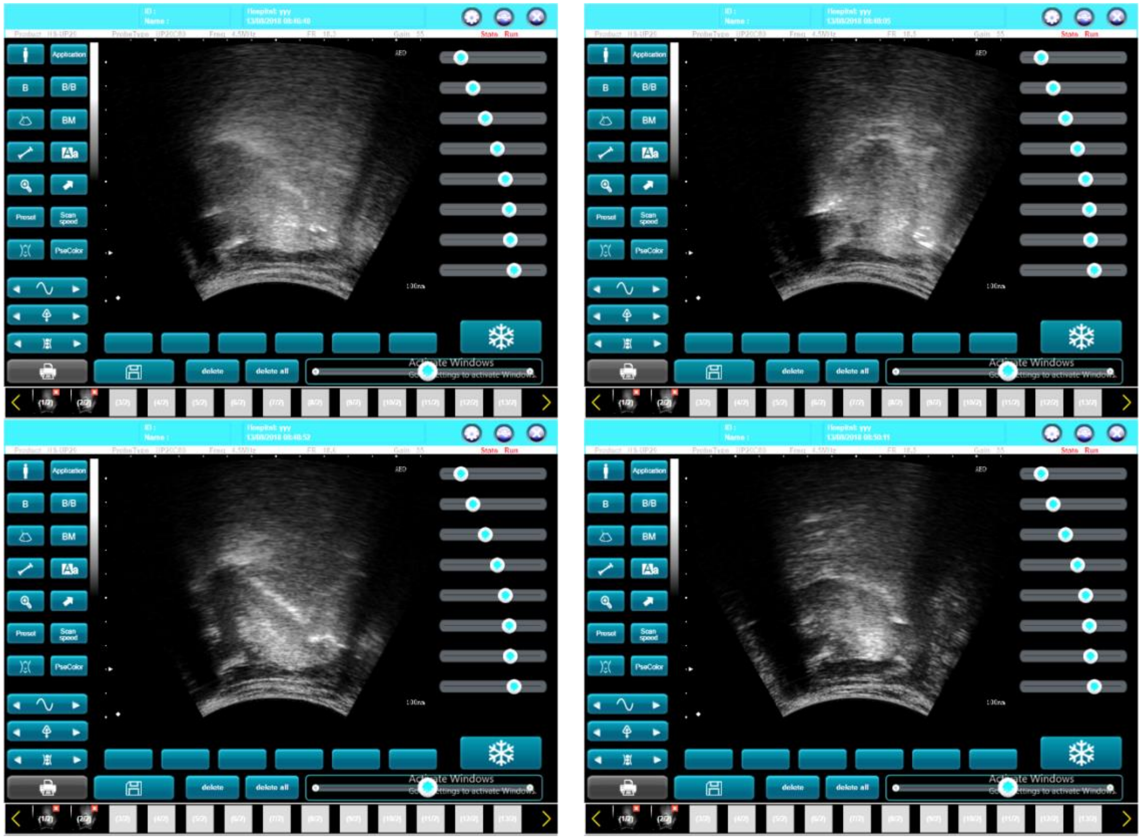

Figure 2. Midsagittal tongue images for /y/ and /u/ (top) and /e/ and /ɛ/ (bottom). The white curved line represents the tongue. The front of the mouth is on the left.

Analyses

Considering that the current study only focused on the /y/-/u/ contrast, backness was the sole point of interest for the analysis. Therefore, vowel F2 in the target words in the orals was measured at mid-point in Praat (Boersma & Weenink, 2018). Three linear mixed effects models (Baayen et al., 2008) were fit, with the difference in the F2 between /u/ and /y/, /y/ F2, and /u/ F2 as dependent variables. The independent variables were Time (Oral 1, 2, & 3), Group (1 & 2), and Vowel duration. The random intercepts were Participant, to control for speaker-specific effects, and Word, to control for word-specific effects such as word frequency and phonological environment. The models were pruned of non-significant interactions and main effects.

RESULTS

Distance Between /y/ and /u/

The results of model 1 (Table 2) looked at the difference in the mean F2 between /y/ and /u/. Note that higher numbers in Estimate indicate a bigger separation between the two vowels. The only significant predictor was Time (Figure 3), such that there was significant improvement from Oral 1 to Oral 2. For Oral 2 to Oral 3, although there was a numerical difference here as well, this difference was not statistically significant. There was, however, no interaction between Time & Group.

Table 2
Final model for /y/-/u/ F2

	Estimate	Std. error	df	t value	Pr(>|t|)	Sign.
(Intercept)	834.6	101.57	48.22	8.217	0.0000	***
Oral1	-176.13	87.83	54.97	-2.005	0.0498	*
Oral3	61.23	88.11	54.43	0.695	0.4901	

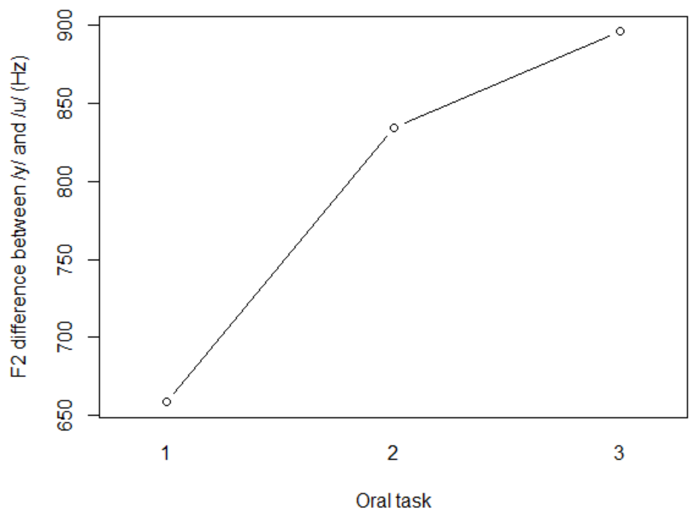

Figure 3. Predicted mean F2 difference between /y/ and /u/ for the three oral tasks.

Learners' Production of /y/

Findings for /y/ show that both factors, Group and Oral recordings, were not significant, although there was a numerical difference. Figure 4 illustrates numerical improvement from Oral 1 to Oral 2 and then regression from Oral 2 to Oral 3 for both groups (and more so for Group 1). Note that larger numbers indicate a fronter vowel.

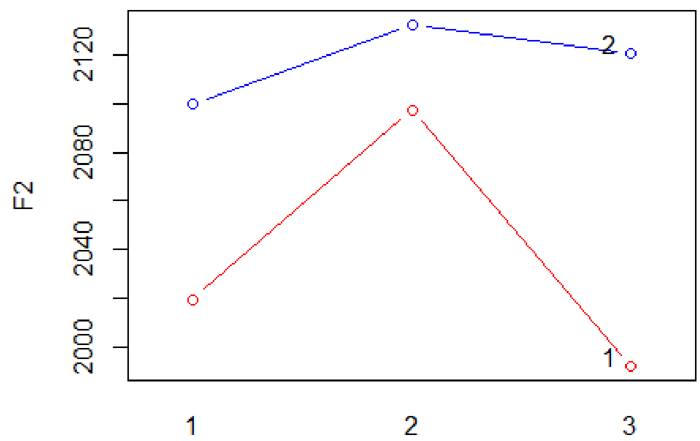

Figure 4. Predicted mean F2 for /y/ by the two groups. Group 1 (red) was trained on the /y/-/u/ contrast between Oral 1 and Oral 2 and Group 2 (blue) was trained on the contrast between Oral 2 and Oral 3.

Learners' Production of /u/

The final model for /u/ F2 in Table 3 indicates that there was a significant improvement (a backer realization of /u/; lower F2) between Oral 2 and Oral 3 for Group 1 (which received ultrasound training between Oral 1 and Oral 2) and between Oral 1 and Oral 2 for Group 2 (which received ultrasound training between Oral 2 and Oral 3). There was also an effect of vowel duration such that longer vowels had a backer realization, indicating a centralization effect of shorter vowels.

Table 3.
Final model for /u/ F2

	Estimate	Std. error	df	t value	Pr(>|t|)	Sign.
(Intercept)	1835.90	121.62	32.51	15.10	0.0000	***
Oral1	-26.83	75.91	28.36	-0.35	0.7263	
Oral3	-130.20	54.24	79.82	-2.40	0.0187	*
Group2	-54.82	149.50	29.16	-0.37	0.7165	
Duration	-1.82	0.39	426.42	-4.62	0.0000	***
Oral1:Group2	222.45	103.16	28.38	2.16	0.0397	*
Oral3:Group2	139.19	76.33	81.86	1.82	0.0719	

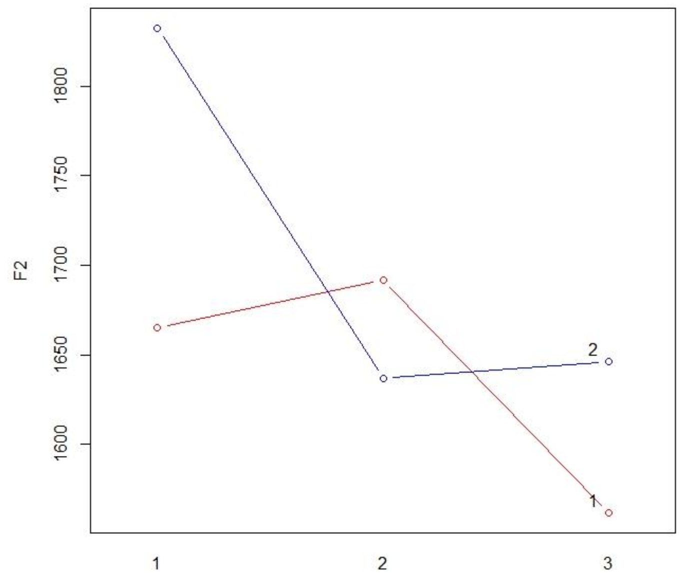

Figure 5. Predicted mean F2 for /u/ by the two groups. Group 1 (red line) was trained on the /y/-/u/ contrast between Oral 1 and Oral 2 and Group 2 (blue line) between Oral 2 and Oral 3.

Learners' Perceptions of Ultrasound Imaging.

The second goal of the current study was to examine students' perceptions of ultrasound imaging as a tool to practice French pronunciation. Qualitative comments were obtained from bi-weekly journal entries that students submitted as a part of their required assessment for the course. Overall, 28 (out of 31) reports suggest a high learner engagement and positive perception of the ultrasound as it allowed students to clearly see their tongue placement and understand their mistakes. This is illustrated in the following student's journal entry:

> *"I think the ultrasound activity was useful, I still find it difficult to have the correct tongue placement for sounds but seeing it through the ultrasound gave me a clearer understanding of how to produce sounds and where I make mistakes".*

In addition, students commented on the benefits of being able to see their articulation, and therefore their pronunciation, in addition to hearing it. This enabled them to feel how a sound should be produced and to use this feeling:

> *"A mon avis c'était très utile, parce qu'on pourrait regarder (et non seulement entendre) la différence dans la production du son. Comme ça, on pourrait se souvenir de la sensation de prononcer le son correctement".* [I think it was very useful, because we could see (and not only hear) the difference in the production of the sound. This way, we could remember the feeling of pronouncing a sound correctly].

The few students who could already accurately produce vowel contrasts also enjoyed using the ultrasound noting that this helped them confirm what they were doing right:

> *"It was very interesting - particularly for demonstrating the subtle differences and actually having evidence of what was correct and what my mouth was doing - rather than relying on what I 'thought' I was doing."*

Using the ultrasound in a classroom is not without challenges, but being able to let all the learners practice and get visual feedback was an important aspect of the lesson. This was possible thanks to the portability and ease of use of the ultrasound, and students were really enthusiastic about it:

> *"I thought it was really useful! (...) I thought it was really good that everyone got the chance to do it, (...)."*

In addition, two students offered comments which were still positive but with some caveats. The first student noted that practicing with the ultrasound in a classroom of about 15 students was distracting:

> *"This technology was a useful tool in understanding how the tongue position changes in close-mid and open mid-vowels [e] and [ɛ]. It gave me a better visual understanding of what is happening when I produce certain sounds. However, I did find the busyness of the classroom a bit distracting, and too noisy. I think it is a practical tool for pronunciation learning in smaller classroom settings."*

And another one expressed some disappointment that the ultrasound did not provide more help:

> *"It helped me to physically see what I was doing wrong, but it didn't really help me to fix my pronunciation. It is a good tool to observe progress but it didn't teach me how to produce the sounds correctly."*

DISCUSSION

The first goal of the current study was to explore the potential benefits of using ultrasound imaging in intact pronunciation classes rather than in a lab, as it has mostly been done in past research (e.g., Pillot-Loiseau et al., 2015). The linguistics focus was the French /y/ and /u/, a contrast that has been reported as challenging for English L1 speakers (e.g., Levy & Law, 2010; Ruellot, 2011). Overall, results showed significant positive changes for /u/ and no change for /y/, and more importantly a significant increase in the distance between the two vowels.

The lack of change in the production of /y/ can be explained by the fact that the vowel was already front enough, with a mean F2 similar to the one produced by native speaker (i.e., 2046-2091 Hz) (Georgeton et al., 2012). There was, therefore, no need (or room) for improvement, as opposed to the learners' productions of /u/, which was produced with a mean F2 of 1657 Hz (Group 1) and 1754 Hz (Group 2) at pretest. Despite positive changes in terms of backness—and a mean F2 of around 1600 Hz after instruction—it still remained too fronted in comparison to the native speakers' means of 779-1153 Hz (Georgeton et al., 2012). These results are also in line with previous observations that Anglophones produce the French /u/ as a front vowel due to their assimilation of the vowel into the existing English /u/ category, but accurately produce the /y/ since it is a novel category (Flege & Hillenbrand, 1984; Levy & Law, 2010).

A surprising finding of the study is that the improvement in the production of /u/ was not observed in the oral recording that immediately followed ultrasound training. What the current study showed was that the group trained on mid-vowels during Week 4 improved on their production of /u/ that same week, while the group that received /y/-/u/ ultrasound feedback during Week 4 only showed significant pronunciation improvement in Week 7—after the second session of ultrasound imaging on mid-vowels. Arguably, learners' improvement in the absence of ultrasound training can be explained by the fact that they were enrolled in a pronunciation course, which contributed to phonological awareness raising and explicit learning—especially through (peer) feedback, speech analysis and phonetic transcription activities, etc. We do not, however, have explanations for why the first group that received /y/-/u/ ultrasound training did not improve at all after that first session. It would be interesting to see whether the same pattern is observed for the acquisition of the /e/-/ɛ/ contrast.

From a pedagogical perspective, the two ultrasound sessions were a success, but they presented challenges that are absent in lab settings and that are worth mentioning here. The most important difficulty was to manage the rest of the classroom while each student practiced with the ultrasound. This can be done with the help of an assistant who can facilitate the practice and the clean-up. With careful lesson planning, the instructor can also assign (small group) activities where the teacher's input is minimal. A second challenge was the variability of learner proficiency as it is often the case in language classrooms, and the time spent on task until each participant could produce accurate contrasts. Each student went through the same list of practice stimuli, and all were able to produce a native-like contrast at least once, which highlights the benefits of ultrasound training, at least in the short-term. However, some who initially struggled did so in less than a minute whereas others needed almost the whole five minutes.

To conclude, interpreting the ultrasound images was not a challenge, and every single learner was easily able to understand the feedback and monitor whether their pronunciation was accurate. Along with the learners' enthusiasm, and thanks to the portability and increasing affordability of this technology, this shows that ultrasound imaging can have a place in the language (pronunciation) classroom. In terms of future directions, it would be interesting to work on implementing more ultrasound sessions in class. This is, of course, a challenge considering the short length of the semester (i.e., 12 weeks, including a midterm and a final exam) and the amount of material that has to be covered besides segmental contrasts. A possible way to increase learners' use of ultrasound technology would be to bring more ultrasound probes to the classroom and have learners work in small groups. Monitoring each other's productions and providing feedback to classmates might lead to an increase of awareness and have strong pedagogical implications.

ABOUT THE AUTHORS

Solène Inceoglu is a Lecturer (Assistant Professor) in French in the School of Literature, Languages and Linguistics at the Australian National University. She received her Ph.D. in Second Language Studies from Michigan State University. Her research focuses on second language speech perception/production, pronunciation instruction, the role of individual differences in SLA, and psycholinguistics. Contact information: solene.inceoglu@anu.edu.au

Ksenia Gnevsheva is a Lecturer in Linguistics in the School of Literature, Languages and Linguistics at the Australian National University. Her research lies at the intersection of sociophonetics and second language acquisition with a focus on sociophonetic variation in bilingual speakers in production and perception. Contact information: ksenia.gnevsheva@anu.edu.au

REFERENCES

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Bacsfalvi, P., Bernhardt, B. M., & Gick, B. (2007). Electropalatography and ultrasound in vowel remediation for adolescents with hearing impairment. *Advances in Speech Language Pathology*, *9*, 36–45.

Bliss, H., Abel, J., & Gick, B. (2018). Computer-assisted visual articulation feedback in L2 pronunciation instruction: A review. *Journal of Second Language Pronunciation*, *4*, 129–153.

Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer (Version 6.0.42) [Computer software]. Retrieved from http://www.praat.org.

Cleland, J., Scobbie, J. M., Nakai, S., & Wrench, A. A. (2015). Helping children learn non-native articulations: The implications for ultrasound-based clinical intervention. *Proceedings of the 18th International Congress of Phonetic Sciences*, 2–6.

Flege, J. E., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America*, *76*, 708–721.

Georgeton, L., Paillereau, N., Landron, S., Gao, J., & Kamiyama, T. (2012). Analyse formantique des voyelles orales du français en contexte isolé: à la recherche d'une référence pour les apprenants de FLE. In *Proceedings Journées d'Étude sur la Parole -TALN-RECITAL* (Vol. 1, p. 145-152). Grenoble, France.

Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, *24*, 495–522.

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, *47*, 360–378.

Inceoglu, S. (2016). Effects of perceptual training on second language vowel perception and production. *Applied Psycholinguistics*, *37*, 1175–1199.

Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. (2015). The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *The Journal of the Acoustical Society of America*, *138*, 817–832.

Kocjančič Antolík, T., Pillot-Loiseau, C., & Kamiyama, T. (2019). The effectiveness of real-time ultrasound visual feedback on tongue movements in L2 pronunciation training. *Journal of Second Language Pronunciation*, *5*, 72–97.

Levis, J., & Pickering, L. (2004). Teaching intonation in discourse using speech visualization technology. *System*, *32*, 505–524.

Levy, E. S., & Law, F. F. (2010). Production of French vowels by American-English learners of

French: Language experience, consonantal context, and the perception-production relationship. *The Journal of the Acoustical Society of America*, *128*, 1290–1305.
Okuno, T., & Hardison, D. M. (2016). Perception-production link in L2 Japanese vowel duration: Training with technology. *Language Learning & Technology*, *20*, 61–80.
Olson, D. J. (2014). Phonetics and technology in the classroom: A practical approach to using speech analysis software in second-language pronunciation instruction. *Hispania*, *97*, 47–68.
Ouni, S. (2013). Tongue control and its implication in pronunciation training. *Computer Assisted Language Learning*, *27*, 439–453.
Patten, I., & Edmonds, L. A. (2015). Effect of training Japanese L1 speakers in the production of American English /r/ using spectrographic visual feedback. *Computer Assisted Language Learning*, *28*, 241–259.
Pillot-Loiseau, C., Kamiyama, T., & Antolík, T. K. (2015). French /y/-/u/ contrast in Japanese learners with/without ultrasound feedback: Vowels, non-words and words. *Proceedings of the 18th International Congress of Phonetic Sciences*, *1*, 1–5.
Preston, J. L., McAllister Byun, T., Boyce, S. E., Hamilton, S., Tiede, M., Phillips, E., … Whalen, D. H. (2017). Ultrasound images of the tongue: A tutorial for assessment and remediation of speech sound errors. *Journal of Visualized Experiments*, *119*, 1–10.
Ruellot, V. M. (2011). Computer-assisted pronunciation learning of French /u/ and /y/ at the intermediate level. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference* (pp. 199–213). Ames, IA: Iowa State University.
Tateishi, M., & Winters, S. (2013). Does ultrasound training lead to improved perception of a non-native sound contrast?: Evidence from Japanese learners of English. *Proceedings of the 2013 Annual Conference of the Canadian Linguistic Association*, 2–15.
Tsui, H. M.-L. (2012). *Ultrasound speech training for Japanese adults learning English as a second language*. Unpublished master thesis, Simon Fraser University.
Wilson, I. (2014). Using ultrasound for teaching and researching articulation. *Acoustical Science and Technology*, *35*, 285–289.
Wilson, I., & Gick, B. (2006). Ultrasound technology and second language acquisition research. In M. G. O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference* (pp. 148–152). Somerville, MA: Cascadilla Proceedings Project.
Wu, Y., Gendrot, C., Hallé, P., & Adda-decker, M. (2015). On improving the pronunciation of French /r/ in Chinese learners by using real-time ultrasound visualization. *Proceedings of the 18th International Congress of Phonetic Sciences*, 1–5.