International Conference
ICT for Language Learning

Ultrasound Articulatory Training for Teaching Pronunciation of L2 Vowels

Bianca Sisinni[1], Sonia d'Apolito[1], Barbara Gili Fivela[1], Mirko Grimaldi[1]

Abstract

The paper describes a pilot study on the contribution of ultrasound tongue imagining in second language (L2) vowel production training. A ultrasound (US) training has been performed on American English (L2) contrast /ɑ-ʌ/ (e.g., cop-cup) by native speakers of Italian (L1). Three experimental subjects (ES) received a one-hour US training session while three control subjects (CS) did not. Both the ES and the CS were recorded in the pre- and post-training sessions. Acoustic analysis of their productions (F1, tongue height; F2, tongue backness) revealed an improvement in terms of F1 and/or F2 for the ES only. Thus, US training has a beneficial effect on the pronunciation of L2 /ɑ-ʌ/ by Italian learners.

1. Introduction

Second language (L2) learning is largely influenced by the native language (L1) and, indeed, L2 adult learners often fail to correctly perceive and produce sounds that do not occur in their L1 phonological system. This is particularly true in formal context, where L2 teachers are often L1 native speakers who offer non-native pronunciation of L2 segments [1].

Pronunciation teaching methods provide acoustic or articulatory feedback about learners' productions [2]. The former may be not suitable, as some learners may fail to map acoustic information onto articulatory movements [3]; the latter may be very useful when learners face with articulatory movements that cannot be seen directly, as for central/back vowels [3]. In fact, movies displaying teacher's or a native speaker's tongue movements during speech or visual biofeedback of one's own tongue moving in real-time on the screen can be very successful [4]. For this reason, US represents a useful tool for teaching pronunciation allowing (i) to directly see teachers' or native speakers' articulation and (ii) to get a visual biofeedback during L2 production as in [5]. Moreover, US has been found to help improving pronunciation even after very short (30 min.) training [6].

In the present study, we focus on the English contrast /ɑ-ʌ/ (e.g., cop-cup) which is difficult to perceive and produce by Italian leaners [7,8]. /ɑ-ʌ/ are characterized by both tongue dorsum and root movements, easily visible with US (sagittal view). They differ in tongue height and backness since /ɑ/ shows a more retracted tongue root and a lower tongue dorsum than /ʌ/ which, in turn, shows a more anterior tongue root and a higher tongue dorsum (Figure 2). We take also into account their orthographic representation since it represents an additional difficulty. As users of phonologically transparent orthography, native Italians tend to rely on orthographic form and to mispronounce English written words which, in turn, are phonologically opaque [9].

The aim of this work is twofold: (i) to explore the efficacy of articulatory instructions and ultrasound technique; (ii) to observe learners' improvement in the realization of /ɑ-ʌ/ even after a short training session. We hypothesize that US can be very helpful for teaching the English contrast /ɑ-ʌ/ to adult Italian learners, who can improve their pronunciation even after a short training.

[1] University of Salento, Centro di Ricerche Interdisciplinare sul Linguaggio (CRIL), Laboratorio diffuso di ricerca interdisciplinare applicata alla medicina (DREAM), (Italy)

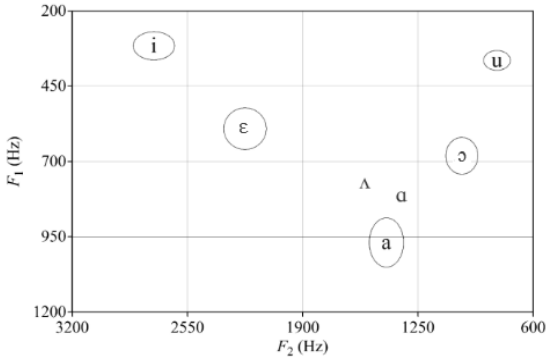

Figure 1: Salento Italian acoustic vowel space (ellipses) and American English vowels /ɑ-ʌ/ produced by female native speakers (ellipses=±1 SD).

2. Method

2.1 Subjects

Six female subjects (mean age: 23, s.d. 0,63) from Salento (southern Apulia) participated. They are monolingual speakers, have never been in a foreign country for longer than a month and have started studying English as a foreign language at the mean age of 8,5. Their exposure to English was limited to school setting with mainly L1-accented teachers. Three of them were the experimental subjects (ES) i.e., SPK1, SPK2, SPK3, and the remaining three were the control subjects (CS), i.e., SPK4, SPK5, SPK6. One female native speaker of American English (AES; 21 years old, Oregon) was recorded to obtain native speaker US data to be presented during the training.

2.2 Procedure

The ES individually received a one-hour of audio-articulatory training where the US machine (Toshiba Aplio XV) was used to offer a real-time biofeedback on the tongue position [4]. US videos showing the AES's productions were used as a visual model.

First, the grapheme-to-sound correspondences of target AE vowels and the phonetic differences between them (we focused on vowel spectral properties and neutralized the duration difference as in [10]), as well as with the closest native vowels, i.e., /a-ɔ/, were explained. The graphic representation, in the F1 (tongue height)/F2 (tongue backness) acoustic space, of the Italian vowels and of the trained AE vowels (Figure 1) were provided with verbal instructions their about articulation. Then, the US training began with videos of AE /ɑ-ʌ/ presented in isolation (Figure 2), successively in CVC words and finally in real sentences. Once the subjects proved they have detected the acoustic and articulatory differences between the target vowels (i.e., tongue root and dorsum) and between them and the native vowels /a-ɔ/, they started to practice with the US probe under their own chin until they thought their production matched the given example.

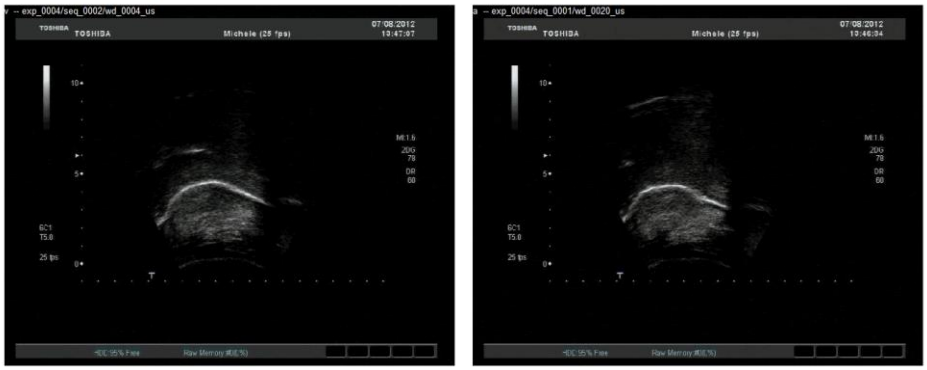

Figure 2: US frames from the US recording of AE /ʌ/ (left) and /ɑ/ (right) produced by the AES.

Both ES and CS recorded their L1 and L2 productions in pre-and post-test sessions (before/after the training session by the ES). The CS did the recordings with a one-hour interval (no training). The US machine was used to acquire the US video with the synchronized audio signal by means of an external a/v analog-to-digital acquisition card. A convex probe, placed under the subjects' chin was

used to get images of the tongue on the midsagittal plane (25Hz). A special set was used to stabilize and secure the head of the subjects and to fix the probe.
The L1 native vowels /i, ɛ, a, ɔ, u/ were produced 12 times in /pV_1pV_2/ words and pseudo-words (V_1 is one of the five native vowels and V_2 is /i/ or /a/) in the carrier sentences “Dicevi pVpi in su” or “Diceva pVpa a Ken”. The L2 corpus consisted of 12 repetitions of AE /ɑ-ʌ/ in /pVp/ words (V is /ɑ/ or /ʌ/) in the carrier sentence “I see pVp inside”. Acoustic data (22050 Hz, 16 bits) were segmented offline and the first three formants were measured in the central 40% of the vowel duration [11]. The average F1 and F2 values elicited in the pre- and the post-test were plotted for each subject on a Cartesian F1-F2 plane (Figure 3-4) and a series of independent t-test was run ($p<0.05$) to compare: (i) /ɑ-ʌ/, respectively, in pre- *vs* post-test; (ii) /ɑ-ʌ/ in pre- and post-test *vs* L1 /a-ɔ/ respectively.

3. Results

ES’s productions are showed in Figure 3. In the pre-test SPK1 realizes L2-/ʌ/ as L1-/a/ even if slightly lower [$p<0.05$]. In the post test, she raises and advances the tongue following the correct direction (Figure 3(a)) [/ʌ/-/a/ F1 and F2 $p<0.05$; /ʌ/-/ʌ/ pre-test F1 and F2 $p<0.05$]. With regard to L2-/ɑ/, she produces it very close to L1-/ɔ/ in the pre-test, even if lower and more anterior [F1 and F2 $p<0.05$], while in the post-test she moves away from both L1-/ɔ/ and L2-/ɑ/-pre-test, lowering and advancing the tongue [/ɑ/-/ɔ/ F1 $p<0.05$; F2 $p=0.05$; /ɑ/-/ɑ/ pre-test F1 and F2 $p<0.05$]. Thus, SPK1 seems to produce both L2 sounds adequately after the training.
SPK2 does not produce any difference between L2-/ʌ/ and L1-/a/ in the pre-test while, after the training, she raises and advances the tongue towards the correct position (Figure 3(b)) [/ʌ/-/a/ F1 and F2 $p<0.05$; /ʌ/-/ʌ/ pre-test F1 and F2 $p<0.05$]. Like SPK1, she produces L2-/ɑ/ close to L1-/ɔ/, though as lower and more anterior [F1 and F2 $p<0.05$] while, after the training, just as more anterior than L1-/ɔ/ [/ɑ/-/ɔ/ F2 $p<0.05$; /ɑ/-/ɑ/ pre-test F2 $p<0.05$]. SPK2, therefore, seems to correctly produce L2-/ʌ/ but not L2-/ɑ/. However, she is able to distinguish her production from L1-/ɔ/.
SPK3 realizes both L2-/ʌ/ and /ɑ/ nearly as L1-/a/, but significantly more open [F1: /ʌ/-/a/ and /ɑ/-/a/ $p<0.05$] in the pre-test. In the post-test, she realizes L2-/ʌ/ as different from L1-/a/, lowering and also advancing the tongue, and form L2-/ʌ/ in pre-test only for advancing (Figure 3(c)) [/ʌ/-/a/ F1 and F2 $p<0.05$; F2 /ʌ/-/ʌ/ pre-test $p<0.05$]. Concerning L2-/ɑ/, she lowers the tongue with regard L1-/a/ and L2-/ɑ/ [F2 /ɑ/-/a/ and /ɑ/-/ɑ/ pre-test $p<0.05$]. Consequently, SPK3 is able to move away from native /a/ for both L2 sounds even if in a partial correct way.

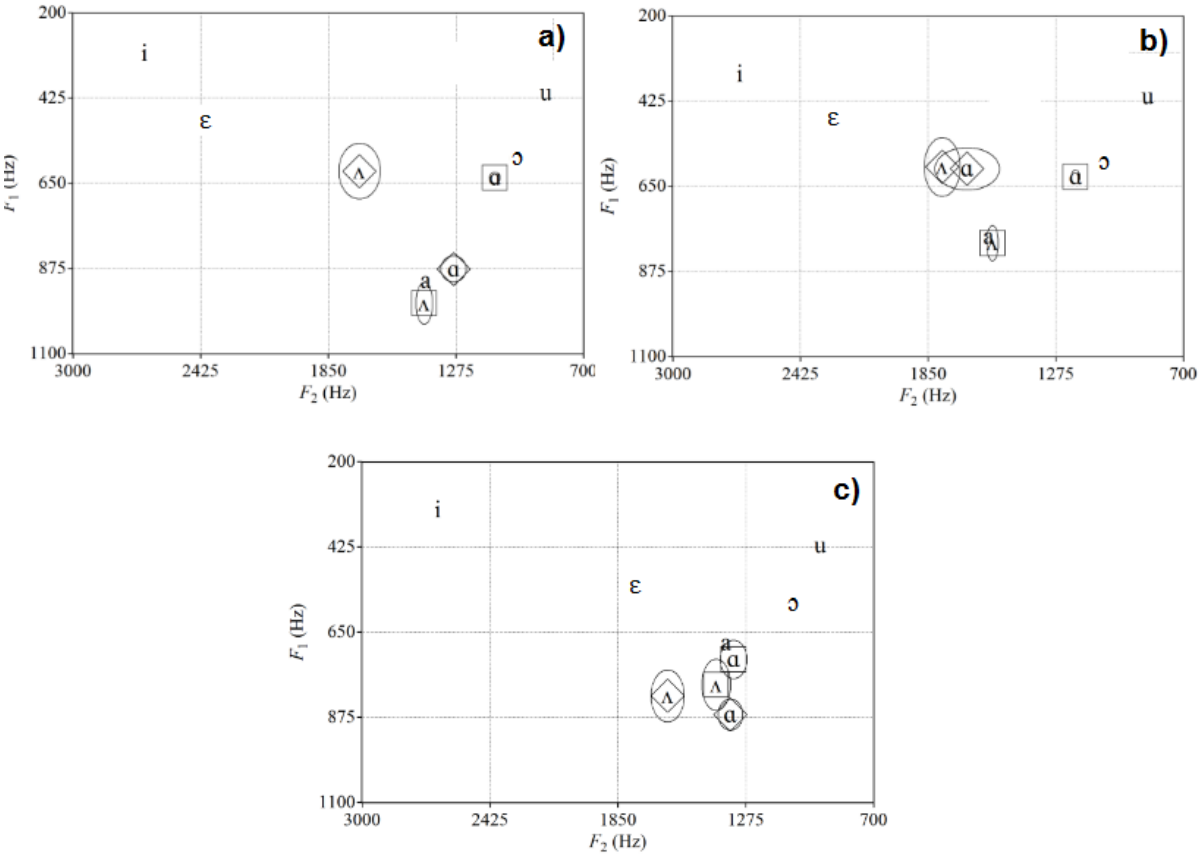

Figure 3. Native and L2 vowels produced by SPK1 (a), SPK2 (b), SPK3 (c) in the pre-test (square) and the post-test (diamond) (ellipses=±1 SD)...

Figure 4 shows the productions of the CS. SPK4 and SPK5 produce L2-/ʌ/ differently from and in-between L1-/a/ [SPK4: F1 $p<0.05$; F2 $p=0.05$; SPK 5: F1 and F2 $p<0.05$] and /ɔ/ [SPK4 and SPK5: F1 and F2 $p<0.05$] confirming this position also in the post-test [SPK4 and SPK5 /ʌ/-/a/: F1 and F2 $p<0.05$; /ʌ/-/ɔ/ F1 and F2 $p<0.05$], and rising the tongue also with regard L2-/ʌ/-pre-test [SPK4 and SPK5: F1 $p<0.05$]. On the contrary, they produce L2-/ɑ/ close to L1-/ɔ/, even if SPK4 realizes it as slightly more anterior and lower in the pre-test [SPK4: F1 and F2 $p<0.05$], and both lower and more advanced in the post-test [SPK4 and SPK5: F1 and F2 $p<0.05$]. Nevertheless, the target position is not reached. On the whole, both SPK4 and SPK5 do not articulate the L2 sounds drastically differently between the two sessions. SPK6 realizes L2-/ʌ/ close to L1-/a/ even if significantly lower in both pre-test and post-test [$p<0.05$]. She produces L2-/ɑ/ as lower and more anterior than L1-/ɔ/ in the pre-test [F1 and F2 $p<0.05$], while just as slightly more anterior in the post-test [F2 $p<0.05$]. SPK6 never differentiates significantly the L2 vowels in the two test sessions (complete statistical analysis at http://www.cril.unisalento.it/papers/Supplemental_Material_ICT_2016_Sisinni_et_al.pdf).

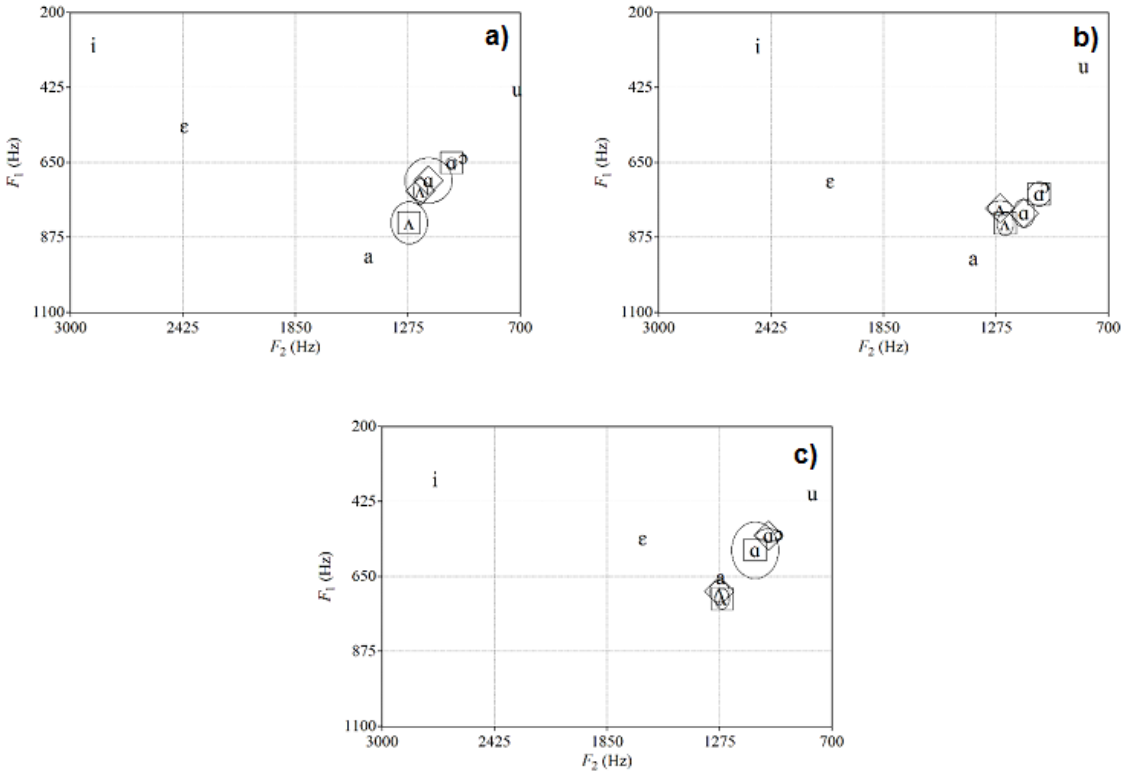

Figure 4. Native and L2 vowels produced by SPK4 (a), SPK5 (b) and SPK6 (c) in the pre-test (square) and the post-test (diamond) (ellipses=±1 SD).

4. Discussion and conclusion

Although AE /ɑ-ʌ/ are difficult to produce for Italian speakers [8], the US training proves to be very helpful for improving learners' productions of the AE contrast. The ES, who received a one-hour audio-articulatory training, by using US as a visual model and as visual real-time biofeedback, are able to differentiate in the post-test the L2 vowels from each other, as well as from the native ones, more than the CS do.

SPK1 and SPK2 initially produce L2-/ʌ/ as L1-/a/ and L2-/ɑ/ as L1-/ɔ/. After the training session, they are able to shift them away from the native vowels, moving the tongue in the correct direction. This is evident especially for L2-/ʌ/, for which both SPK1 and SPK2 rise and advance their tongue, while is less evident for L2-/ɑ/ since SPK1 only is able to lower the L2 vowel in the acoustic-articulatory space.

Special attention needs to be paid to SPK3. Differently from the previous subjects, her acoustic vowel space between /a/-/ɛ/ and /ɔ/ is restricted because her L1-/a/ is very high. If in the pre-test she realizes both L2 /ɑ-ʌ/ nearly as L1/a/, in the post-test she moves them in the lowest and empty part of her acoustic space, in order to differentiate them from native /a/ as well as from any native vowels.

On the contrary, the control subjects, who did not receive any training, do not show improvements or reliable L2 vowels' shifting from pre-test to post-test, even if they show some differences between L2 and L1 vowels.

Moreover, all speakers' productions (but not SPK3) of L2-/ɑ/ suggest the existence of a strong influence of orthography since it is read as [o] as the grapheme is <o>.

The results show that pronunciation training by means of US yields an improvement in L2 vowel production after one-hour session, though they confirm that for L2 adult learners it is not easy to finely control tongue body gesture [12]. Moreover, they confirm that it is important to consider the starting point of each student in order to adjust the L2 teaching process and that, even after receiving formal instruction in L2 English for many years, Italians tend to rely on orthographic forms when reading L2 words and pseudo-words [9].

References

[1] Best C. T., Tyler M. D. “Nonnative And Second Language Speech Perception: Commonalities And Complementarities”, Second Language Speech Learning: The Role Of Language Experience In Speech Perception And Production, Amsterdam, John Benjamins, 2007,13-34.
[2] Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. “The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds”, JASA, 138(2), 2015, 817-832.
[3] Wilson, S. M., Gick, B. “Ultrasound Technology And Second Language Acquisition Research”, Proceedings Of The 8th Generative Approaches To Second Language Acquisition Conference, 2006,148–152.
[4] Wilson I. “Using Ultrasound For Teaching And Researching Articulation”, Acoustical Science & Technology, 35(6), 2014, 285-289.
[5] Pillot-Loiseau, C., Antolík T. K., Kamiyama T. “Contribution Of Ultrasound Visualization To Improving The Production Of The French /y/-/u/ Contrast By Four Japanese Learners”, Phonetics, Phonology, Languages In Contact. Contact: Varieties, Multilingualism, Second Language Learning, 2013, 86-89.
[6] Gick, B., Bernhardt, B.M., Bacsfalvi, P., Wilson, I., “Ultrasound imaging applications in second language acquisition”. Phonology and Second Language Acquisition, 2008, 309-322.
[7] Escudero, P., Sisinni, B., Grimaldi M. “The Effect Of Vowel Inventory And Acoustic Properties In Salento Italian Learners Of Southern British English Vowels”, JASA”, 135(3), 2014, 1577-1584.
[8] Flege, J., Mackay, I., & Meador, D. “Native Italian Speakers’ Production And Perception Of English Vowels”, JASA, 106, 1999, 2973-2987.
[9] Bassetti, B. & Atkinson, N. “Effects Of Orthographic Forms On Pronunciation In Experienced Instructed Second Language Learners”, Orthographic Effects In Second Language Phonology. Special Issue. Applied Psycholinguistics, 36(1), 2015, 67-91.
[10] Escudero P., Benders T., Wanrooij K. “Enhanced Vowel Distributions Facilitate The Learning Of Second Language Vowels”, JASA, 130, 4, 2011, El206-El212.
[11] Chládková K., Escudero P., Boersma P. “Context-Specific Acoustic Differences Between Peruvian And Iberian Spanish Vowels”, JASA, 130, 2011, 416–428.
[12] Ouni S.,“Tongue Control and its implication in Pronunciation Training”, Computer Assisted Language Learning, 27(5), 2014, 439-453.