Training of Spanish monophthongs

Carmen Lie
University of Amsterdam
lie@uva.nl

Problem

When speaking an another language, you will always almost have a foreign accent. We are aware of our accent, but find it hard to get rid of. That's strange, because we have no problem making the sounds of our own language, and we can mimic all kinds of other sounds. However, it is scientifically proven that after puberty we have difficulty perceiving the subtle differences between the vowels of an unfamiliar language. We tend to interpret all vowels we hear exclusively in relation to the sounds of our own language. Even if we perceive a foreign vowel, we cannot reproduce it easily.

Solution

The *Fix Your Vowels* programme makes it possible not only to hear differences, but also to see them. This computer programme helps second language learners to improve their pronunciation. It shows how and where vowels are formed in the mouth. It is like playing darts with your voice. You imitate the word you hear from the native speaker, and you see on the screen how far off you are from the original.

1. Introduction

It is generally known that the production and perception of second language (L2) sounds is usually a difficult tasks for adult learners of a new language, and that the acoustic similarities and differences between the first language (L1) and L2 inventories are an important factor affecting L2 learning (e.g. Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Best, McRoberts, & Sithole, 1988; MacKay, Meador, & Flege, 2001; MacKay, Flege, Piske, & Schirru, 2001; Flege, Birdsong, Bialystok, Mack, Sung, & Tsukada, 2006; Hojen, & Flege, 2006). For example, the studies of Mackay et al (2001) and Brannen (2002) have shown that the production and perception of sounds in a L2 are strongly influenced by a speaker's first language. In speech production this influence typically manifests itself as an accent in the speaker's L2. Many theories have claimed that learner's accents actually have a perceptual basis and that native like production can only be achieved if perception is accurate (Flege, 1995; Rochet, 1995; Escudero, 2005) and that learners' initial perception of L2 phones (or contrasts) is formed entirely by their native-language perceptual system (Escudero, 2005; Best, 1995; Best & Tyler, 2007). This was, for instance, shown by Strange (1995) in a cross-linguistic speech perception research in the 1990s: L2 learners exposed 'perceptual foreign accents', i.e. their perception was shaped by the perceptual system of their L1. These difficulties learners have in perceiving L2 vowels can be predicted and explained by the comparison between the L1 and L2 vowel systems.

Therefore, our lesson plan starts with the perception of the Spanish vowels in the *absolute* beginning stage, as the learners create a copy of their L1 system and will reuse their L1 perceptual mappings at the onset of their L2 acquisition process. Once we know this initial stage we can predict the problems they will have in production. Therefore, we start with two perception tasks, an assimilation and a discrimination task, the students can analyse the results and focus on his/her problems in production.

On production we use our program Fix Your, a standalone[1] program, the basic format of which was designed by engineer D.J. Vet, who is an electronic and audio-visual engineer at the department of Phonetic Science of the University of Amsterdam. It 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[2]. The data were checked for obvious formant tracking errors, and the corresponding samples were deleted 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

[1] ED Standalone program is a computer program that does not load any external module, library function, or program and that is designed to boot with the bootstrap procedure of the target processor. Selection for this program was made because of the possibility to design our own experiment. Furthermore, obtained data is saved automatically.

[2] Boersma, Paul & Weenink, David (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.60, retrieved 8 December 2013 from http://www.praat.org/

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 1 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 black, 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 program, they can aim at the correct vowel position
as if it were on a dartboard.

2. Lesson plan

2.1. Perception

Objective: to analyse the perception problem and to develop a personal plan for improvement of the perception of the Spanish vowels.

Much of L2 perception research is based on Vowel Assimilation (Escudero & Chládková, 2010), Discrimination tasks (Cebrian, 2006) or a combination of both (Cebrian, 2008; Flege & Mackay, 2004; Rallo Fabra & Romero, 2012). Cebrian et al. (2010) points out that using a variety of tasks to measure cross-linguistic distance will give us a, "more complete and reliable measure of cross linguistic similarity". Cross-language assimilation tasks form the basis for many L2 perception studies (Mayr, 2005; Cebrian, 2008; Rallo Fabra & Romero, 2012). In this type of task participants classify L2 sounds in terms of their L1. The PAM-L2 and the L2LP model use the similarity relations to predict the relative difficulty that will occur in the discrimination of two L2 sounds. A way to see if language learners have established new L2 phonetic categories is often used by researchers in discrimination tasks (Flege & MacKay, 2004; Cebrian, 2008). During this task two L2 sounds are presented to the participant and they must determine whether the two sounds are instances of the same sound or if they represent two different sounds. Since the establishment of new L2 categories increases language learners' sensitivity to between-category differences while the decreases sensitivity to within-category differences that can occur between speakers (i.e. pitch, tone etc.), it is expected that language learners will have greater difficulty in discriminating the tokens of an L2 contrast if they do not have the necessary L2 categories formed (Flege, 1995).

Task 1 – Cross-language assimilation (10 minutes)

The goal of this vowel assimilation task is, as we mentioned before, to assess how native English speakers perceive the relation between L2 Spanish vowels and L1 English vowels. The three most influential theories in this respect are the following ones: Flege's *Speech Learning Model* (SLM; Flege, 1981, 1987,1995), Best's *Perceptual Assimilation Model* (Best 1995;PAM-L2; Best & Tyler, 2007) and Escudero's *Second Language Linguistic Perception Model* (L2LP model) (Escudero & Boersma 2004, elaborated in Escudero, 2005) suggest that predictions can be made about the level of discrimination difficulty of two L2 sounds by analyzing how each is categorized onto the L1 sound system. Therefore, the patterns of categorization found here will provide the foundation for making and assessing the predictions of the four Spanish vowel contrasts /a-e/, /e-i/, /a-o/ and /o-u/ that will be under investigation in the AXB categorical discrimination task (Task 2).

Task 2: Categorical discrimination (7 minutes)

Since vowels are often more easily distinguished in relation to one another than by any external standard of measurement, the perception of contrasts is an essential starting point in the teaching of vowels. Tuning students' ears to the subtle differences between vowels is critical. In this AXB discrimination tasks, subjects have to compare two stimuli and decide whether one of the vowels in A or B is identical to the vowel in X.

2.2 Production

Objective: training Spanish vowels to improve vowel production and perception.

2.2.1. Controlled practice and feedback

The training last for three weeks with half-hour sessions in Fix Your Vowels[3] (FYV), two exercises per week. The training consists of recordings made of the five Spanish monophthongal vowels, /a e i o u/. All target vowels are produced in separate words. The words have the following generic structure (C=consonant, V= vowel): CVC, CVCV, CVCVCVC. The initial consonants are specific voiceless consonants, /p, t, k, f, s /, which are chosen for better formant detection. Students' last names, language backgrounds, and gender are specified because the program 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 have to normalize 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. Real-time feedback is given by the vowel triangle (F1: height and F2: front/back), the pitch and the duration of the vowel.

The training contains the following listening and repeating exercises:

1. 20 nonsense CVC words like [fef], [sos] etc.
2. 20 nonsense CVCV words like [fefe], [seso]
3. 25 Spanish CVC words
4. 25 Spanish CVCV words
5. 25 Spanish CVCVC words
6. 25 Spanish CVCVCV words

All the students' productions are recorded and can be analyzed with Praat scripts. The results, i.e. the formant values can be compared with the Spanish natives' production. Projecting the formant values of the student in a Spanish vowel triangle[4], they can perceive their production

[3] Warning: this computer program FYV does not work on a Mac computer.

[4] For both proceedings we have made Praat scripts that are easy to use.

and be aware of their “mistakes” that might occur. Practicing with the program FYV will improve their ability to produce the Spanish vowels.

N.B. This programme can be converted to any language.

References

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (ed.), *Speech Perception and Linguistic Experience: issues in cross-language research*, 171–204. Timonium, MD: York Press.

Best, C.T. & Tyler, M.D. (2007). Nonnative and second language speech perception. In O. S. Bohn, & M. L. Munro (eds.), *Language experience in second language speech learning: In honor of James Emil Flege.* (pp. 13). Portland, OR: John Benjamins Publishing Co.

Boersma, P. & Escudero, P. (2004). *Learning to perceive a smaller L2 vowel inventory: An optimality theory account*. Ms. University of Amsterdam. Rutgers Optimality Archive

Boersma, P. & Escudero, P. (2008). *Learning to perceive a smaller L2 vowel inventory: An Optimality Theory account.* In Avery, P., Dresher, E. & Rice, K. (eds.) *Contrast in phonology: Theory, Perception, Acquisition*. Berlin & New York: Mouton de Gruyter, 271-301.

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

Carbó, C., Llisteri, J., & Machuca, M. J. (2003). Estándard oral y enseñanza de la pronunciación del español como primera lengua y como lengua extranjera. *Estudios De Lingüística De La Universidad De Alicante,* 17, 161-180.

Celdrán, E. M. (1963). En torno a las vocales del español. *Nueva Revista De Filología Hispánica,* 17 (1/2), 197.

Cervera, T., Miralles, J. L., & González-Alvarez, J. (2001). Acoustical analysis of Spanish vowels produces by laryngectomized subjects. *Journal of Speech, Language and Hearing Research,* 44, 988-996.

Díaz-Campos, M. (2004). Context of learning of Spanish second language phonology. *Studies in Second Language Acquisition, 26*(2), 249-273.

Escudero, P. (2005). Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization. Utrecht: LOT Dissertation series, 113.

Escudero P, & Boersma, P (2002). The subset problem in L2 perceptual development: Multiple- category assimilation by Dutch learners of Spanish. In B. Skarabela, S. Fish, & A. H.-J. Do, Eds. *Proceedings of the 26th annual Boston University Conference on Language Development,* pp. 208– 219. Somerville, MA: Cascadilla Press.

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

Escudero, P., Benders, T. and Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics,* 37, 452-465.

Escudero, P.,Boersma, P., Rauber, A. S., & Bion, R. A. H. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *Journal of Acoustical Society of America,* 126(3), 1379-1393.

Escudero, P. &. Chládkóva, K. (2010) "Spanish Listeners' Perception of American and Southern British English Vowels." *Journal of Acoustical Society of America* 128: EL254-EL260

Flege, J. E. (1987). The production of new and similar phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47–65.

Flege, J. E., O. S. Bohn, and S. Jang (1997). "Effects of Experience on Non-Native Speakers' Production and Perception of English Vowels." *Journal of phonetics* 25.4: 437-70.

Flege, J. E., E. M. Frieda, and T. Nozawa (1997). "Amount of Native-Language (L1) use Affects the Pronunciation of an L2." *Journal of Phonetics* 25.2: 169-86. Print.

Flege, J. E., I. R. A. MacKay, and D. Meador (1999). "Native Italian Speakers' Perception and Production of English Vowels." *The Journal of the Acoustical Society of America* 106: 2973.

Flege, J. E. (1991)."Perception and Production: The Relevance of Phonetic Input to L2 Phonological Learning." *Crosscurrents in Second Language Acquisition and Linguistic Theories.* Eds. T. Hueber and C. Ferguson. Amsterdam: John Benjamins,. 249-289.

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

Flege, J. E. (1995). Second language speech learning theory, findings and problems. In W. Strange, (ed): *Speech Perception and Linguistic Experience,* 233-277.

Flege, J. E. &. Munro,I. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition,* 26, 1-34.

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

García-Bayonas, M. (2008). "Perception of English Vowels as First and Second Language." *RAEL: revista electrónica de lingüística aplicada*.7

Guirao, M., and Borzone de Manrique A.M. (1975). "Identification of Argentine Spanish Vowels." *Journal of Psycholinguistic Research* 4: 1.

Hualde, J. I., and M. Prieto. (2002). "On the diphthong/hiatus Contrast in Spanish: Some Experimental Results." *Linguistics* 40.2; ISSU 378: 217-34.

Lobanov, B. M. (1971). Classification of Russian Vowels Spoken by Different Speakers. *J. Acoust. Soc. Am*. 49, (2B), pp. 606-608

Lozano, A. R. (2005). Sobre el vocalismo y la pronunciación. *Phonica, 1*

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

Rochet, B. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (ed.) *Speech Perception and Linguistic Experience*,379-410.

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

Zampini, M. L. (1998). The relationship between the production and perception of L2 Spanish stops. *Texas Papers in Foreign Language Education, 3*(3), 85-100.

Zampini, M. L. (2010, Feb.). *An eye on the present and a look to the future in Spanish and Portuguese L2 phonology research*. Plenary speech at Current Approaches to Spanish and Portuguese Second Language Phonology Conference, Gainesville, FL.

Zampini, M. L., Clarke, C. M., & Green, K. P. (2000). Language experience and the perception of stop consonant voicing in Spanish: The case of late English-Spanish bilinguals. In R. P. Leow & C. Sanz (Eds.), *Spanish Applied Linguistics at the Turn of the Millennium* (pp. 194-209). Somerville, MA: Cascadilla Press.