MIAPARLE: Online training to improve discrimination of stress contrasts

Jean-Philippe Goldman, Sandra Schwab
University of Geneva
jeanphilippegoldman@gmail.com, sandra.schwab@unige.ch

Abstract

MIAPARLE is a public web application that is designed to offer a range of CAPT (computer-aided pronunciation teaching) tools for L2 learners. Besides helping language learners to reduce their foreign-accentedness, the goal of the platform is to test these tools, gather feedback and improve them according to their educational impact. In this article, we describe one particular training tool that focuses on stress perception. This tool is particularly useful for speakers whose L1 is a fixed-stress language, such as French. These speakers have difficulties perceiving and discriminating stress contrasts. To help them with this so-called stress 'deafness', the methodology used in the training is based on successive questions in which a visual pattern is associated with the sound of a lexical item. After successively completing their pre-tests, training and post-tests, the participants are given their improvement score. The performance of the training is evaluated by comparing the learner's results at the pre- and post-test stages. Various methodological parameters, such as the number of training items or the number of visual patterns are tested in parallel in order to quantify their teaching efficiency, and to optimise the overall teaching impact.

Keywords: CAPT, pronunciation training, prosody, stress deafness

1. Introduction

Second language learners tend to imprint the prosody of their mother language onto the second language (L2) (e.g., Barquero et al., 2014; Muñoz, 2010). This prosodic cross-language transfer is often combined with segmental transfers, which can lead to the presence of a foreign accent. A foreign accent can not only hamper communication between learners and natives (like in Anderson, Munro & Derwing 1988), but it can also affect the credibility of learners and how they are evaluated by others; this can sometimes lead to social discrimination (e.g., Dailey et al., 2005, Lev-Ari & Keysar, 2010; , Purnell et al. 1999). Moreover, despite the crucial role of prosody in speech flow segmentation (e.g., Cutler 1984; Cutler & Butterfield 1992 for English; Bagou 2002 for French), it is rarely taught in language courses, even in foreign-language pronunciation courses.

In the framework of *computer-assisted pronunciation teaching*, the goal of this project is to develop a methodology which improves the prosodic skills of L2 learners in order to reduce cues of foreign accent. After many years of research (Neri et al. 2002, Busà 2008, Chun 2012), there has recently been a growing interest in CAPT applications (e.g. for French[1], Spanish[2], English[3],and Norwegian[4]). However, there have been few applications that focus specifically on prosodic features (among others,, Vakil & Trouvain, 2015 for German stress; Sztahó et al, 2014, for Hungarian intonation, stress and speech rhythm; Niebuhr et al. 2017, for visualization methods of prosody).

This paper focuses on the tool used to assist in discrimination of stress contrasts. Although this method is designed for learners with any first language (L1), training on stress contrasts is mostly useful for speakers with a fixed lexical stress native language (such as French, Czech, or Hungarian). For example, French-speaking listeners tend to experience difficulties in perceiving stress contrasts in a second language such as Spanish. These difficulties are the basis of the *stress deafness* hypothesis (e.g., Dupoux et al., 1997.). According to this hypothesis, the degree of *stress deafness* is related to the stress properties of the L1; more specifically, to the nature of lexical stress (free or fixed). In a free-stress language such as Spanish or English, lexical stress has a distinctive function, since it distinguishes segmentally identical words, such as in Spanish *<u>nú</u>mero* (['numeɾo], English *(the) number*) and *nu<u>me</u>ro* ([nu'meɾo], English *I number*)[5] . As a consequence, speakers of a free-stress language need to encode stress position in their mental representation of the words. On the other hand, the position of stress in a fixed-stress language such as French is not variable, and thus not contrastive. Consequently, the stress information does not need to be stored in the lexical representation. The *stress deafness* hypothesis claims that speakers of fixed-stress languages have difficulties in perceiving stress contrasts in free-stress L2s since they are not able, or at least not trained, to encode stress information in their mental lexicon (e.g., Dupoux et al., 1997). Nevertheless, as showed by our recent research (Schwab & Dellwo, 2016b), French learners are able to

[1] http://www.coelang.tufs.ac.jp/mt/fr-swiss/dmod/index_en.html
[2] http://www.enterate.unam.mx/Articulos/2007/abril/sarahi.html
[3] http://www.carnegiespeech.com/products/nativeaccent.php
[4] https://orgesuniversitetet.no/prosjekt/computer-assisted-listening-and-speaking

[5] The underlined syllable in these examples, and in the rest of the paper, corresponds to the stressed syllable.

improve their stress detection ability after a 4-hour training session. Although the current methodology focuses on isolated words only, our goal is to extend it to larger groups of words and continuous speech flow. Based on its promising results, our aim is that our tool for training stress-contrast discrimination will give L2 learners the bootstrapping required to start encoding stress information in their mental lexicon.

2. MIAPARLE

MIAPARLE is a web application that hosts various tools and activities dedicated to pronunciation training. It is based on PyBossa, a micro-tasking crowdsourcing platform. PyBossa is developed and maintained by SciFabric, runs on flask micro-framework and is well designed to be responsive on smaller screens such as those in tablets and smartphones.

In this paper, we describe the training tool for learning stress contrasts in Spanish. The application presents teaching material such as audio samples of lexical items pronounced by native speakers and corresponding clip arts. The whole procedure is organized as follows: the participants' initial performance is evaluated with 2 different pre-tests, the participants then follow one of two possible full training pathways (A or B), and eventually take post-tests in order to quantify their improvement level in the perception of lexical stress.

2.1 Pre- and post-tests

The pre- and post-tests are used to evaluate the training and comprise similar tasks. Two types of exercises are available: Localisation and Odd-One-Out, as described below. Each exercise has 15 items in the pre-tests (so the participants go through 30 items in total during two pre-tests). During the post-test, each activity shows the same 15 items and 15 additional items, in order to test the generalisation of the training. Thus the learner is shown 60 items during the post-tests.

	Pre-test	Training	Post-test
	localisation (15)	A or B	localisation (15+15)
	odd-one-out (15)		odd-one-out (15+15)
Total	30	104 or 90	60

Table 1. In parenthesis, the number of items for pre-test, training and post-test

2.1.1 Localisation of a given stress pattern

In this exercise, learners have to localise a given stress pattern. They hear trials of different words produced by a native Spanish female speaker. Each word is associated with a representative drawing. They have to answer to the following question: "Which word has stress on the X syllable". For example, as shown in Figure 1, the learners hear the words "*río*" (river), "*mesa*" (table) and "*champú*" (shampoo) and have to indicate which word has stress on the final syllable (i.e., "dernière syllabe " in French). The learners indicate their response by clicking on the drawing corresponding to the word with the given stress pattern. Before answering, they have the opportunity to listen to the words as many times as they want. They do not receive any feedback after each trial, but their score is displayed at the end of the exercise.

The task can target the final, penultimate or antepenultimate syllable. The difficulty of the task increases during the exercise when words produced by another (male) speaker are introduced, and by increasing the number of words presented in the trials (2, 3, 4 words) or/and the number of syllables in the words (2, 3, and 4 syllables).

Fifteen trials were used in the pre-test. In addition to these 15 trials, we used 15 new trials in the post-test in order to evaluate the generalisation of the training method to new items.

Figure 1: Localisation of a given stress pattern. Question "Click on the word which has stress on the final syllable"

2.1.2 Odd-One-Out

In this exercise, the learners perform an Odd-One-Out task (Schwab & Dellwo, 2016a). They hear a trial of three segmentally identical stimuli (e.g., *numero*). Among them, two stimuli present the same stress pattern (e.g., stress on the penultimate syllable) and one (i.e., the *odd*) presents a different stress pattern (e.g., stress on the final syllable). After hearing each trial, learners have to indicate which of the three elements is the *odd* one by clicking on the corresponding option on the screen (Mot 1, Mot 2, Mot 3; in English Word 1, Word 2, Word 3).

We used triplets of trisyllabic CV.CV.CV and CVC.CV.CV Spanish words that differ with respect to the stressed syllable. Each triplet is composed of a proparoxytone word (i.e., stress on the antepenultimate syllable; in trisyllabic words, it means that the first syllable is stressed, e.g., *número*), a paroxytone word (i.e., penultimate syllable stressed word, e.g., *numero*) and an

oxytone word (i.e., final syllable stressed word, e.g., *numeró*). Two native speakers of Spanish (one female and one male) produced the six words twice. The trials presented to the learners are composed of two words with the same stress pattern (e.g., *número*) and of one word with a different stress pattern (e.g., *numero*). The word with the stress accentual pattern is the odd element.

All combinations of stress contrasts are tested (i.e., proparoxytone target word paired with paroxytone words or with oxytone words; paroxytone target word paired with proparoxytone words or with oxytone words; oxytone target word paired with proparoxytone words or with paroxytone words). The position of the odd word within the trial is assigned randomly. The difficulty of the task increases in the second part of the exercise when words produced by the male speaker are introduced. Fifteen trials were also used in the pre-test for this task. In addition to these 15 trials, we used 15 new trials in the post-test for the same reason as in 2.1.1.

2.2 Training

The training *per se* is the principal, and therefore the most time-consuming, activity in the whole process. We present two versions of the training here. The first version is an adapted replication of the training used in Schwab & Dellwo (2016b), while the second version has a more explicit methodology with respect to prosodic patterns.

2.2.1 Version A

In this training phase, learners perform a shape/word matching task. They hear a word and 4 of 6 possible shapes appear on the screen (see Figure 2). Learners have to click on the shape which they think corresponds to the word they hear. After giving their response they receive feedback: they hear the word again and the correct shape is indicated with a green frame. If they had clicked on an incorrect shape, it is indicated with a red frame. This feedback enables the learners to learn the correspondence between the words and the shapes.

Two triplets of trisyllabic Spanish words are used in this training: *cáscara* (shell), *cascara* (en. that he cracked), *cascará* (he will crack) and *módulo* (module), *modulo* (I modulate), *moduló* (he modulated); these are associated to 6 the shapes as shown in Figure 2.

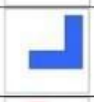
Shape	Word
	cáscara
	cascara
	cascará
	módulo
	modulo
	modulo

Figure 2: Shapes and words used in the training.

As can be observed, each triplet is composed of a proparoxytone (i.e., *cáscara* and *módulo*), a paroxytone (i.e., *cascara* and *modulo*) and an oxytone word (i.e., *cascará* and *modulo*). The six words are produced by another female native Spanish speaker (i.e., different to the one who recorded the stimuli in the pre-tests).

The training is divided into three blocks. Each word (e.g., cáscara) is presented six times per block (6 words x 6 times = 36 times per block). Among the four shapes that appear on the screen, only one corresponds to the word and the three others are distractors. Among the three distractors, one corresponds to a word with a different stress pattern (e.g., *cascara*) and the two other shapes correspond to two words from the other triplet (e.g., *módulo* and *modulo*). The position of the shapes on the screen is assigned randomly. The order of presentation of the trials is determined semi-randomly in such a way that no more than two same stress patterns and no more than two members of each triplet follow each other.

2.2.2 Version B

In version B of the training phase, learners perform a shape/stress pattern matching task. They hear a word and three shapes appear on the screen (see Figure 3). The shape iconically represents a trisyllabic word with stress on one of the three syllables (see Figure 3). The participants have to click on the shape they think corresponds to the stress pattern they hear. After giving their response, they receive feedback: they hear the word again and the correct shape is indicated with a green frame. If they had clicked on an incorrect shape, it is indicated with a red frame. The feedback enables the learners to learn the correspondence between the stress patterns and the shapes.

Ten triplets of trisyllabic Spanish words are used in this training pathway. Each triplet is composed of a proparoxytone (e.g., *cálculo*), a paroxytone (e.g., *calculo*) and an oxytone word (e.g., *calculó*). The 30 words are produced by a female native Spanish speaker (the same who produced the stimuli in version A). The three stress patterns are associated to three shapes, as shown in Figure 3.

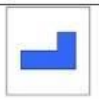
Shape	Stress pattern
	Proparoxytone (e.g., cálculo)
	Paroxytone (e.g., calculo)
	Oxytone (e.g., calculo)

Figure 3: Shapes and stress patterns used in the training version B.

Similarly to version A, the training is divided into three blocks. Each word (e.g., *cálculo*) is presented once per block. Among the three shapes that appear on the screen, only one corresponds to the target stress pattern and the two others correspond to the other two stress patterns. The position of the shapes on the screen is assigned randomly. The order of presentation of the trials is determined semi-randomly in such a way that no more than two same stress patterns, and no more than two members of each triplet, follow each other.

3. Preliminary results

The platform was launched in mid-September 2017. Six first-testers went through the whole training process (pre-tests, training and post-tests) and gave us some useful qualitative feedback in respect of user experience.

The platform uses an A/B testing mechanism, so that every other registered participant is led alternatively to version A or version B of the training. For version A, the mean scores (i.e correct identification of the appropriate shape) increased from 56% at the pre-test to 72% at the post-test, i.e., an improvement of 16%. For version B, the mean scores increased from 66% to 74%, i.e., an improvement of 8%. In both versions, we could notice that the best improvement was as high as 20%.

We also noticed a ceiling effect at 80% for the post-test scores, which limited the degree of possible improvement for speakers with a higher ability level in L2 Spanish. Finally, we estimated that the total time that the participants spent on the whole training was 30 minutes.

4. Conclusion

With the implementation of the web platform, we now have a framework which allows us to test several methodological configurations in parallel and compare their teaching impact. For example, we can address the question of implicit vs explicit teaching (as seen in Schwab 2016b) in order to test the effectiveness of explicit instructions. In the explicit training, explicit instruction and explanations about Spanish lexical stress contrasts would be given to the participants, whereas no mention to Spanish lexical stress would be made in the implicit training.

Our two main short-term tasks are 1) to continue taking the qualitative feedback of our testers into account in order to refine the didactic instructions and the general user experience, and 2) to deploy the platform to a larger number of participants in order to consolidate our preliminary results. We are also considering expanding this platform for any L1 by localising it in other languages. Further we also aim to expand the L2 material in other free-lexical-stress languages such as English and German, as well as tone languages such as Mandarin. Finally, another set of exercises that focus on production - rather than perception - is currently under development and will allow participants to speak and have their intonation automatically compared with reference stimuli.

5. Acknowledgements

This project is supported by the INNOGAP proof-of-principle fund from UNITEC, University of Geneva.

6. References

Anderson-Hsieh, J. & Koehler, K. (1988). The Effect of Foreign Accent and Speaking Rate on Native Speaker Comprehension. *Language Learning*, 38, 561–613.

Bagou, O., Fougeron, C. & Frauenfelder, U. H. (2002). Contribution of prosody to the segmentation and storage of "words" in the acquisition of a new mini-language. *Speech Prosody 2002*, Aix-en-Provence, France, 159-162.

Barquero, Mª A., Racine, I., Baqué, L. & Schwab, S. (2014). La estructuración acentual: estudio comparativo en la interlengua español-francés. Caso de la lectura. In: Y. Congosto Martín, M. L. Montero Curiel & A. Salvador Plans (Eds.), *Fonética experimental, educación superior e investigación*. (Vol. II, pp. 9-28). Madrid: Arco/Libros.

Busà, M. G. (2008). New perspectives in teaching pronunciation. In *From DIDACTAS to ECOLINGUA. An ongoing research project on translation and corpus*

linguistics. Trieste: Università degli Studi di Trieste, 165-182.

Chun, D. M. (2012). Computer-Assisted Pronunciation Teaching. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Oxford: Wiley-Blackwell. doi:10.1002/9781405198431.wbeal0172

Cutler, A. & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, 31, 218-236.

Cutler, A. (1984). Stress and accent in language production and understanding. In: D. Gibbon & H. Richter (eds.) *Intonation, Accent and Rhythm: Studies in Discourse Phonology*. Berlin: Walter de Gruyter.

Dailey, R. M., Giles, H. & Jansma, L. L (2005). Language attitudes in an Anglo-Hispanic context: the role of the linguistic landscape. *Language and Communication*, 25, 27–38.

Dupoux, E., Pallier, C., Sebastián, N. & Mehler, J. (1997). A destressing 'deafness' in French? Journal of Memory and Language, 36(3), 406-421.

Lev-Ari, S. & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, 46, 1093-1096.

Muñoz Garcia, M. (2010). *La perception et la production de l'accent lexical de l'espagnol par des francophones: aspects phonétiques et psycholinguistes*. Thèse de doctorat, U. Toulouse 2/Universitat Autònoma de Barcelona.

Munro M. J. & Derwing T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45, 73–97.

Neri, A., Cucchiarini, C., Strik, H., Boves, L. (2002). The pedagogy-technology interface in Computer-Assisted Pronunciation Training, *Computer-Assisted Language Learning* 15, 5: 441-467.

Niebuhr, O. (2017), Comparing visualization techniques for learning second language prosody. First results. *International Journal of Learner Corpus Research* 3:2 (2017), 250–277.

Purnell, T, Idsardi W. & Baugh J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology*, 18, 10-30.

Sztahó, D., Kiss, G. Czap, L. & Vicsi, L. (2014). A Computer-Assisted Prosody Pronunciation Teaching System. Fourth Workshop on Child Computer Interaction (WOCCI 2014), Singapore, September 19, 2014.

Schwab, S. & Dellwo, V. (2016a). The use of the Odd-One-Out task in the study of the perception of lexical stress in Spanish by German-speaking listeners. *Proceedings of the 6th International Conference on Speech Prosody*, Boston, May, 31 – June, 3, 2016.

Schwab, S. & Dellwo, V. (2016b). Explicit versus implicit training methods in the learning of Spanish accentual contrasts by French- and German-speaking listeners. *Second Language Prosody Workshop*, November, 18-19, 2016, Graz, Austria.

Vakil, A. S. & Trouvain, J. (2015). Automatic classification of lexical stress errors for German CAPT. SLaTE 2015, Leipzig, September 4–5, 2015.