Fix your vowels: Computer-assisted training by Dutch learners of Spanish

Lie-Lahuerta, C.

Publication date
2012

Document Version
Final published version

Published in
Tijdschrift voor Skandinavistiek

Citation for published version (APA):
Lie-Lahuerta, C. (2012). Fix your vowels: Computer-assisted training by Dutch learners of Spanish. *Tijdschrift voor Skandinavistiek*, *32*(1-2), 69-88. http://dpc.uba.uva.nl/tvs/vol32/nr01/art05

Carmen Lie-Lahuerta

Fix Your Vowels: Computer-assisted training by Dutch learners of Spanish

It is common knowledge that foreign learners struggle when it comes to producing the sounds of the target language accurately. Research in L2 speech acquisition has shown that in order to achieve native-like production of sounds in a foreign language, the learner must first become proficient in perceiving these sounds. That is, if Dutch listeners perceive Spanish vowel sounds through their first language categories, they are likely to produce them through their L1 categories, i.e. in a non-native way, thus with a foreign accent. This study analyses the effect of the vowel training programme *Fix your vowels* on the production of Spanish vowels by native speakers of Dutch. The findings confirm that computer training exerts a positive effect on the production of Spanish vowels and that this effect is also related to the desire to acquire a native accent.

1. Introduction

The most challenging task for the adult learner of a second language is acquiring new vowel sounds. It requires a great deal of

© *TijdSchrift voor Skandinavistiek* vol. 32 (2011), nr. 1-2 [ISSN: 0168-2148]

time and individual attention from teachers[1] and, in most learning contexts, the time that can be dedicated to practicing with individual students is generally small, or altogether non-existent. Furthermore, vowels that differ from the native language (L1) are difficult to teach, because their articulatory properties cannot always be clearly described, and vowel articulation is difficult to observe without special instrumentation.[2] Consequently, vowels may be excellent candidates for computer-assisted pronunciation training (CAPT).[3] However, specific theoretical problems during speech learning, such as vowel pronunciation, have not yet been solved in CAPT settings.[4]

This pilot study was initiated to investigate the effect of a training tool on pronunciation, particularly the vowel system. Therefore, we developed a training programme using the speech signal processing programme Praat[5] (www.praat.org.), which is different from a number of software packages that are available for speech analysis. Most freeware programs for the elaboration of spectrograms do not allow the in-depth study and treatment of formant values. Furthermore, software programmes are expensive. This new tool was developed using the Praat programme, which can be downloaded for free and allows the in-depth study and treatment of formant values.

In this paper, we report our research on the effects of the *Fix*

---

1 Wang, 'Acoustical analysis of English vowels produced by Chinese, Dutch and American speakers', 1999, pp. 30-42.

2 Lord, '(How) can we teach foreign language pronunciation? On the effect of a Spanish phonetics course', 2005, pp. 557-567.

3 Warschauer and Healey, 'Computers and language learning: An overview', 1998, pp. 57-71.

4 Brett, 'Computer generated feedback on vowel production by learners of English as a second language', 2004, pp. 103-113.

5 The Praat programme was elaborated by Paul Boersma, Professor of Phonetic Sciences at the University of Amsterdam.

Your Vowels programme on Spanish vowel production of Dutch students at the University of Amsterdam. The Northern Standard Dutch vowel system consists of nine monophthongal vowels / i ɪ y ʏ ɛ a ɑ ɔ u/ and vowel duration is a contrastive feature. In contrast, the Spanish vowel system consists of only five steady-state vowels /a e i o u/, but vowel length is not a contrastive feature. Dutch learners simply reuse five of their L1 vowels for representing their L2 Spanish lexemes. Reusing the existing categories leads to a mismatch when producing the Spanish sounds. To facilitate the learning of this new vowel sound system, we used this pilot study to develop the *Fix your vowels* programme, which endeavours to provide reliable, clear and useful feedback on vowel production to learners of Spanish as a second language.

In this paper, we demonstrate the results of our research on the suitability of the Praat programme for vowel production training. The purpose of this study is to show that CAPT, if properly adjusted for specific pedagogical goals, can be effective in improving pronunciation skills despite occasional errors. We describe the research context (Section 1), the procedure adapted to analyse vowel tokens (Section 1. 2), the architecture of the system we developed (Section 2), and the pilot study with its correspond-ding results (Section 3).

1.2. Research context

Languages differ greatly with respect to the number and types of vowels in their phonemic inventory; consequently, they provide a wealth of opportunities for researchers in second language acquisition (L2). In phonological terms, vowels are classified and distinguished in part by the relative position of the tongue in the mouth during articulation; that is, vowels may be classified in terms of tongue height (e.g., high, mid, low) and frontness/backness (e.g., front, central, back). These properties are reflected acoustically, to some degree, in the formant frequencies associated with each vowel. The formant frequencies refer to the characteristic "pitch overtones" of a given vowel as a function of the size and shape of the articulatory tract.[6] There are two primary formants that distinguish vowels: the first formant for vowel height (F1), and the second formant for vowel backness (F2). To illustrate this point, Figure 1 shows the average F1 and F2 values (in Hertz) of the five Spanish vowels /a e i o u/:

Spanish vowels	**First formant (Hertz)**	**Second formant (Hertz)**
/a/	F1 = 699	F2= 1471
/e/	F1 = 457	F2= 1926
/i/	F1= 313	F2 = 2200
/o/	F1= 495	F2 = 1070
/u/	F1 = 349	F2 = 877

Figure 1: The average F1 and F2 values of Spanish vowels of men[7]

In addition to the acoustic or spectral quality of vowels, quantity may also play a distinctive or phonetically prominent role in a given language. To this extent, certain languages, for example, Dutch, demonstrate phonological contrasts between long and short vowels that have otherwise similar spectral properties.[8] Other languages, such as English, have long and short vowels, but the long-short pairs also exhibit spectral differences (e.g., the English /i/ has a lower F1 value and higher F2 value than /ɪ/. Still other languages, such as Spanish, do not show any significant durational differences for vowels whatsoever.[9] Both vowel quality and quantity may be

6 Ladefoged, *Vowels and consonants: An introduction to the sounds of languages*, 2001, p. 34.

7 Martínez, 'En torno a las vocales del español', 1963, p. 197.

8 Wang, 2006, pp. 237-248.

9 Chládková, 'Context-specific acoustic differences between Peruvian and

measured fairly readily through acoustic analysis. Subsequently, in L2 speech research, a common methodological approach is to examine the L2 vowels produced by learners and to compare characteristics, such as average formant frequencies or duration of articulation, to those of monolingual speakers.

In our study, we investigated the Spanish vowel system of Dutch students. Both systems differ considerably as we shall see below, both in the number of vowels in the inventory and in the details of their positions within the articulatory vowel space. Potentially, they may also differ in terms of their durational characteristics (this part was excluded from this pilot study but will be investigated in further research). When native Dutch speakers speak Spanish as a foreign language, their pronunciation of vowels deviates from the native norms of Spanish, producing a "foreign accent". This "foreign accent" can be defined as deviations from the expected acoustic (e.g., formants) and prosodic (e.g., intonation, duration, and rate) norms of a language. Several studies have hypothesised that adult L2 learners perceive new L2 sounds through their L1 sound categorisation. That is, they discriminate and identify speech sounds on the basis of language-specific combinations of acoustic cues. Learning to perceive speech, therefore, consists not only of learning to identify the relevant acoustic cues in the speech signal but also of learning to combine and weight them appropriately.[10] To that end, many studies claim that the deficiencies arising during the process of perception account for many of the production problems that non-native speakers encounter.[11] Several studies hypothesise that L2 phonetic

Iberian Spanish vowels', 2011, pp. 416-428.

10 Boersma, 'Empirical tests of the gradual learning algorithm', 2001, pp. 45-86.

11 Bent (e.a.), 'Segmental errors in different word position and their effects on intelligibility of non-native speech', 2007, p. 331; Best and Tyler, 'Nonnative speakers and second language speech perception', 2007, p. 13; Flege, 'Production and perception of a novel, second-language phonetic contrast',

segments cannot be produced accurately unless they are perceived accurately.[12] Cross-linguistic speech perception research performed in the 1960s showed that L2 learners also have 'perceptual foreign accents,' i.e., their perception is shaped by the perceptual system of their L1.[13]

This finding seems to suggest that the origin of a foreign accent is the use of language-specific perceptual strategies that are rooted in the L2 learner and cannot be avoided when encountering L2 sound categories.[14] Problems producing L2 sounds could originate in particular from difficulties in perceiving such sounds accurately; that is, in a native-like way.[15] Research in L2 speech perception has shown that in order to achieve native-like production of sounds in a foreign language, the learner must first become proficient in perceiving these sounds.[16] Many researchers have elaborated perception and production models, such as Best's Perceptual Assimilation Model (PAM)[17] and Flege's Speech Learning Model (SLM).[18] However, one of the most specific models is the Second Language Linguistic Perception model (L2LP).[19] The L2LP, in contrast to the other models, describes the learning scenarios of a L2 learner. Escudero (2005) predicts in her L2LP model that there

1993, pp. 1589-1608.

12 Kuhl, 'A new view of language acquisition', 2000, pp. 11850-11857.

13 Strange, 'Cross-language study of speech perception: a historical review', 1995, pp. 3-45.

14 Escudero, *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*, 2005, p. 85.

15 Flege, *Phonetic interference in second language acquisition*, 1982, p. 51.

16 Rochet, 'Perception and production of second-language speech sounds by adults', 1995, pp. 75-79.

17 Best and Tyler, 2007, p. 13.

18 Flege, 'Second language speech learning theory, findings and problems', 1995, pp. 233-277.

19 Escudero, 2005, p. 85.

can be three possible relations between the L1 and L2 sounds and that these will result in three different learning scenarios:

- NEW: the second language has a contrast that the first language does not possess, but whose members are acoustically similar to one L1 phoneme, p.e. Spanish learners of English /i/ and /ɪ/, while in Spanish there is only the phoneme /i/ or Dutch learners of Englisch /æ/ and /ɛ/, while in Dutch there is only the phoneme /ɛ/. These learners only perceive one phoneme, that of their L1.

- SUBSET: (also Multiple Category Assimilation): the learner is being faced with a language whose phonemic categories constitute a subset of the L1 ones. In the initial state of the learning process, the learner will perceive more categories than the L2 listener, e.g. the Dutch learner will experience a learning problem, while Spanish has two front vowels, /i/ and /e/, Dutch has three corresponding categories, /i/, /ɪ/ and /ɛ/. Spanish learners of Dutch perceive the Spanish /e/ as both /ɪ/ and /ɛ/.

- SIMILAR: two L2 phonemes are equated with two L1 phonemes, which poses a learning problem, because there will often be a mismatch between the L1 and L2 perception of the two sounds in question, e.g. the learning of English /i/ and /ɪ/ by Spanish listeners, they perceive /i/ and /e/.

Furthermore, she predicts acquisition for the four learning stages: the initial state, the task-learning stage, the development stage, and the end stage. For all three scenarios, the L2 learner will attain optimal L2 perception and, at the same time, maintain optimal L1 perception. If an L2 learner fulfils the predictions of Escudero's L2LP model in production, this result would merit investigation in a future longitudinal study. For this study, we will concentrate only on production.

One example of Multiple Category Assimilation is the vowel system of Dutch native speakers, who possess a larger vowel system than native Spanish speakers. In this case, the learner perceives more sounds than those produced in the target language. The

Northern Standard Dutch vowel system consists of nine monophthongal vowels /i ɪ y ʏ ɛ a ɑ ɔ u/ with steady-state characteristics, and three long mid-vowels (/e o ø/) possessing more dynamic character (also called "potential diphthongs", see Escudero/Williams 2011) and three diphthongs /ɛi, oey, ʌu/,[20] several of which differ in length. On the other hand, the Spanish vowel system consists of only five steady-state vowels (/a e i o u/) and fourteen diphthongs; however, vowel length is not a contrastive feature. These five monophthongal vowels in Spanish are different in location in F1-F2 space than those in other languages with five monophthongs, such as Japanese, which has an (articulatory unrounded and therefore) acoustically fronter /u/, which is traditionally transcribed as /ɯ/.[21]

Upon producing the Spanish vowels, Dutch learners maintain at least twelve vowel categories from their native lexical representation.[22]

Learners simply reuse five of these L1 vowels to represent L2 Spanish lexemes. The reuse of existing categories leads to a mismatch of perceiving and producing the Spanish sounds. It is expected that, based on the acoustic comparison of the Spanish and Northern Standard Dutch vowels, learners will produce /i y ɛ ɑ u/ in terms of their acoustically closest Spanish counterparts, /i i e a u/. In Figure 2 (Boersma/Escudero 2008), we see the Spanish vowels circled among the twelve Dutch vowels. It is expected that the Dutch learners will have to categorise these new Spanish sounds by reducing or increasing, either the first formant (vowel height), the second formant (vowel backness), or both. To facilitate learning this new vowel sound system in this pilot study, we developed a programme that endeavours to provide reliable, clear, and

20 Adank (e.a.), 'An acoustic description of the vowels of Northern and Southern Standard Dutch', 2003, pp. 1729-1738.

21 Chládková (e.a.), 2011, pp. 416-428.

22 Boersma, 'Learning to perceive a smaller L2 vowel inventory: An Optimality Theory account', 2008, pp. 271-301.

useful feedback on vowel production for learners of Spanish as a second language.

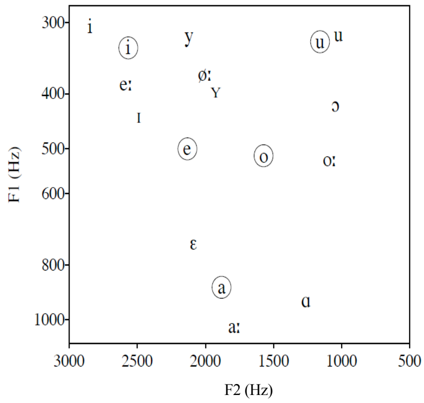

Figure 2: The Spanish vowels (circled) amongst the twelve Dutch vowels

1.3. Analysis of vowel tokens

This pilot study has been carried out using Praat, which is a programme that can be downloaded free from www.praat.org. This programme has many possibilities for speech analysis, and can be easily modified for specific research purposes; results can be exported to Excel-compatible spreadsheets. Use was made of Praat's inbuilt Linear Predictive Coding (LPC) formant analysis function. The scripts were written by Dirk Jan Vet and Ton Wempe, automating a number of steps in order to make the use of the programme as straightforward as possible. In reality, Praat was not designed to be used as a training programme; however, with certain modifications, it can function as an instrument of speech learning. The present study examines the programme Fix Your Vowels (FYV) software programme for providing feedback on learners' vowel production based on the analysis of formant data. For those unfamiliar with the topic, a brief explanation may be necessary. An analysis of the acoustic qualities of vowels shows peaks at certain frequencies. The frequency at which these peaks appear differs from one vowel sound to another. Furthermore, the values, when plotted inversely (i.e., as negative values) on a graph with F2 and F1 on the x- and y-axes, respectively, bear a resemblance to the traditional vowel chart (Fig. 3), which, in turn, is directly connected to articulation. In other words, visual information, extracted from a produced vowel sound, bears a direct (albeit inverse) relationship to the articulatory position the speaker adopted when producing the sound and vice versa. In Figure 3, we see the Spanish vowels (circled) plotted amongst the Dutch counterparts.

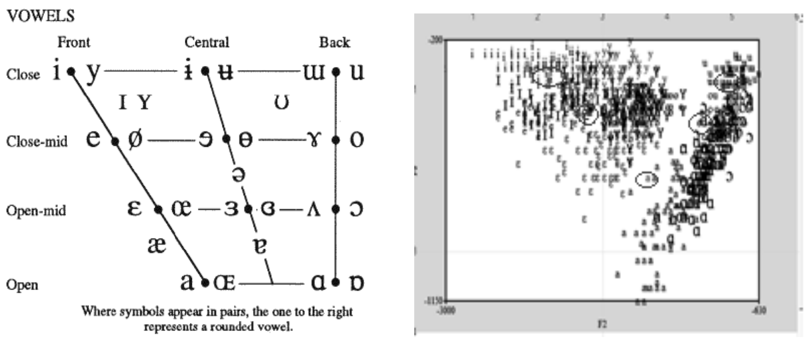

Figure 3: Comparison between the vowel chart (left) and a graph (right) plotting formants values of the median speakers of Spanish (circle) and Dutch learners (black), x-axis= F2 (Hz), y-as = F1 (Hz).

In FYV, the main purpose of the vowel similarity system is to determine if a given student's vowel falls within a vowel space derived from a target set of vowels, the latter being produced by a group of native speakers. In order to derive the target vowel spaces, vowel data were collected from three female and three male native speakers of Spanish with each speaker producing words containing the target vowels. The formant data used to create a target set of vowels spaces were derived as described in the next paragraph. The data were checked for obvious formant tracking errors within Praat, and the corresponding samples were deleted from the data. Extreme outliers were also identified and excluded from the data. The acoustic vowel analysis provides a representation of vowel tokens in terms of the normalised formant parameters that are put into the vowel system. A given speech token is submitted to the segment scripts to isolate the vowel from any surrounding speech or silence. The isolated vowel is subsequently analysed to produce estimates of the first three formant frequencies.[23] The most stable region of the vowel is located by a "steady state" finder algorithm in Praat and projected in the vowel triangle.

2. The architecture of the programme Fix Your Vowels

For the training, we used the programme from the pilot study Fix Your Vowels*,* which was made with Praat[24] in collaboration with Dirk Jan Vet and Ton Wempe. This programme has been developed to practice monophthongal vowels in Spanish. The main purpose of the vowel similarity metric is to determine if a given student's vowel token falls within a vowel space derived from a target set of vowels produced by native speakers. In order to derive the target vowel spaces, vowel data were collected from six speakers (three male and three female) with each speaker producing 150 different vowel tokens in different word positions. The formant data used to create the target vowel spaces were derived with scripts written for Praat. The data were checked for obvious formant tracking errors, and the corresponding samples were deleted

---

23 Escudero, 'Perceptual assimilation of Dutch vowels by Peruvian Spanish listeners', 2011, pp. 254-260.

24 Praat version 18 is a programme that can be downloaded free of charge from www.praat.org. This programme has many possibilities for the analysis of speech.

from the database. Extreme outliers were also identified and excluded from the data. The final target vowel spaces were subsequently derived using a script in Praat, which generated a two-dimensional vowel triangle; targets were calculated using spreads of 1, 1.5, and 2 standard deviations either side of the mean values. The final decision of the metric determines if the input formants from the student's vowel token fall within the equivalent target vowel space (this component of our research has to be improved in a future study). Unfortunately, formant values measured for the same vowel differ when different individuals with distinct vocal tract shapes and cavity sizes produce the tokens. Thus, in the present study, we have opted for a straightforward vowel normalisation, also called calibration procedure, first used by Lobanov (1971), which is simply a z-normalisation of the F1 and F2 frequencies over the vowel set produced by each individual speaker. In the z-normalisation, the F1 and F2 are transformed to z-scores by subtracting the individual speaker's mean F1 and F2 values from the raw formant values and dividing the difference by the speaker's standard deviation. We applied Hertz values to the calibration procedure. Figure 4 shows a typical example of the vowel-teaching module's user interface for a female Dutch student. The vowel triangle is placed above a small prompt window and four user buttons. The main display shows 1) the vowel triangle (to provide a reference for the articulatory position of vowel targets), 2) the vowel, exposed in the learner's triangle in yellow, 3) the vowel, exposed in the native speaker's triangle in red, and 4) a real-time feedback indication, produced two seconds after pronouncing the word and aimed at improving the position of the vowel. The students can improve their pronunciation by reducing or increasing the first formant (vowel height), the second formant (vowel backness), or both. With this programme, they can aim at the correct vowel position as if it were on a dartboard. When the right position is hit, i.e., in the correct, native speaker's vowel space, a green light will turn on (this has to be improved in the future).

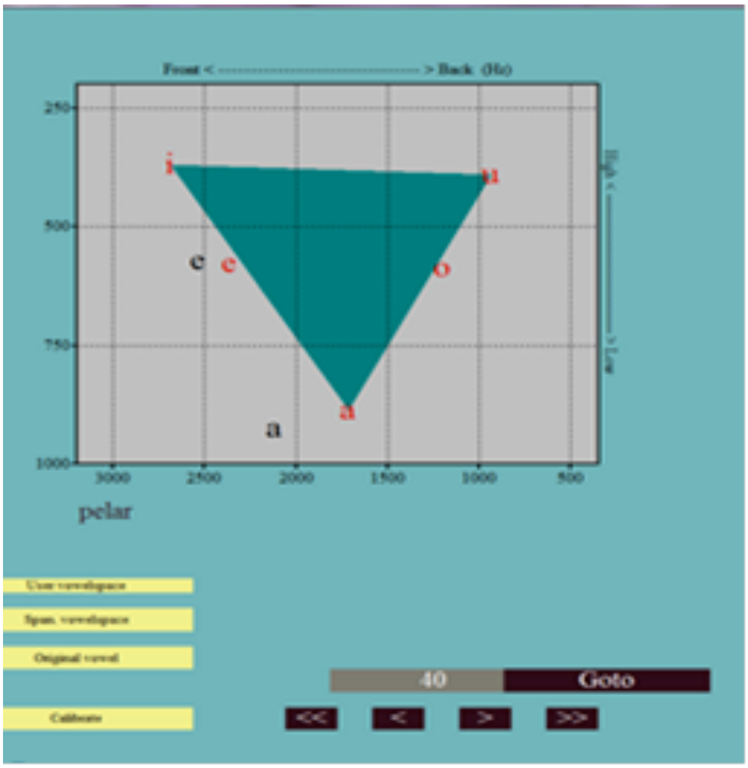

Figure 4: The Fix Your Vowels programme. The triangle shows the Spanish vowels in red, and the L2 learner's vowels in black.

3. The pilot study

The pilot study involved 19 participants (5 men and 14 women, with the mean age of 23), all of whom were first-year students at the University of Amsterdam who had previously taken two semesters of language acquisition classes. The subjects voluntarily practiced their vowel production with the computer programme that was created in this pilot study. The training lasted for four weeks with half-hour sessions in Fix Your Vowels, itself within Praat. The training consisted of recordings made of the five Spanish monophthongal vowels, /a e i o u/. All target vowels were produced in separate words. The words had the following generic structure (C= consonant, V= vowel): CVC, CVCV, CVCVCVC. The initial consonants were specific voiceless consonants, /p t c k q f θ s h/, which were chosen for better formant detection. Students' last names, language backgrounds, and gender were specified because the programme makes use of different parameter settings for the acoustic analyses of male and female speakers. Navigation through the exercises is undertaken freely, and users can complete an exercise at their own pace before proceeding to the following one.

Before starting the training, every student had to normalise his or her vowels; we used a calibration method to correct the deviations. The deviations are recorded in a so-called correction table. Through the digital processing of measured values, the correction values are calculated such that an accurate result is obtained. Based on the calibration, one can determine whether the measuring device (in this case, the vowel triangle) remains true to specifications.

3.1 Results

We ran a linear mixed model on the F1 and F2 values of the non-native speakers' first and last attempts with vowel category as the within-subject factor and gender and word as the between-subject factors. The Dutch participants were measured acoustically in a pre-test, and their values were compared with those of six native speakers of Spanish (three men from Madrid and three women, two from Barcelona and one from Valencia), all of whom were lecturers of Spanish at the University of Amsterdam.

The non-natives differed in spectral values with the natives for the vowel token /e/ (F1 was higher, t-value=3.22, df=18) as well as the vowel token /a/ (F1 was lower, t-value=2.30, df=12). As for the results of the training test, we ran the same linear mixed model on the F1 and F2 values with vowel category as the within-subject factor as well as gender and language as the between-subject factors. The analysis principally shows the effect of vowel category on both measures. The analysis demonstrated a significant improvement in scores obtained for F1 with the vowel /a/ (t- value= 5.38, p=0.000043, df= 18) and a lesser improvement for /u/ (t-value= 2.12, p= 0.058, df=3); meanwhile, there was a significant improvement in F2 for the vowels /e/ (t-value= 3.5, p= 0.0040, df= 12) and /u/ (t-value= 3.39, p= 0.015750, df=3).

Anonymous questionnaires were used in which participants indicated whether they agreed with a number of statements on a 1-5 Likert-scale; additionally, they had to answer two open-ended questions. The answers indicate that the students enjoyed working with the provided programme and, furthermore, that participants found the training to be useful. Eleven of the fourteen participants who provided comments on the system said that it was helpful, mostly in improving their pronunciation and in making them aware of specific pronunciation problems. We can conclude that after only four weeks there was a significant improvement of some, but not all, vowel tokens.

4. General discussion

Several perception training studies had shown that learners could be successfully trained to redirect their attention to acoustic cues, normally unnoticed because they do not mark phonetic contrasts in their native language. For example, Japanese and Korean learners of English improved their perception of the consonant /l/-/r/ contrast.[25] Furthermore, it has been shown that beginner-level Spanish learners of English can achieve the English /i/and /ɪ/

[25] Hazan (e.a.), 'Effect of audiovisual perceptual training', 2005, pp. 54-59.

through practice.[26] However, few studies have investigated production resulting from computer training undertaken by adult learners.[27] In Cucchiarini's study (2009), an automatic speech recognition programme was used, which provided limited feedback (e.g., "you had a problem with the red sound") on a circumscribed number of well-selected, problematic phonemes. Brett's study (2004) of computer-generated feedback on vowel production by learners of English as a second language also used Praat with another application but concluded that the feedback wasn't user-friendly. Learners couldn't start practising immediately, as a series of readings must be taken first; typically, ten readings for vowels at the extreme ends of the chart were required. These values could then be exported and loaded each time the learner started the exercise. Feedback was given in the form of a phonetic transcription of the sound they pronounced.

Computer training programmes have a long way to go before an individual learner can easily use it, i.e., without qualified help, to gain useful, clear feedback on vowel production. However, in this pilot study, we developed a training tool that gives appealing feedback that is not only easily understandable for any learner of Spanish but can also be applied to any vowel system. The scripts are open-source and can be modified, developed, and tailored to the specific needs of the training situation. Furthermore, the possibility of training and analysing speech simultaneously without speech recognition is a step forward in computer training development.

5. Conclusion

In this paper, we have presented a system for providing automatic,

---

26 Flege, 'Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English', 2003, pp. 90-99.

27 Neri, *The pedagogical effectiveness of ASR-based computer assisted pronunciation training*, 2002, pp. 143-147.

corrective feedback on pronunciation errors in Spanish, focusing especially on vowel detection, scoring accuracy, and feedback effectiveness. We have shown that while this system could be improved in terms of error detection, it was nonetheless effective in improving the pronunciation of vowels after just a few hours of use over a one-month period; furthermore, learners enjoyed using it. Nevertheless, the results from this pilot study are an indication of differences between non-native and native speakers' vowel production, and a future longitudinal study would be needed to demonstrate whether non-native speakers improve with training and retain such improvement over time.

References

Adank, Patricia, Van Hout, Roeland, Smits, Roel (2004): 'An acoustic description of the vowels of Northern and Southern Standard Dutch', in: *Journal of Acoustical Society of America*, 116 (3), 1729-1738.

Bent, Tessa; Bradlow, Anne; Smith, Bruce L (2007): 'Segmental errors in different word position and their effects on intelligibility of non-native speech' in: Bohn, Ocke; Munro, Murray (eds.), *Language experience in Second Language Speech Learning: In honor of James Emil Flege.* Portland, OR: John Benjamins, 331-340.

Best, Catherine T. (1995): 'A direct realist view of cross-language speech perception' in: Strange, Winifred (ed.), *Speech Perception and Linguistic Experience: issues in cross-language research*. Timonium, MD: York Press, 171–204.

Best, Catherine; Tyler, Michael (2007): 'Nonnative and second language speech perception' in: Bohn, Ocke; Munro, Murray (eds.), *Language experience in second language speech learning: In honor of James Emil Flege*. Portland, OR: John Benjamins, 13-19.

Boersma, Paul; Hayes-Harb, Bruce (2001): 'Empirical tests of the gradual learning algorithm' in: *Linguistic Inquiry*, 32, 45-86.

Boersma, Paul; Escudero, Paola (2008): 'Learning to perceive a smaller L2 vowel inventory: An Optimality Theory account' in: Avery (et al, eds.) *Contrast in phonology: Theory, Perception, Acquisition*. Berlin: Mouton de Gruyter, 271-301.

Boersma, Paul; Weenink, David (1992-2011): *Praat: Doing phonetics by computer*. Amsterdam: Institute of Phonetic Sciences, University of Amsterdam.

Brett, David (2004): 'Computer generated feedback on vowel production by learners of English as a second language' in: *ReCALL*, Cambridge: Cambridge University Press, 16 (1), 103-113.

Chládková, Katerina; Escudero, Paola; Boersma, Paul (2011): 'Context-specific acoustic differences between Peruvian and Iberian Spanish vowels' in *Journal of the Acoustical Society of America*, 130, 416-428.

Cucchiarini, Catia; Neri, Ambra; Strik, Helmer (2009): 'Oral Proficiency Training in Dutch L2: the Contribution of ASR-based Corrective Feedback' in: *Speech Communication*, 03 003.

Escudero, Paola (2005): *Linguistic perception and second language acquisition:*

Explaining the attainment of optimal phonological categorization. [PhD thesis]. Utrecht: LOT Dissertation series, 113.

Escudero Paola; Boersma, Paul (2002): 'The subset problem in L2 perceptual development: Multiple- category assimilation by Dutch learners of Spanish' in: Skarabela (et al, eds.), *Proceedings of the 26th annual Boston University Conference on Language Development*, Somerville, MA: Cascadilla Press, 208– 219.

Escudero, Paola; Boersma, Paul (2004): 'Bridging the gap between L2 speech perception research and phonological theory' in: *Studies in Second Language Acquisition*, 26, 551-585.

Escudero, Paola; Williams, Daniel (2011): 'Perceptual assimilation of Dutch vowels by Peruvian Spanish listeners' in: *Journal of the Acoustical Society of America*, 129, EL1–EL7.

Flege, James (1993): 'Production and perception of a novel, second-language phonetic contrast' in: *Journal of the Acoustical Society of America*, 93, 1589-1608.

Flege, James (1995): 'Second language speech learning theory, findings and problems' in: Strange, Winifred (ed): *Speech Perception and Linguistic Experience*, 233-277.

Flege, James; MacKay, Ian (2003): 'Perceiving vowels in a second language' in: *Studies in Second Language Acquisition,* 26, 1-34.

Flege, James (1982): *Phonetic interference in second language acquisition*. Ann Arbor, Mich.: University Microfilms International.

Hazan, Valerie; Sennema, Anke; Iba, Midori; Faulkner, Andrew (2005): 'Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English' in: *Speech Communication*.

Kuhl, Patricia (2000): 'A new view of language acquisition' in: *PNAS*, 97 (22), 11850-11857.

Ladefoged, Peter (2001): *Vowels and consonants: An introduction to the sounds of languages*. Malden: Blackwell Publishers.

Lobanov, Boris (1971): 'Classification of Russian Vowels Spoken by Different Speakers' in: *Journal of the Acoustical Society of America*, 49, (2B), pp. 606-608.

Lord, Gillian (2005): '(How) can we teach foreign language pronunciation? On the effect of a Spanish phonetics course' in: *Hispania*, 88 (3), 557-567.

Martinez Celdrán, Eugenio (1963): 'En torno a las vocales del español' in: *Nueva Revista De Filología Hispánica*, 17 (1/2), 197.

Morrison, Geoffrey (2006): 'L1 & L2 production and perception of English and Spanish vowels: A statistical modeling approach' in: *Journal of Linguistics*. Edmonton, AB.

Navarro-Tomás, Tomás (1967): *Manual de pronunciación española.* Madrid: Consejo Superior de Investigaciones Científicas

Neri, Ambra (2007): *The pedagogical effectiveness of ASR-based computer assisted pronunciation training.* [PhD thesis]. Utrecht: LOT.

Rochet, Bernard (1995): 'Perception and production of second-language speech sounds by adults' in: Strange, Winifred (ed.), *Speech Perception and Linguistic Experience*, 379-410.

Strange, Winifred (1995): 'Cross-language study of speech perception: a historical review' in: Strange, Winifred (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* 3-45. Timonium, MD: New York Press.

Trubeckoj, Nikolai (1958): *Grundzüge der Phonologie.* Göttingen: Vandenhoeck/Ruprecht.

Wang, Hongyan; Van Heuven, Vincent (2006): 'Acoustical analysis of English vowels produced by Chinese, Dutch and American speakers' in: *Linguistics in the Netherlands*, 237-248.

Warschauer, Mark; Healey, Deborah (1998): 'Computers and language learning: An overview' in: *Language Teaching*, 31, 57-71.