Recherches anglaises et nord-américaines

55 | 2022
Pronunciation Matters

Teaching Pronunciation with Direct Visual Articulatory Feedback: Pedagogical Considerations for the Use of Ultrasound in the Classroom

Barbara Kühnert and Claire Pillot-Loiseau

Electronic version
URL: https://journals.openedition.org/ranam/364
DOI: 10.4000/ranam.364
ISSN: 3000-4411

Publisher
Presses universitaires de Strasbourg

Printed version
Date of publication: July 20, 2022
Number of pages: 9-24
ISBN: 9791034401161
ISSN: 0557-6989

Electronic reference
Barbara Kühnert and Claire Pillot-Loiseau, "Teaching Pronunciation with Direct Visual Articulatory Feedback: Pedagogical Considerations for the Use of Ultrasound in the Classroom", *Recherches anglaises et nord-américaines* [Online], 55 | 2022, Online since 01 February 2024, connection on 08 February 2024. URL: http://journals.openedition.org/ranam/364 ; DOI: https://doi.org/10.4000/ranam.364

The text only may be used under licence CC BY-NC-SA 4.0. All other elements (illustrations, imported files) are "All rights reserved", unless otherwise stated.

Teaching Pronunciation with Direct Visual Articulatory Feedback: Pedagogical Considerations for the Use of Ultrasound in the Classroom

Barbara Kühnert♦
Claire Pillot-Loiseau♦♦

Pronunciation is an essential component of communicative competence and hence an important aspect of foreign language (L2) learning and teaching. Several studies have shown that effective communication is difficult when non-native speakers' pronunciation falls below a certain threshold level, even when the vocabulary and grammar are mastered sufficiently (e.g. Derwing and Munro, 2015). Various training methods have been proposed with the aim to improve the pronunciation of suprasegmental units (stress, tone, rhythm, intonation, cf. Gluhareva and Prieto, 2017; Estebas-Vilaplana, 2017) and segmental units (vowels and consonants, cf. Hazan et al., 2005) of an L2. In the present paper we will concentrate on the latter.

Broadly speaking, methods of pronunciation teaching can be categorized as being either more comprehension- / perception-based or articulation- / production-based. An extensive overview about the ongoing debate as to whether perception- or production-based methods are more appropriate and efficient in L2 instruction can be found in Lee et al. (2019). Importantly, different methods of pronunciation teaching provide different forms of supportive and/or corrective feedback. Feedback is information provided by an agent regarding

♦ Barbara Kühnert, *CNRS et Université Sorbonne Nouvelle, UMR 7018, Laboratoire de phonétique et phonologie.*
♦♦ Claire Pillot-Loiseau, *CNRS et Université Sorbonne Nouvelle, UMR 7018, Laboratoire de phonétique et phonologie.*

aspects of a person's performance or comprehension. This agent can be a teacher or parent who can provide corrective information, a peer who can provide an alternative strategy, a book that can provide information for clarification, or a device that can provide information about the teacher's performance (Hattie and Timperley, 2007). Thus, it is an aid provided to improve control or performance of a procedure, as part of the learning process. Effective feedback must first of all be meaningful. It should not rely solely on the learner's perception; it should also allow verification of the accuracy of the response, highlight specific errors, and possibly suggest a remedy (Neri et al., 2002).

Corrective feedback may initially be verbal from the teacher, in the form of implicit corrections (rephrasing, repetition, and request for clarification) or explicit corrections (including explicit correction with or without metalinguistic explanation, metalinguistic cueing, elicitation, or paralinguistic cueing to elicit self-correction from the learner, see Lyster et al., 2013). Corrective feedback can also be auditory by having the learner use the audio-phonatory loop to verify whether the current production of speech is in accordance with his/her intention, either in real time or based on a previously made recording. Moreover, it can be kinesthetic and proprioceptive, such as feeling the position of the articulators—for instance, the rounding of the lips or the exhaled breath on the hand produced by the production of aspirated stop consonants.

Finally, corrective feedback in L2 learning and teaching can also be visual: on the segmental level, it can include direct feedback that provides visual models of articulation, and indirect feedback that uses visualized acoustic information as a means of informing the learner about different articulatory positions. This indirect feedback then provides visual information about the articulation of L2 sounds derived from the acoustic analyses (Bliss et al., 2018). However, phonetic correction through such acoustic information has not always been shown to be effective with a population of learners, as spectrographic information is not always easily interpretable by the student (Ruellot, 2011).

Devices that enable direct visual corrective feedback with respect to learners' articulatory performance include electropalatography (e.g. Bernhardt et al., 2003) which measures lingual contact with a custom-made artificial palate, electromagnetic articulography (e.g. Levitt and Katz, 2010) where small sensors are glued to the speaker's articulators, and lingual ultrasound, which allows for the imaging of the shape and the movements of the tongue during speech in real time (e.g. Stone, 2005; Wilson and Gick, 2006). One of the advantages of the use of ultrasound technology as a direct visual biofeedback is that is less costly and less invasive than other articulatory imaging devices. There are also no known harmful effects associated with the use of standard diagnostic ultrasound instruments (Wilson, 2014).

The purpose of the present contribution is to explore whether diagnostic ultrasound imaging can be used as a valuable tool to improve L2 pronunciation

in a regular adults' classroom setting. Does the provision of visual feedback about correct and incorrect tongue movements and configurations help the learners' understanding and production of critical L2 speech sounds when exposed to a series of short-time training sessions? The study is largely explorative, and hence qualitative, in nature. The primary aim is not to come up with systematic measures as evidence for or against improvement but rather to gain an overall understanding about the feasibility of ultrasound in regular pronunciation teaching, the possible experimental setup or the most appropriate speech material, and to develop guidelines and hypotheses for potential further qualitative research.

In the following, we will first present the basic principles of the ultrasound imaging technique, as well as the advantages and limitations of the device. Then, a short review of previous research on the contribution of lingual ultrasound to the teaching of L2 sounds will be given. In the main part, a pilot study on the learning of two English vowel contrasts by French-speaking learners using lingual ultrasound will be presented. The final section will discuss the results and some of their pedagogical implications.

Principle, Advantages and Limitations of Ultrasound

Ultrasound is a medical imaging technique that uses high frequency sound waves and their echoes. It works in the following way: first, the ultrasound scanner transmits high frequency (1 to 15 MHz) sound pulses into the body, using a probe. Then, waves propagate and touch a boundary between tissues (e.g. soft tissues and bones): some waves are reflected back to the probe, while others continue to propagate. Then, the reflected waves are picked up by the probe and relayed to the machine or computer. The machine calculates the distance of the probe from the boundary using the speed of sound in the tissues and the time of return of each echo. Finally, the ultrasound scanner displays the distance and intensity of the echoes on the computer screen (Wilson, 2014).

Ultrasound is a safe and non-invasive visualization technique which can provide continuous images of the tongue movement during speech production or rest positions. In this way, ultrasound is a biofeedback method which allows for both direct visual feedback on the *execution* of the articulatory movement and direct visual feedback on the *result* of the targeted sound production, i.e. whether a certain articulatory configuration has been reached correctly (or not). It allows to visualize in real time the contour, shape, position and the direction of movements of the tongue, an essential articulator of most of the sounds of the world's languages, and yet an invisible organ during this articulation (Kocjančič Antolík, 2020): this device provides a unique perspective on the

complex deformations of the moving tongue, in the sagittal and coronal planes. It requires minimal training and its startup is fast. Moreover, the images are relatively easy to explain to participants. During teaching, the lingual illustration is provided in addition to tongue proprioception, i.e. vision of the tongue in the articulatory space and morphology of the learner, and not only that of a reference subject (teacher). Besides, it does not necessarily require bulky equipment. Some researchers have used systems with head stabilization techniques which allow for articulatory recordings with subsequent analysis. Other devices, like the one used in our study and most studies in speech therapy (e.g. Preston et al., 2014) consist of a simple program to be installed on a computer and an ultrasound probe which can be inserted into the computer's USB connection. It is hence easily portable and can be used with any laptop but only allows for visualization and the video recording of the tongue movements, without recording the associated sound production. The setup of the system, together with an illustration of the resulting tongue image are given in figure 1.

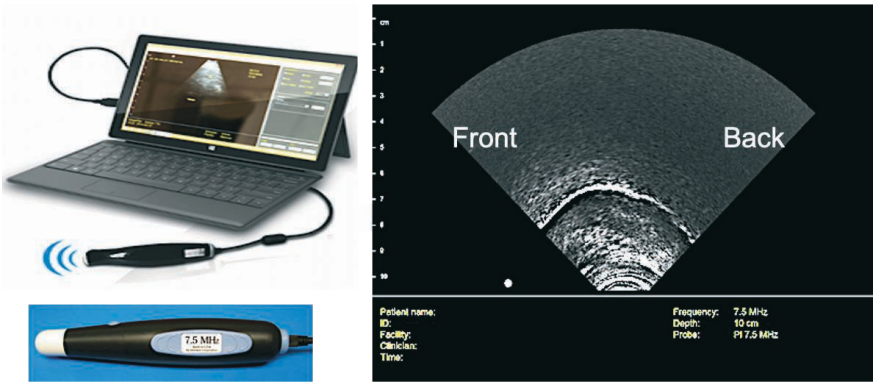


Fig. 1: SeeMore probe (https://seemore.ca/portable-ultrasound-products/pi-7-5-mhz-speech-language-pathology-99-5544-can/) with illustration of portable ultrasound system (left), and display of tongue contour (sagittal plane) during the production of /æ/ in "*cat*" by speaker FR3 (right) (photo by Barbara Kühnert and Claire Pillot-Loiseau).

There are, however, several well-known limitations to the use of lingual ultrasound. One primary shortcoming is that only the tongue is imaged. Thus, the relationship and interaction of the tongue with other structures (palate, pharyngeal walls, or lips) is not obvious as there is no reference structure. Furthermore, the quality of the obtained tongue images depends on several factors: the probe must be of good quality with a sufficient frequency and depth of field (in the present study we used 7,5 MHz microconvex probe with 10 cm depth), and be positioned in a stable way under the oral floor of the subject without deviating from the medio-sagittal plane. The morphology and the size of the speaker can alter the image if the subject is too tall and/or with a notable cervical corpulence. The quality of the image also varies with age and sex of the

subject. Furthermore, ultrasound does not always provide information on the tongue tip due to the possible presence of the mandible and hyoid bones. All these factors can render the comparison between speakers and between sessions a delicate matter.

Use of Ultrasound in Pronunciation Teaching

The use of ultrasound visual feedback is by now a common practice in speech therapy for assessment and remediation of speech errors (Preston et al., 2013). However, thus far few studies have examined the benefits of using visual ultrasound feedback in the acquisition of second language learning. Tsui (2012), for example, trained the lingual configurations of English /l/ and /ɹ/ of six Japanese native speakers, a contrast which is absent in Japanese phonology. After each participant had four 45-minute ultrasound training sessions over a two-week period, all participants were rated by expert listeners as having more accurate productions of both sounds. In Tateishi and Winter's study (2013) who were examining the same contrast with ten native Japanese learners in five separate training sessions of 30 minutes each, only the production of English onset /l/ seemed to improve. Kocjančič Antolík et al. (2019) provided three individual 45-minutes lessons to four Japanese leaners and report some evidence of an enhanced differentiation between French /y/ and /u/, even two months after the end of the training. Roon and colleagues (2020) investigated nine naïve English learners of palatalized Russian consonants and report improvements in both production and discrimination when the subjects were presented with ultrasound videos of the respective sounds.

A comprehensive review of the studies available to date can be found in Bliss et al. (2018). However, the results of all studies at present seem to suggest that ultrasound feedback might be an effective tool for L2 learning. For the time being though, it seems to lend itself best to the training of small groups with learners receiving individual coaching sessions.

Pilot Study

In most regular classes of modern language teaching, individual training sessions—as the ones used in the studies outlined above—are rather unrealistic and would be difficult to implement for reasons of time and resources. The aim of the present pilot study was therefore to explore whether the use of visual ultrasound feedback might also be beneficial and facilitate speech sound acquisition when applied in a series of short-time training sessions with a number of people simultaneously, i.e. in a regular classroom setting. To this end, we conducted a longitudinal study over one semester at the Sorbonne Nouvelle

University in which the use of the portable ultrasound was integrated into regular oral language teaching classes.

Subjects

The participants were monolingual French first year students at the English Department of the Sorbonne Nouvelle University. According to the European Framework of Reference for Languages their level corresponded to either B1 or B2. All subjects had at least eight years of previous experience with English in secondary education. None of them reported any hearing and speech impairments. While we initially recorded 15 French participants, only seven of them attended every single classroom session during the semester (average age: 21 yrs.) and we will therefore concentrate on their results here (5 female, 2 male speakers). A native British subject (male, 23 yrs.) was recorded as a control speaker.

Materials and Recordings

Several studies suggested that ultrasound might be particularly effective for improving learners' vowel articulations, which provide little proprioceptive feedback (e.g. Cleland et al., 2015). The present pilot study therefore focused on the pronunciation training of two vowel contrasts in English known to be problematic for French learners (Huart, 2010): the contrast between tense /iː/ and lax /ɪ/, and the contrast between the front and central open vowels /æ/ and /ʌ/.

Contrary to French, English uses both length and tenseness to distinguish two high front vowels with the close front vowel /iː/ often slightly diphthongized and with prominent spread lips whereas for /ɪ/ the tongue is nearer to the centre and the lips only loosely spread (see figure 2). The French vowel system only contains a single monophthong /i/, and learners have been shown to frequently assimilate the two non-native phones equally well or poorly to their single native category /i/ (Best and Tyler, 2007). The opposition between /æ/ and /ʌ/ is slightly more complex. During the articulation of /æ/ the mouth and lips are slightly more open than for the production of French /ɛ/ and the front tongue body is raised midway between open an open-mid. Moreover, the contrast between the two vowels is prone to a potential interference with the spelling system of both languages. In fact, in the

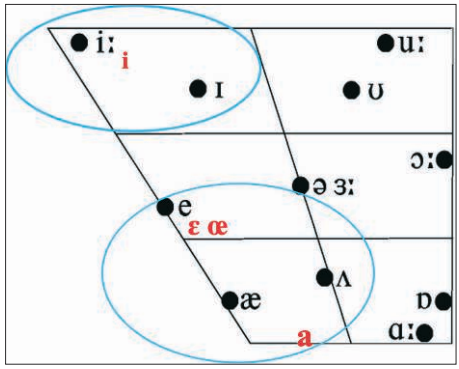


Fig. 2: The IPA vowel diagram for Standard Southern British English (in black) and the four French vowels (in red) used for comparison. The target vowels are indicated by the ellipses.

standard variety of English taught in the French system (commonly referred to as Standard Southern British English), the articulation of /ʌ/ as in "*cup*" is quite similar—even though more centralized—to the articulation of French /a/ as in "*cape*" (Huart, 2010). But while in French /a/ is always associated with the letter <a>, this is never the case in English. Rather the sound /ʌ/ is commonly associated with the letters <u> or <o>. Hence, French learners might have a tendency to add some lip rounding to the production of /ʌ/.

The English target vowels /iː, ɪ, e, æ, ʌ/ together with the French control vowels /i ɛ a œ/ were selected in words with a C_1VC_2 structure with C_1 being the voiced plosive /b/ in all contexts. The English test items were "*beat, bit, bet, bat, butt*", the French test items were "*bide, bettes, batte, boeuf*". All test items were embedded in short carrier sentences. The speech material can be found in table 1. Each speaker produced 10 repetitions of each sentence in randomized order and the experiment was split into two parts, with recordings of the English corpus first, followed by the French one.

Speech Material			
English Test Sentences		French Test Sentences	
Target Vowel		Target Vowe1	
/ iː /	We danced to the **beat** to the music	/ i /	Le film a fait un **bide** en salles.
/ ɪ /	She ate a **bit** of cake.	/ ɛ /	Elle prépare les **bettes** au four.
/ e /	He had a **bet** on the game.	/ a /	Il joue avec sa **batte** en bois.
/ æ /	There is a **bat** in the garden.	/ œ /	Il a mangé du **boeuf** hier soir.
/ ʌ /	He dropped the **butt** of his cigarette.		

Table 1: English (left) and French (right) test sentences used for each target vowel.

The participants were recorded in a soundproof studio twice, once before the beginning of the semester (pre-recording), once four weeks after the end of the semester (post-recording). For two of the speakers, additional recording sessions were included immediately before and immediately after the fourth ultrasound training session.

All acoustic recordings were semi-automatically aligned with the Munich MAUS system (Kisler et al., 2012) and then manually corrected with Praat (Boersma and Weenink, 2014). In order to evaluate the participants' vowel productions, first (F1) and second (F2) formant measurements were automatically extracted at 25%, 50% and 75% into the duration of the vowel segment using a PRAAT script, and then averaged to provide a single value.

Ultrasound Training

During regular language laboratory classes each participant received an individual 10-minute-long ultrasound training using the *SeeMore* portable ultrasound system with a PI 7,5 MHz probe by Interson (see figure 1) on a fortnightly basis, which means that every student had five sessions in total. Each training included an explanatory discussion to incorporate explicit awareness of the tongue movements associated with the target sounds as well as the production of all English vowels to increase the understanding of the personal vocal tract dimensions. Finally, the participants repeatedly practiced the English target vowels, both in isolation and in CVC syllables in different phonetic contexts. During the training period the students were working in pairs together with the teacher in order to encourage cooperation and confidence.

Results

In the following, the speakers' productions will be presented in form of acoustic vowel spaces. While this evidently constitutes a simplified mapping there is nevertheless a well-known correspondence between the articulatory and acoustic vowel space in which F1 is indicated on the vertical axis corresponding to tongue height (higher values for lower, more open vowels) and F2 is indicated on the horizontal axis corresponding to tongue advancement (lower values for back vowels). All the data are represented in form of Fisher confidence ellipses (90%) in order to also assess the variability of the speakers' productions.

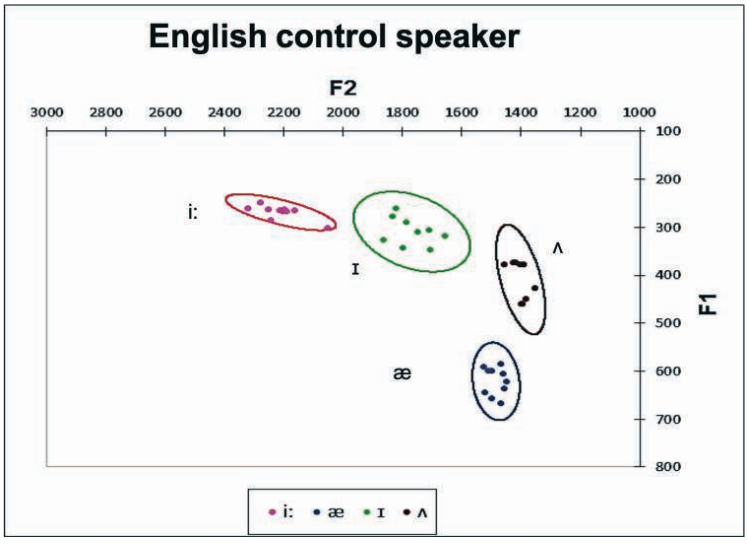


Fig. 3: F1/F2 ellipse plot of English target vowels as produced by the British control speaker.

Figure 3 shows the native speaker's vowel productions. As can be seen, the realizations of all four target vowels are clearly distinct from each other without there being any overlap between the formant values of the different vowel articulations. The vowel /iː/ is articulated further forward and with a higher tongue configuration than its centralized lax counterpart /ɪ/. During the articulation of the vowel /æ/ the main tongue constriction is lower and situated further in the front cavity than the more centralised and higher tongue constriction of the vowel sound /ʌ/. As such, the data closely resemble the data reported in the literature on the formant values of Standard Southern British English monophthongs (e.g. Wells, 1962; Deterding, 1997).

Turning to the French learners' distinctions between the tense and lax high front vowel, two different patterns can be observed. First and foremost—and somewhat surprisingly in light of the subjects' previous experience of English—none of the French learners showed a clear difference between the two vowels in the pre-recording, with both vowel productions being very similar to the realizations of the French counterpart /i/, as illustrated in figure 4 by the data of Speaker FR1 (top) and FR4 (bottom). In other words, the words *beat* and *bit* were virtually indistinguishable. A similar overlap in terms of vowel quality between the two categories by French learners of English has been reported in the literature before (cf. Krzonowski et al., 2018).

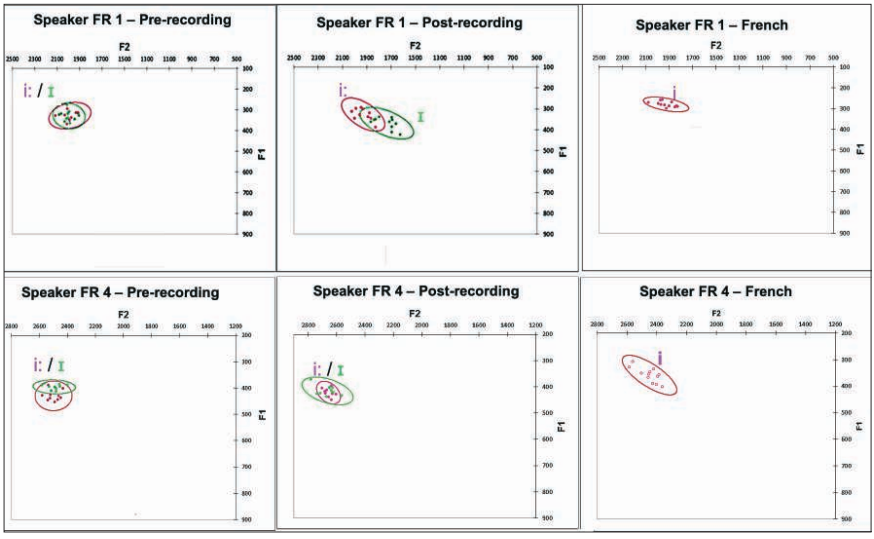


Fig. 4: F1/F2 ellipse plot of /iː/ and /ɪ/ as produced by speaker FR1 (top) and FR4 (bottom) in the pre-training (left) and post-training (middle) condition, together with the control productions of French /i/ (right).

However, after the ultrasound training sessions three of the seven subjects showed an increased tendency to distinguish between tense /iː/ and lax /ɪ/ in the post-recording as illustrated by the data of speaker FR1. While a certain

overlap between the two vowels still can be observed, there is notwithstanding an apparent trend for an improved differentiation between the two vowel articulations, with /ɪ/ being now produced slightly more centralized and more open than /iː/. For the remaining four speakers, though, no clear difference between the two vowel sounds emerged, neither prior- nor post-training as can be seen in the productions of speaker FR4.

The French learners' productions of the vowels /æ/ and /ʌ/ display a similar twofold pattern. Three of the seven speakers positively distinguished the two vowels from the outset, with little change between their productions prior or post ultrasound training as shown in figure 5 (speaker FR3). The main effect is some reduced variability in the productions at the end of the semester. It should also be noticed that—even though the distinction between the two sounds is maintained—the vowel /æ/ is considerably more centralised than commonly produced by native speakers. The difference seems to be primarily accomplished by a difference in tongue height rather than by a difference in tongue height and tongue fronting, with a similar acoustic representation of English /ʌ/ and French /a/.

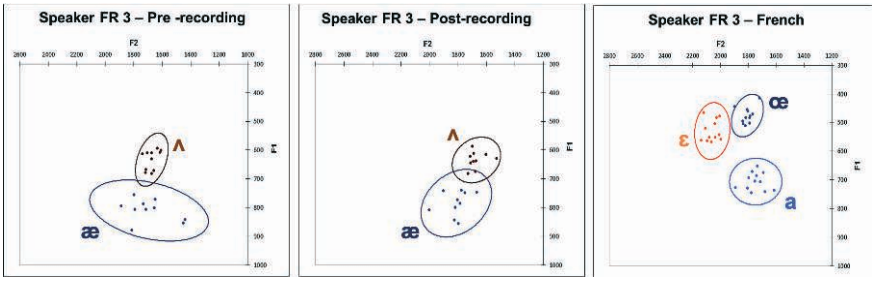


Fig. 5: F1/F2 ellipse plot of /æ/ and /ʌ/ as produced by speaker FR3 in the pre-training (left) and post-training (middle) condition, together with the control productions of French /ɛ, a, œ/ (right).

The remaining four speakers realised /æ/ with an even more distinct posterior position than /ʌ/. As a comparison with the French control data suggest (see figure 6, speaker FR1), tongue advancement for the vowels was reversed and there seems to be a noticeable confusion between English /æ/ and French /a/, and English /ʌ/ and the French rounded vowel /œ/. While the F1/F2 plots do not provide us with any concrete information about lip rounding, the close similarity of the acoustic values—and hence presumed underlying tongue configurations—between the French vowel productions /œ/ and the English vowel productions /ʌ/ seem to suggest that some lip rounding for /ʌ/ might have occurred. Lip rounding lengthens the vocal tract and as a consequence the main constriction location of the tongue moves further forward as a compensatory contribution (Lee et al., 2019). When listening to the recordings, the two vowels in the English word *butt* and French word *boeuf* sound indeed were much alike. Again, this observation matches the data reported elsewhere. For instance, a similar distinction and

confusion between the production of English /ʌ/ and French /œ/ have been found in the acoustic study of 48 French L2 learners of English by Krzonowski et al., 2018.

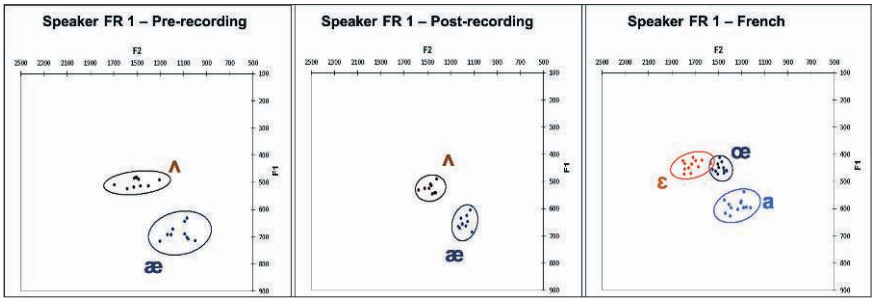


Fig. 6: F1/F2 ellipse plot of /æ/ and /ʌ/ as produced by speaker FR1 in the pre-training (left) and post-training (middle) condition, together with the control productions of French /ɛ, a, œ/ (right).

Finally, an example of the immediate impact of an ultrasound training session is provided in F3's data in figure 7. To recall, here the speaker has been recorded at the beginning of the semester, immediately after the training session in week 4 and one month after the end of the semester. As can be observed, subject F3 displays a more accurate distinction between the two front closed vowels when recording the target vowels directly after the ultrasound practice. However, as the data in the post-recording reveal, there was no lasting effect of an increased differentiation between the two vowels. Two months after the training sessions ended the subject's articulations of /i:/ and /ɪ/ were once more almost identical.

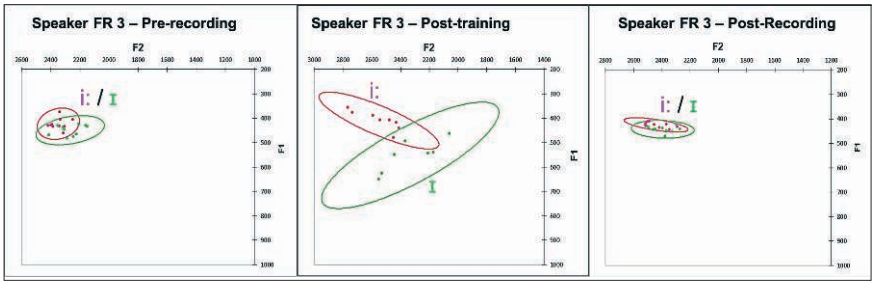


Fig. 7: F1/F2 ellipse plot of /i:/ and /ɪ/ as produced by speaker FR3 in the pre-training condition (left), immediately after the ultrasound training (middle), and in the post-training condition (right).

Discussion and Pedagogical Implications

This exploratory study investigated the effects of ultrasound visual feedback for improving the pronunciation of language learners in a regular classroom setting. Each participant received five training sessions on a fortnightly basis, addressing each of the four target sounds produced in isolation and in context. To sum up, the results of the study showed that the effects of the short-time

ultrasound training were speaker- and/or sound-dependent. In the recordings one months after the training, three speakers showed a lasting effect towards an increased differentiation between the vowels /iː/ and /ɪ/ in terms of F1/F2 values while for the remaining speakers no change in their production patterns could be observed. The distinction between /æ/ and /ʌ/—even though some speakers' data revealed a certain trend in the right direction and a decrease in variability—did not show any prominent enhancement between the initial and the post-training sessions.

There are several observations and considerations to be taken into account which could partially explain the performance of the learners. The present study differs in a number of ways from prior research on the use of ultrasound in second language learning. First, previous studies used between 30 to 45 minutes training time for the sounds under consideration, delivering individual coaching sessions (e.g. Tsui, 2012; Kocjančič Antolík et al., 2019). The length of the training time therefore may not have been enough for all subjects to demonstrate an effect and the duration of the coaching session might be a factor that should be re-considered and, if possible, extended. Nevertheless, that three participants responded to the very short 10 minutes interventions in the present study with an enhanced distinction between the tense and lax high front vowel still after two months suggests that in some cases even short visual biofeedback might be helpful to successfully establish improved pronunciation accuracy.

Likewise, studies in speech therapy suggest (e.g. Allen, 2013) that increasing the frequency of training intervention and having multiple training sessions in a short time period might lead to a better outcome than having the same number of sessions over a longer time period, even though the exact frequency and intensity of the sessions is still subject to debate (Preston et al., 2014). We were partially restricted by our class structure but ultimately fortnightly sessions should be replaced by weekly or even twice a week interventions.

Another consideration from the study concerns the speech material. Previous studies targeted either selected consonants with a clear anchor point in the vocal tract (for example, /l/ versus /r/, Tateishi and Winter, 2013) or vowel contrasts with rather distinct tongue configurations in the vocal tract, as for example the contrast between the two rounded vowels /y/ (front) and /u/ (back) in French (Kocjančič Antolík et al., 2019). In particular the articulatory differences in the tongue configurations between the vowels /æ/ and /ʌ/ might be too subtle to be captured and interpreted by the participants in the visual ultrasound display and hence could not be incorporated into their own productions. Further studies may consider first starting with vowel articulations that are more distinct in terms of lingual articulations, before tackling finer differences.

The performance related to the production of the vowels /æ/ and /ʌ/ draws the attention to a further aspect of the study. As noted previously, four of the subjects realized the sound /ʌ/ with almost identical acoustic values to the French

counterpart /œ/, and with a clear auditory perceptible degree of lip rounding. This is a well-known production strategy of French learners of English, in particular in the early phases of L2 acquisition. All our participants had at least 8 years of previous experience with English at secondary education and as a consequence already well-established pronunciation patterns in their L2. Some of those patterns may in fact have turned into fossilized pronunciation errors over time as a result of the incorrect acquisition of pronunciation of the L2, usually affected by the L1 (Wei, 2008). While researchers disagree whether and how those fossilized pronunciations are 'undoable' (for an overview, see Demirezen and Topal, 2015), it is evident that chronic articulation mistakes cannot easily be solved. In the case of French learners with a strong preceding exposure to English, as most likely some of our participants, we find ourselves therefore not only in a situation of acquisition and learning but sometimes rather in a situation of remediation, in which mature production patterns have to be unlearned first. Thus, short-time ultrasound interventions might have a stronger impact on less advanced language learners.

A good example to this effect can be seen in data of speaker FR3 for whom we compared the sound productions immediately after the ultrasound training and four weeks after the end of the study. While the speaker did not show any differentiation between the vowels /i:/ and /ɪ/ in the initial baseline recording, there was some evidence of refined acquisition directly after the visual imaging session with an increase in the distinctiveness between the two vowels. However, there was no evidence of retention and learning in the long term. One month after the training the speaker had fallen back to the initial pronunciation habits. How to move from acquisition to retention and learning is a crucial but still largely unresolved issue, not only for visual ultrasound instruction but any pronunciation teaching method.

Leaving aside the direct and / or long-term impact, the use of visual ultrasound biofeedback in the classroom clearly contributed the students' awareness and understanding of their proper articulatory behaviour. While the students had been exposed to articulatory diagrams or animations during lessons in the past, the visual experience of seeing the movements of their *own* tongue and hearing the consequences in the corresponding sound productions far exceeded their previous descriptive knowledge. The work in pairs with another student which allowed them to comment on and correct possible deviant pronunciations contributed to this understanding. Hence, in general the sessions were much appreciated by the students and, if nothing else, prominently increased their interest in the subject matter.

Conclusion

In conclusion, the present pilot study aimed at evaluating the feasibility and potential benefits of using ultrasound in a regular classroom setting. Due to the small number of subjects and the absence of a control group of speakers, any results drawn from the study must, of course, be considered very tentative, and more work is necessary to prove the validity of our interpretations. Nevertheless, our pilot study suggests that the use of the new portable ultrasound devices could be an enriching experience for pronunciation learning, in particular for students at the earlier stages in the acquisition process of a new language, with more frequent training sessions than used in our pilot study, and with speech material that allows for clear articulatory distinctions.

Ultrasound imaging is part of an explicit and analytical method of foreign language teaching but, of course, this approach does not prevent or replace the practice of other more global teaching methods oriented towards auditory perception. Any method that increases the positive attitude of the students towards pronunciation teaching classes and their interest in willing to improve their own pronunciation can only be helpful.

References

• ALLEN, M. (2013): "Intervention efficacy and intensity for children with speech sound disorder", *Journal of Speech Language and Hearing Research* 56, p. 865-877.

• BERNHARDT, B., GICK, B., BACSFALVI, P., & ASHDOWN, J. (2003): "Speech habilitation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners", *Clinical Linguistics and Phonetics* 17.3, p. 199-216, https://doi.org/10.1080/0269920031000071451.

• BEST, C. & TYLER, M. (2007): "Nonnative and second-language speech perception: commonalities and complementarities", in Munro M. & Bohn, O.-S. (eds.), *Language experience in second language speech learning: In honor of James Emil*, John Benjamins Publishing Company, p. 13-34.

BLISS, H., ABEL, J. & GICK, B. (2018): "Computer-assisted visual articulation feedback in L2 pronunciation instruction: A review", *Journal of Second Language Pronunciation* 4.1, p. 129-153, https://doi.org/10.1075/jslp.00006.bli.

• BOERSMA, P. & WEENINK, D. (2014): Praat: doing phonetics by computer [Computer program]. Version 6.0.37. Retrieved from http://www.praat.org/.

• CLELAND, J., SCOBBIE, J. M. & WRENCH, A. (2015). "Using ultrasound visual biofeedback to treat persistent primary speech sound disorders", *Clinical Linguistics and Phonetics* 29.8-10, p. 575-597.

• DEMIREZEN, M. & TOPAL, H. (2015): "Fossilized pronunciation errors from the perspectives of Turkish teachers of English and their implications", *Procedia - Social and Behavioral Sciences* 199, p. 793-800.
• DERWING, T. & MUNRO, M. (2015): *Pronunciation Fundamentals: Evidence-based perspectives for L2 teaching and research*, John Benjamin Publishing Company, 208 p.
• DETERDING, D. (1997): "The formants of monophthong vowels in Standard British English Pronunciation", *Journal of the International Phonetic Association* 27.1-2, p. 47-55.
• ESTEBAS-VILAPLANA, E. (2017): "The teaching and learning of L2 English intonation in a distance education environment: TL_ToBI vs. the traditional models", *Linguistica* 57.1, p. 73-91.
• GLUHAREVA, D. & PRIETO, P. (2017): "Training with rhythmic beat gestures benefits L2 pronunciation in discourse-demanding situations", *Language Teaching Research* 21.5, p. 609-631.
• HATTIE, J. & TIMPERLEY, H. (2007): "The power of feedback", *Review of educational research* 77.1, p. 81-112, https://doi.org/10.3102/003465430298487.
• 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* 47.3, p. 360-378.
• HUART, R. (2010): *Nouvelle Grammaire de l'Anglais oral*, Paris, Ophrys, 258 p.
• KISLER, T., SCHIEL, F. & SLOETJES, H. (2012): "Signal processing via web services: the use case webmaus", *Proceedings Digital Humanities 2012*, Hamburg, p. 30-34.
• KOCJANČIČ ANTOLÍK, T., PILLOT-LOISEAU, C. & KAMIYAMA, T. (2019): "The effectiveness of real-time ultrasound visual feedback on tongue movements in L2 pronunciation training: Japanese learners improving the French vowel contrast /y/-/u/", *Journal of Second Language Pronunciation* 5.1, p. 72-97, https://doi.org/10.1075/jslp.16022.
• KOCJANČIČ ANTOLÍK, T. (2020): "Ultrasound Tongue Imaging in Second Language Learning", *Studie z aplikované lingvistiky* 11.1, p. 109-119.
• KRZONOWSKI, J., PELLEGRINO, F. & FERRAGNE, E. (2018): "Étude acoustique de la production de voyelles de l'anglais par des apprenants francophones", in *Actes des XXXII^e Journées d'Études sur la Parole*, 4-8 June 2018, Aix-en-Provence, France, p. 525-531.
• LEE, B., PLONSKY, L. & SAITO, K. (2019): "The effects of perception- vs. production-based pronunciation instruction", *System* 88, p. 1-13.
• LEVITT, J. S. & KATZ, W. F. (2010): "The effects of EMA-based augmented visual feedback on the English speakers' acquisition of the Japanese flap: A perceptual study", in *Eleventh Annual Conference of the International Speech Communication Association* Interspeech 2010, Makuhari, p. 1862-1865, https://doi.org/10.21437/Interspeech.2010-538.

• LYSTER, R., SAITO, K. & SATO, M. (2013): "Oral corrective feedback in second language classrooms", *Language Teaching* 1, p. 1-40, https://doi.org/10.1017/S0261444812000365.
• NERI, A., CUCCHIARINI, C. & STRIK, H. (2002): "Feedback in computer assisted pronunciation training: Technology push or demand pull?", *ICSLP-2002*, Denver, USA, p. 1209-1212.
• PRESTON, J. L., BRICK, N. & LANDI, N. (2013): "Ultrasound biofeedback treatment for persisting childhood apraxia of speech", *American Journal of Speech-Language Pathology* 22, p. 627-643, https://doi.org/10.1044/1058-0360(2013/12-0139).
• PRESTON, J.L., MCCABE, B., RIVERA-CAMPOS, A., WHITTLE, J., LANDRY, E. & MAAS, E. (2014): "Ultrasound visual feedback treatment and practice variability for residual speech sound errors", *Journal of Speech Language and Hearing Research* 57.6, p. 2102-2115.
• ROON, K.D., KANG, J. & WHALEN, D.H. (2020): "Effects of ultrasound familiarization on production and perception of nonnative contrasts", *Phonetica* 77.5, p. 350-393.
• RUELLOT, V. (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,* Sept. 2010, Ames, IA, Iowa State University, p. 199-213.
• STONE, M. (2005): "A guide to analysing tongue motion from ultrasound images", *Clinical Linguistics and Phonetics* 19.6-7, p. 455-501, https://doi.org/10.1080/02699200500113558.
• TATEISHI, M. & WINTER, S. (2013): "Does ultrasound training lead to improved perception of a nonnative sound contrast? Evidence from Japanese learners of English", in LUO, S. (ed.), *Proceedings of the 2013 Annual Conference of the Canadian Linguistic Association,* Victoria, BC, p. 1-15.
• TSUI, H. M. (2012): "Ultrasound speech training for Japanese adults learning English as a second language", *Unpublished MSc thesis,* University of British Columbia. Retrieved from https://enunciate.arts.ubc.ca/research-and-case-studies/other-research/.
• WELLS (1962): "A study of the formants of the pure vowels of British *English*", *Unpublished MA thesis,* University of London. Retrieved from https://www.phon.ucl.ac.uk/home/wells/formants/index.htm.
• WILSON, I. (2014): "Using ultrasound for teaching and researching articulation", *Acoustical Science and Technology* 35.6, p. 285-289, https://doi.org/10.1250/ast.35.285.
• WILSON, I. & GICK, B. (2006): "Ultrasound technology and second language acquisition research", *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference*, p. 148-152.
• WEI, X. (2008): "Implication of IL fossilization in second language acquisition", *English Language Teaching* 1.1, p. 127-131.