



Integrating information technology in the teaching of English pronunciation: Designing and implementing an online course to teach word and sentence stress to tertiary level students.

Nadia Kebboua, Joaquín Romero

Rovira i Virgili University, Spain

nadia.kebbouac@estudiants.urv.cat, joaquin.romero@urv.cat

Abstract

This study explores students' progress in English pronunciation as far as stress is concerned, by creating online materials to develop their pronunciation skills. The study was carried out with two groups of first year university students. The activities were divided into four sections: overview, background, perception and production. While the experimental group recorded themselves using an integrated recording tool in the website, listened to the native speaker and then reflected on the differences between their production and the native speaker production, the control group listened to a native speaker and repeated the words afterwards. Both groups took a level test and a pre-test. After finishing the training, they took a post-test. Their recordings were evaluated using acoustic analysis with the Praat speech analysis software [1]. Each syllable in the words and phrases was identified in the acoustic signal and a reading of the intensity peak was obtained. These values were then processed in order to obtain the difference in amplitude between the stressed and the unstressed syllables. Preliminary results show that while both groups improved their pronunciation, the improvement was more pronounced in the experimental group. This seems to provide evidence of the benefits of the use of information technologies in the learning of pronunciation in the classroom.

Keywords: Keywords- CAPT (computer assisted pronunciation training), word stress, sentence stress, speech analysis software

1 Introduction

Teaching pronunciation has often been a neglected area of teaching English as a second or foreign language [2, 3]. The majority of teachers justify the little room devoted to teaching pronunciation because of its difficulty [4, 5]. Consequently, the negligence of teaching this important aspect of the language

leads many learners to communication breakdown with native speakers [6]. Furthermore, most textbooks and syllabi in which pronunciation is dealt with focus mainly on segmental features; suprasegmental features are hardly ever dealt with unless particular manuals are used. However, and according to many researchers, the misuse of suprasegmentals may lead to misunderstandings between native and non-native speakers of English. For instance, the communication failure between native and non-native speakers of English was blamed on the misplacement of stress in words [7].

Conversely, many advanced learners of English are sensitive now to the importance of acquiring if not a native-like pronunciation at least a correct and an appropriate pronunciation which will enable them to communicate effectively in the target language. Most of them have already had frustrating experiences in which they failed to make their messages understood because of their English pronunciation. Unfortunately, and in many cases, teachers give pronunciation a low priority even if students are willing to work hard to improve it.

On the other hand, a large body of literature claims that computer assisted pronunciation teaching can be advantageous and beneficial to support and assist learners acquire new skills by offering visuals, natives speaker voices and models for students to follow and compare their production with [8, 9, 10]. Following this notion, this study investigates the effectiveness of the use of computer assisted pronunciation training by testing how students at tertiary level can improve their pronunciation learning skills by recording themselves using an integrated recording tool in an online pronunciation course.

2 Method

In order to test the hypothesis that using computer-assisted language learning (CALL) and new technologies in the classroom while teaching/learning

pronunciation will benefit students in improving their pronunciation, an experiment was conducted that included two groups of participants: experimental and control groups. Before they commenced the treatment, participants carried out an English level test, and a speaking pre-test. After the training, they answered a qualitative questionnaire and carried out a post-test. Learning Pronunciation Skills was the name given to the training course, which was integrated in their usual course of English. The course was not graded though, as the aim was mainly to train them rather than to test them. After collecting the data, two different approaches were carried out to analyse it: quantitative and qualitative.

2.1 The participants

Thirty-four first year university students participated in this study. They were enrolled in either a Catalan or a Spanish language and literature degree at Rovira i Virgili University, in Tarragona (Spain). Almost all the participants were Spanish and Catalan bilingual speakers. They were aged between 19 and 21 years old. At the beginning of the course there were 34 participants, 4 males and 30 females, but due to external factors, only 27 students completed the study, of which 3 males and 24 females.

2.2 Stimuli (teaching materials)

All the materials were original, designed specifically for university Spanish and Catalan students, taking into account the typical mistakes some of these first language speakers make while speaking in English as far as word and sentence stress are concerned. After creating all the materials to be used for this particular group of participants, they were made available online using the Scorm 1.2 compliant package; they were developed to be delivered through LMS files. These files are able to store/deliver students' learning progress and interactions. As far as the activities are concerned, they were divided in four major units, i.e., overview, background, perception and production. In each unit, students were required to accomplish a set of interactive activities. The course was taught in 8 sessions, one hour per week.

Both the experimental and the control groups were monitored and guided throughout the whole course by the instructor/researcher; both groups did exactly the same activities, the only difference between the two groups was that, while the control group did the activities in a regular classroom and in a conventional way, using pen and paper, the experimental group did the activities online using computers; they were taught in a computer lab and used external microphones to record themselves.

Throughout the training course, participants of both groups received constant and detailed instructions

and guidance. However, the experimental group received extra guidance on how to use the website, the recording tool and the interactive activities.

Figure 1 shows a screen shot of the recording tool integrated in the website the experimental group made use of.

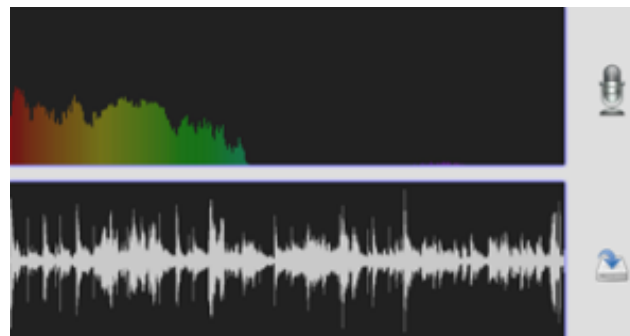


Figure 1. The html 5 recording tool showing a spectrum form display (upper panel) and a waveform (lower panel).

To record themselves, participants had to click on the “microphone icon”, which became red instantly. Once they started recording, they could see waves which showed the amplitude of their voices. When they finished recording, they clicked on the “save icon” shown in Figure 1 to save the file in their computers, which they uploaded afterwards on Moodle so that they could go back to it whenever they needed to.

Throughout the four units, the experimental group recorded a set of words, phrases, sentences and short texts. The control group, on the other hand, listened to a native speaker and repeated the words, phrases and sentences.

2.2.1 Overview

In this section, the main objective was to draw students' attention to the differences in word and sentence stress between Spanish and English. As the rules of stress in Spanish are quite predictable, it was very important to make them aware that stress in English does not have as well-established rules as in Spanish.

2.2.2 Background

This section was the longest one of the whole training course, its objective being to make students reflect on the differences in the rules of stress between Spanish and English, Spanish being a syllable-timed language and English being a stressed-timed language. Most of the participants had the tendency to apply the well established rules of stress of Spanish while saying words in English, mainly the words of Latin origin that have a counterpart in Spanish. They were provided with some patterns which most of the words of the same category share as far as stress is

concerned. It was hoped that students would realise that stress in English is not as random and difficult to predict as it may seem.

2.2.3 Perception

In this section, students did a set of interactive activities the aim of which was to make them reflect on some of the reasons why they sometimes fail to understand native speakers' speech even though they are familiar with the language used. To achieve it, students did four different activities.

2.2.4 Production

In the last part of the training, production, students were provided with a set of activities to make them reflect on the reason why sometimes native speakers do not understand their speech even if they use the right words and grammar.

2.3 Data analysis

After students had done the pre-test, the training, the post-test, and the questionnaire, all the recordings were edited using the Audacity software. Due to time and space limitation, a decision was made to analyse only a part of the pre-test, the post-test and the questionnaire. In this paper, discussion will be limited to the analysis of the pre- and post-test.

2.3.1 Acoustic analysis of data

A part of the recordings was evaluated using acoustic analysis with the Praat speech analysis software [1]. Even though in the actual test participants had read and recorded ten words, four phrases, ten sentences, and a short text, only the words and phrases were analysed. The analysis was carried out for each participant and then per group at each testing phase (pre-test and post-test). Each syllable in the words and phrases was identified in the acoustic signal and a reading of the intensity peak was obtained. These values were then processed in order to obtain the difference in amplitude between the stressed and the unstressed syllables. Amplitude, rather than pitch values, were used because amplitude is a more relevant aspect of Spanish stress than pitch, so it was thought that it would reflect Spanish/Catalan speakers attempts at pronouncing English stress more precisely than pitch. All the values were transferred by default to an intensity log document, which was then copied to an Excel file in which more detailed info was added: the initials of the name of the participant, the number of activity, the word measured, and the time of the intensity peak of the syllable. In Figure 2 the word *doctor* is shown. The red bars show the peak of the two syllables that the word has. The difference of amplitude between the stressed syllables and the

unstressed syllable is clearly visible.

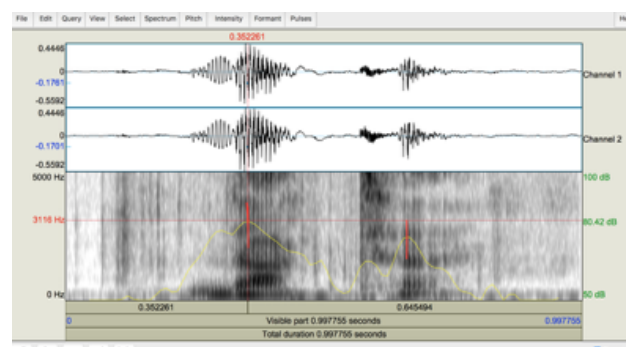


Figure 2. An example showing a waveform display (upper panel) and a spectrogram with intensity curve (lower panel).

The scores obtained were calculated by subtracting the value of the unstressed syllables from the stressed syllables. Thus, the scores were calculated and compared for each participant, in the pre- versus post-test, and per each group, in the control versus the experimental group. The decision of comparing the pre-test and the post-test was made assuming that at the pre-test both groups were homogeneous in terms of their perception and production skills. All scores were submitted to statistical analyses.

Figure 3 shows the mean scores for both the control and the experimental groups. The improvement is clearly more pronounced in the experimental group.

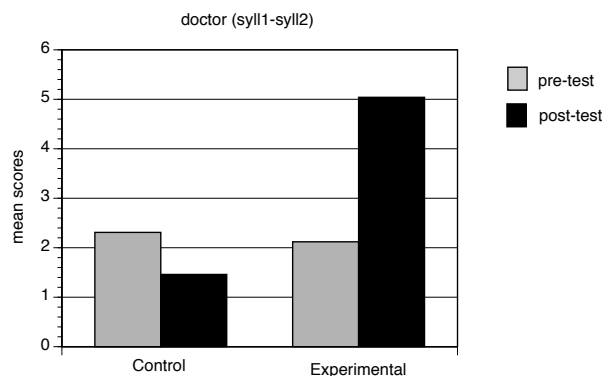


Figure 3. Mean scores of the values obtained in the pre-test and post-test for the word “*doctor*”.

2.3.2 Acoustic measurement of words

The words and phrases recorded in the pre- and the post-tests were analysed and calculated in order to obtain the gain scores by subtracting the stressed syllables from the unstressed syllables. In two syllabic words, only one analysis was required in order to compare the amplitude between the stressed and the unstressed syllables. For instance, the first word analysed was *doctor*. As the stressed syllable falls on the first one, the second syllable was subtracted from the first one. On the other hand, in words with more than two syllables, more than one

analysis was needed. For example, for the word *vegetable*, this being a four syllable word, three different analyses were required. Thus, the first analysis consisted of subtracting the second syllable from the first one; in the second analysis, the third syllable was subtracted from the first one and in the third analysis, the fourth syllable was subtracted from the first one. The third word analysed was *chocolate*, for which two analyses were needed. For the phrase *many vegetables*, as this phrase was out of context, it was assumed that the stressed word, the word which contains the most information, was “*vegetables*”, the stressed syllable of which is the first one. Accordingly, five different analyses were required in order to calculate the difference of amplitude between the stressed syllable, the first one, and the unstressed syllables, the other four syllables. Hence, in the fourth analysis, for example, the third syllable was subtracted from the fifth one, and in the fifth analysis, the sixth syllable was subtracted from the third one.

3 Results

Tables 1 and 2 show the group changes in the pronunciation of the words and phrases analysed as far as word stress is concerned. The numbers in the first column refer to the position of the unstressed syllables which were subtracted from the stressed syllables.

Table 1: An example of the mean scores the control group obtained in the pre- and post-test

| Words and phrases analysed | Control pre-test | Control post-test | Gain scores |
|----------------------------|------------------|-------------------|-------------|
| doctor (1-2) | 2,31 | 1,46 | -0,85 |
| vegetable (1-2) | 2,54 | 5,48 | 2,94 |
| vegetable (1-3) | 2,09 | 5,22 | 3,13 |
| vegetable (1-4) | 1,57 | 3,1 | 1,53 |
| chocolate (1-2) | 2,98 | 6,97 | 3,99 |
| many veg (3-5) | 3,88 | 5,5 | 1,62 |
| many veg (3-6) | 4,9 | 4,2 | -0,7 |

If we compare the mean scores of the values both groups obtained in the pre- and post- conditions, we observe that the level of discrimination between the stressed and unstressed syllables was more pronounced in the experimental group than it was in the control group.

Table 2: An example of the mean scores the experimental group obtained in the pre and post-test

| Words and phrases analysed | Exp pre-test | Exp post-test | Gain scores |
|----------------------------|--------------|---------------|-------------|
| doctor (1-2) | 1,48 | 4,99 | 3,51 |
| vegetable (1-2) | 0,42 | 6,94 | 6,52 |
| vegetable (1-3) | -0,41 | 5,95 | 5,54 |
| vegetable (1-4) | -0,62 | 7,69 | 7,07 |
| chocolate (1-2) | 2,92 | 8,36 | 5,44 |
| many veg (3-5) | 2,9 | 5,93 | 3,03 |
| many veg (3-6) | 2,32 | 8 | 5,68 |

In the word *vegetable*, for instance, mainly in the first and third analyses, even though both groups improved, the gain scores the experimental group obtained were almost three times higher than for the control group. Thus, in the first analysis, the gain score of the control group was 2,94, while in the experimental group it was 6,52. In the third analysis, the gain score of the control group was 1,53, whereas in the experimental group it was 7,07. The same applies to the phrase *many vegetables*, the control group obtained a -0,7 gain score, against 5,68 the experimental group obtained.

3.1 Statistical analysis of words

Four different t-tests were conducted in order to analyse statistically the data obtained from the acoustic analysis: two paired samples t-tests to compare the pre- and post- scores within the same group and two independent samples t-tests to compare scores across the two groups.

Table 3: An example of the results (p. values) obtained using t-tests. The rows correspond to the words and phrases shown in Tables 1 and 2

| Control pre. vs. Control post | Experimental pre vs. experimental post | Control pre vs. experimental pre | Control post vs. experimental post |
|-------------------------------|--|----------------------------------|------------------------------------|
| 0,6601 | 0,1277 | 0,9032 | 0,1434 |
| 0,0177 | 0,0033 | 0,3550 | 0,1795 |
| 0,1301 | 0,0261 | 0,2481 | 0,6360 |
| 0,3907 | 0,0156 | 0,4215 | 0,0890 |
| 0,1588 | 0,0453 | 0,7668 | 0,3436 |
| 0,1779 | 0,0849 | 0,6511 | 0,7812 |
| 0,7682 | 0,0207 | 0,2789 | 0,1648 |

Table 3 shows some of the results obtained from the t-tests. In the first word analysed, *doctor*, even if the results were not significant, the improvement was more pronounced in the experimental group: $M = 1.37$, $SD = 4.59$ in the pre-test and $M = 2.12$, $SD = 1.78$ in the post-test; $t(11) = -1.64$, $p = .127$. The control group, on the other hand, obtained $M = .93$, $SD = 6.44$ in the pre-test, and $M = 1.37$, $SD = 4.59$ in the post-test conditions; $t(11) = .500$, $p = .66$. In the first analysis of the second word, *vegetable*, the

results were significant for both groups in the pre- vs. post- comparisons, favouring the experimental group: $M = 0.42$, $SD = 5.34$ in the pre-test and $M = 6.94$, $SD = 5.60$ in the post-test; $t(11) = -3.44$, $p < .05$ compared to the control group, $M = 2.54$, $SD = 4.77$ in the pre-test and $M = 5.48$, $SD = 4.70$ in the post-test; $t(11) = -2.78$, $p < .05$.

In the second and the third analyses of the second word, the results were significant for the experimental group in the pre- vs. post- comparison: $M = -0.41$, $SD = 5.13$ in the pre-test and $M = 5.95$, $SD = 5.03$ in the post-test; $t(11) = -2.56$, $p < .05$ for the second analysis, and $M = -0.62$, $SD = 5.72$ in the pre-test and $M = 7.69$, $SD = 7.85$ in the post-test; $t(11) = -2.85$, $p < .05$, for the third analysis. The control group, on the other hand, obtained $M = 2.09$, $SD = 4.94$ in the pre-test and $M = 5.22$, $SD = 7.24$ in the post-test; $t(11) = -1.63$, $p = .130$ for the second analysis, and $M = 1.57$, $SD = 5.47$ in the pre-test and $M = 3.10$, $SD = 5.57$ in the post-test; $t(11) = -0.893$, $p = .390$, for the third analysis.

The same situation can be observed in the first analysis of the third word, i.e., *chocolate*, the results for which were significant for the experimental group in the pre- vs. post- comparison: $M = 2.92$, $SD = 4.84$ in the pre-test and $M = 8.36$, $SD = 6.69$ in the post-test; $t(11) = -2.25$, $p < .05$ compared to the control group $M = 2.98$, $SD = 8.03$ in the pre-test and $M = 6.97$, $SD = 5.73$ in the post-test; $t(11) = 1.512$, $p = .158$.

In the fourth analysis of the phrase *many vegetables*, even if the results were not significant for either group, the improvement was more pronounced in the experimental group: $M = 2.90$, $SD = 4.33$ in the pre-test and $M = 5.93$, $SD = 5.13$ in the post-test; $t(11) = -1.89$, $p = .084$, compared with the control group: $M = 3.88$, $SD = 3.38$ in the pre-test and $M = 5.50$, $SD = 4.62$ in the post-test; $t(11) = -1.43$, $p = .177$.

In the last analysis illustrated in tables 1 and 2, the results were significant for the experimental group: $M = 3.15$, $SD = 4.35$ in the pre-test, and $M = 6.08$, $SD = 5.51$ in the post-test; $t(11) = -2.69$, $p < .05$, compared to the control group: $M = 4.90$, $SD = 5.1$ in the pre-test and $M = 4.2$, $SD = 6.03$ in the post-test; $t(11) = .285$, $p = .76$.

4 Discussion

The main research hypothesis of this study dealt with the effects of the use of computer assisted pronunciation training in the classroom. That is to say, the effects obtained from the use of the integrated recording tool in the online pronunciation course. One of the reasons for creating the materials for this specific group of participants was to make them reflect on the differences between stress in Spanish

and English. Thus, most of the words recorded have their counterpart in Spanish, as most participants tend to transfer the position of stress of these words from their first languages to their English counterparts. As the results show, in most of the words and phrases analysed, even though both control and experimental groups show improvement in the production of the difference of amplitude between the stressed and unstressed syllables within words and phrases, the improvement was much more pronounced in the experimental group, which got significant results in most cases.

5 Conclusion

The findings of this study can be taken into consideration while designing a pronunciation course for Spanish and Catalan tertiary level students. Since the duration of the treatment was limited to only eight hours of training, it is speculated that dedicating more time of explicit instruction on pronunciation will result in higher awareness about their mistakes, which, in turn, should result in an improvement of their pronunciation as far as word and sentence stress are concerned. To conclude, the results seem to provide evidence of the benefits of the use of information technologies in the learning of pronunciation in the classroom.

6 Acknowledgment

This research was partly funded by project FFI2017-84479-P from the Spanish Ministry of Economy and Business.

7 References

- [1] Boersma, Paul & Weenink, David (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.40, retrieved 11 May 2018 from <http://www.praat.org/>
- [2] Rubin, D. (1992). Nonlanguage factors affecting undergraduates' judgements of nonnative English-speaking teaching assistants. *Research in Higher Education*, 33, 511-31.
- [3] Scovel, T. (1969). Foreign accent: Language acquisition and cerebral dominance. *Language learning*, 19, 245-254
- [4] Henderson, A., Frost, D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, A., Waniek-Klimeczak, E., Levey, D., Cunningham, U. and Curnick, L. (2012). The English pronunciation teaching in Europe survey: Selected results. *Research in Language*, 10 (1): 5-27.
- [5] Macdonald, S. (2002). Pronunciation – views and practices of reluctant teachers. *Prospect*, 17 (3): 3-18.
- [6] Jenkins, Jennifer (2000) *The Phonology of English as an International Language*. Oxford: Oxford University Press.
- [7] Benrabah, M. E. (1997). Word stress – A source of unintelligibility in English. *International Review of Applied Linguistics*, 35(3), 157 – 165.
- [8] Celce-Murcia, M., Brinton, D. M., Goodwin, J. M. & Griner, B. (2011). *Teaching pronunciation* (2nd ed.). New York: Cambridge University Press.
- [9] Neri, A., Cucchiari, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in Computer Assisted

- Pronunciation Training. *Computer Assisted Language Learning*, 15, (5), 441-467.
- [10] Pennington, M. C. (1999). Computer-aided pronunciation pedagogy: Promise, limitations, directions. *Computer Assisted Language Learning*, 12 (5), 427-440.