

Teaching and Learning English Phonetics and Phonology in the EHEA through Design and Creation of Audiovisual Material at the ULPGC

Francisco Vizcaíno-Ortega, Mercedes Cabrera-Abreu

IATEXT, Instituto Universitario de Análisis y Aplicaciones Textuales, ULPGC, Las Palmas de
Gran Canaria, Spain
{francisco.vizcaino,mercedes.cabrera}@ulpgc.es

Abstract. In this paper we describe the procedure followed in the design and recording of a set of videos for teaching and learning ‘English phonetics and phonology’, a second-year undergraduate course at Universidad de Las Palmas de Gran Canaria. The student’s L1 is Spanish. Two different types of technological support were used: screencast and Powerpoint® presentations. The traditional whiteboard together with the lecturer’s presence also contributed both to the integrated learning of certain acoustic/articulatory aspects of the course contents and to the use of specific software for speech analysis. This video production owns the advantage of being an interactive and autonomous tool which favours a continuous learning process on the student’s side.

Keywords: phonetics and phonology · teaching-learning process · audiovisual material · correlation between acoustic cues and articulatory events

1 Introduction

Fonética y Fonología Inglesas is a second-year level subject within the four-year undergraduate programme in the degree of Modern Languages at Universidad de Las Palmas de Gran Canaria (henceforth ULPGC), where students major in English and minor in either French or Chinese. Students who register for this course, which is compulsory for all the students in the degree, are level B1 plus students who all too often struggle with the pronunciation of English and whose prior knowledge of this topic, if any, is very limited. What is more, even though they may have a relatively good command of the language grammar and vocabulary, most of these students have a hard time pronouncing English words either because they have recurrently mispronounced them or because an adequate pronunciation programme has been largely ignored in the teaching-learning process.

The course, which is part of the second-year curriculum in semester 1, is taught 4 hours per week; it is, therefore, equivalent to 60 hours of class time and counts for 6 ECTS (3 theory and 3 laboratory practice). Given the large number of students en-

rolled in this course every year – around 120 in theory sessions – the class must be divided into two lab groups, both in the morning and in the evening shift, for the purpose of offering students the opportunity to apply the theoretical knowledge acquired to the analysis of their own pronunciation and that of others in terms of perception and production. Due to the aforementioned number of students, theory sessions are mostly lectures generally assumed to proceed in a unilateral fashion and in which it becomes almost impracticable to organize seminar classes which require students to actively engage in the issues studied.

In order to encourage active participation among students, the practical lab sessions have been designed in such a way that they may work individually or in pairs on several different tasks, ranging from articulatory descriptions of sounds and the analysis of their corresponding acoustic indices in both spectrograms and waveforms, to phonemic and phonetic transcriptions of single words and short utterances across different varieties of English. Perception tests dealing with vowels and consonants - especially those that might pose problems for Spanish speakers of English - are accompanied by continuous audio recordings made by the students themselves in an attempt to improve their pronunciation. These lab sessions allow the students to deepen their understanding of the notions and phenomena studied in the large lecture class and, most importantly, to internalize the necessary knowledge to monitor their progress in oral production. Undertaking all these tasks is essential if students are expected, at the end of the semester, to master the competences included in the course syllabus. Some of these core competencies are [1] expertise in both articulatory and acoustic analyses of the sounds of English; [2] ability to diagnose and correct their pronunciation of English; [3] program-specific computer competency (speech analysis software). Bearing in mind the wide variety of tasks in this course as well as the achievement of the goals just listed, the need for audiovisual material that meets the specific learning needs of our students becomes obvious. Committing ourselves to designing and creating video material that complies with the holistic view of teaching English phonetics and phonology as described above served as the departure point for the methodological proposal expounded in the present paper and developed in the next sections.

2 Contents of the course *Fonética y Fonología Inglesas* and the specific needs of our students at ULPGC

For the students of the degree in Modern Languages this course provides the first contact with the formal description of the sounds of English and a great opportunity to identify their pronunciation errors and use the necessary resources to correct them. Internet search results reveal that there is often a division into articulatory phonetics on the one hand and acoustic phonetics on the other hand. There is a large number of websites which focus on the anatomy and physiology of the human vocal tract while

others are mostly acoustic-oriented. In both cases the video material used reflects either of these orientations. Students can also visit many web pages where they can find helpful learning material to get started with phonemic and phonetic transcription. Finding practical exercises on the perception of English sounds, however, proves more difficult, and the presence of educational videos in this area is almost non-existent.

While it is true that articulatory phonetics gives us a reliable description of the production of some sounds (especially those involving contact between the active and passive articulators), we must recognise that it is frequently useless not only for second language learners but also for native speakers, whose perception of their own articulators is far from being clear (Llisterri 1998). Articulatory phonetics requires strong support from acoustic phonetics, for the latter has proved to be extremely useful in the correct identification and classification of sounds – so much so that it performs a crucial role in our course.

Therefore, in view of this scenario, we, as instructors willing to maximize the student's learning process, realized that a series of videos was needed which dealt with both articulatory and acoustic issues as well as with transcription and perception matters. As for the assessment of the contents of our course and its core competencies, such videos are aimed at the evaluation of specific tasks (acoustic cue identification and their correlation with a gross articulatory gesture; production of phenomena like aspiration, devoicing or clipping; vowel and consonant discrimination; differences between orthographic form and pronunciation) rather than at the appraisal of general knowledge.

3 Video production

The videos were produced with Camtasia®, and they exhibit three different types of support: i) screencast, ii) PowerPoint® presentations, and iii) whiteboard together with the lecturer's presence (see *References*). A screencast consists in a digital recording of the computer screen output together with audio-recording. Thus, the lecturer may use either the computer mouse to point at specific sections of what is being displayed on the screen, which is normally a specific software for speech analysis (as can be seen in figure 1 below), or a digital pen to write on a digital board, whose output is being recorded (as, for example in figure 2 below).

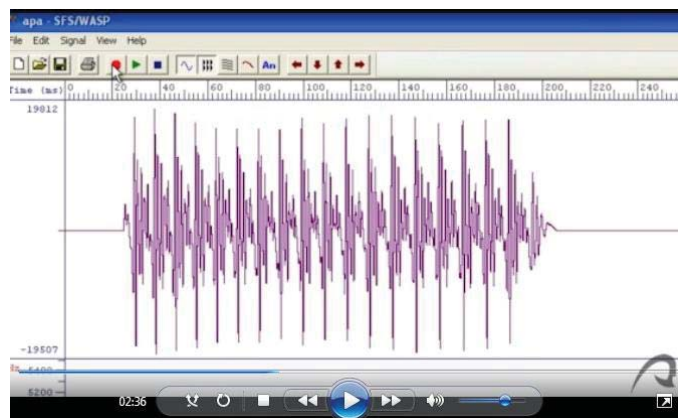


Fig. 1. Screencast video which illustrates the use of WASP, a software for speech analysis. The lecturer uses the mouse cursor to point at specific commands on the toolbar menu.

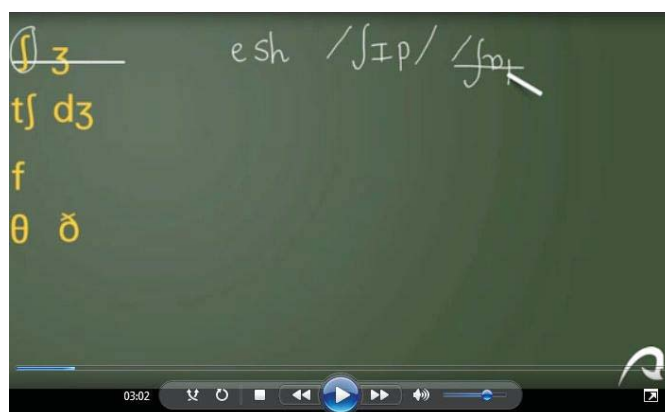


Fig. 2. Screencast video which illustrates the shapes and strokes of symbols for phonemic transcription. The lecturer uses a digital board, and her digital pen is visually represented by a white piece of chalk.

In both cases, the audio is also recorded. In the case of videos based on PowerPoint presentations®, such presentations included plain text and images of spectrograms and speech waveforms. Any kind of animations and transitions are excluded, since we decided to take advantage of the callouts provided with the software Camtasia® to either point or highlight areas of the displayed image wherever it was necessary. For instance, figure 3 offers an instance of a particular acoustic cue which is being described in the audio, and at the same time, it is highlighted in in the video display in order to closely guide the student in its identification. Video extension ranges between 5 and 10 minutes.

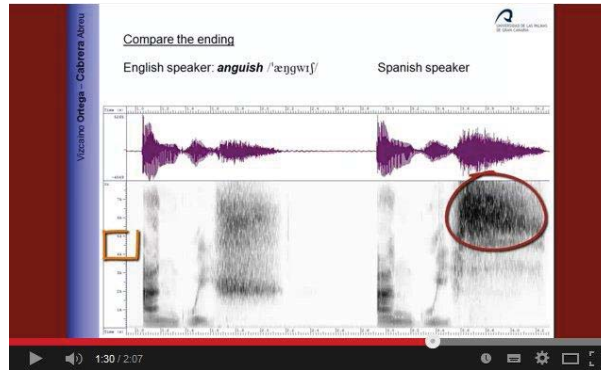


Fig. 3. Illustration of a video supported by a Powerpoint® presentation with callouts from Camtasia®.

The videos were filmed during various sessions. In the first one, the video headers were recorded bearing in mind that a short, clear, and concise sentence was necessary to introduce the topic in the video. In the later sessions, either a script was followed by the lecturer, or his/her performance was unscripted. In the latter case, a specific piece of material (a word for the analysis in terms of parametric diagrams displayed on a whiteboard) was used (see figure 4). The video project was constructed and edited with Camtasia®.

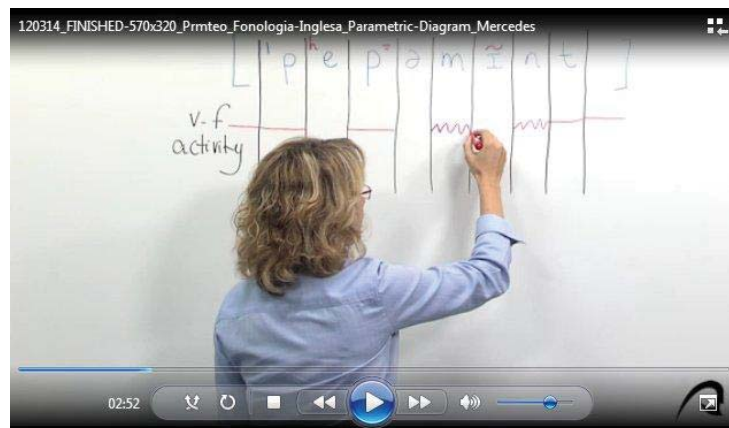


Fig. 4. Illustration of a video in which the lecturer uses the whiteboard for the description on how to draw a parametric diagram.

During the video recordings, there is a small degree of improvisation in order to keep a high degree of naturalness. On one occasion, the lecturer made a mistake during her transcription, but immediately noticed it and corrected it straight away. This sequence of actions could have been erased after the editing process of the video.

However, a decision not to cut such photogram was made, as we assume that students may also learn from others' mistakes.

3.1 Mock exam video with PowerPoint® presentations

Students can acquire the core competency 1 aforementioned, that is, expertise in both articulatory and acoustic analyses of the sounds of English, by following the lecturers' instructions in mock exam videos in which the PowerPoint® presentation format is used. These videos are intended to be uploaded to the virtual learning environment (Moodle platform) used at ULPG. In this way, long before taking the written exam, which is one of the course assessment tools, students can practise with one of the questions included in the exam paper. In the first slide, they are asked to observe two spectrograms displayed on the screen, one corresponding to the production of an English word by a native speaker and the other showing the pronunciation of a learner of English with a strong Spanish accent. Students have to compare these two spectrograms by using the clue provided. Thus, the first step in the procedure to solve the exercise is presented, and its aim is made explicit. Then, in a different slide, the following instructions necessary to complete the exercise continue to be read aloud by the lecturer (male voice): i) draw a circle in the spectrogram to show the differences between the two speakers; ii) use the spectrographic information to describe in acoustic terms the pronunciation error in the word produced by the Spanish learner; and iii) specify in acoustic terms how such error should be corrected by observing the acoustic cues of the word uttered by the native speaker. In the third slide, the first spectrogram is displayed yet again, and another lecturer with a female voice, begins to solve the exercise. Every time she makes reference to an acoustic cue displayed in the spectrogram or a phonetic symbol displayed outside the spectrogram, a callout appears on the screen. This is a ready-made basic, coloured, shape (a circle, a square or an arrow) used to visually communicate specific information (the relevant acoustic cue or the symbol whose sound contains such a cue), so that the student is permanently guided in every individual screen. Camtasia® transitions between slides are also incorporated as a visual effect which indicates the flow of time. Finally, the videos in this section include background music. Such music is found in various size audio clips included in the Camtasia® library which are aligned with different sections of the video; thus the video header, which includes an animated title, is aligned with a short audio clip (33 secs.), while the video body is aligned with a long audio clip (8 minutes).

By resorting to voices from different speakers, callouts, transitions, and music clips, the learning process is improved as communication is enhanced visually and auditorily.

3.2 Symbols for transcription with screencast

In order to achieve competency [2], that is, the ability to diagnose and correct student's own pronunciation of English, one type of rather basic but fundamental video was recorded, 'Symbols for transcription', with a three-fold purpose: 1) to write accurately the symbols for transcribing English; 2) to create awareness of the differences and similarities between symbols for transcription and symbols for writing; and 3) to practise with the transcription of monosyllabic or disyllabic words. Students are recommended to watch these videos during their autonomous learning sessions.

By using screencast, students see the actual stroke for the symbols in a monosyllabic word, together with orthographic representations as it is being drawn on a green board (the digital board) by the lecturer (see figure 2 above). This is accompanied by the lecturer's voice offering some warnings to avoid typical mistakes recurrently observed, like for instance, the stroke for *esh* (ʃ), which is frequently transcribed above the line, instead of crossing the line. In fact, the lecturer makes the mistake herself, and makes sure she crosses it out slowly, so that the student can observe the gesture, something which can serve to enhance the correct learning of transcription practice.

New terms can also be written on the board as the lecturer is talking, and they are recurrently used during the description; the repetition of the term, together with its use in its natural context helps in the acquisition of further knowledge without extra effort; in addition, the actual presence of the context also aids in the latter retrieval of the term.

Video watching is not simply sitting on a comfortable chair and watching a video. The videos we describe in the present paper are conceived as a tool to be used by the student while sitting at his own desk, so that he is expected to have paper and a pen available in order to join in with the lecturer, and perform his own short exercise on transcription. For this, the lecturer allows the student a few seconds to try the transcription, and resorts to the following expressions: 'Would you like to transcribe the word...?', or 'Why don't you try transcribing the word 'yet'?'. Notice, by the way, that there is also a frequent use of short pauses so that they clearly signal when the student is expected to interact in some way. So, the lecturer may resort to expressions like 'I'll give you a couple of seconds for you to do it'. Do you know how to transcribe the word 'taxi'? 'I'll give you one second to do it! Quick!'

Many students assume that transcription is rather boring, and this may be partially true. The point is, we believe, that during the initial stages of the acquisition of an adequate pronunciation, it stands as an essential tool which allows to learn how to pronounce words which are either completely new for the student, or known by the student. Unfortunately, in this latter case, they have been mispronounced ever since they were incorporated to his/her vocabulary. It is specifically at this stage when transcription is highly necessary and most effective. Once the student has developed a good command of the pronunciation of the language, however, he/she may dispense

with phonemic transcription altogether, or leave it for an occasional use when he encounters place names, or a more complex/advanced vocabulary. The videos we describe here have been designed for our students who, while they exhibit a B1 plus level of English, need a high degree of both exposure to pronunciation and encouragement to learn ways to predict systematically the pronunciation of words and short phrases. Thus, expressions like ‘Let’s have fun transcribing another complete word’ are used by the lecturer as a lively and ludic prompt to carry on with the transcription task; then, the exercise stands as yet another small challenge, since there is a movement forward from the transcription of an individual sound to a complete word.

The fact that occasionally the student is offered the opportunity to perform a short, simple task may wrongly lead on to the assumption that his/her attention is distracted from the video itself. Quite the contrary. Special care is taken to allow the student a short time interval (between five and ten seconds to transcribe a monosyllabic word, for instance) to turn away from the screen and into his/her own piece of paper, and also, with expressions like, ‘Alright, please watch the screen now’, he is kindly asked to return to the video display.

One of the great advantages of using videos is that the actions performed by the lecturer can be synchronized with his/her own oral description of the action itself. For instance, the lecturer accompanies the drawing of a stroke for a specific symbol with her voice; or in the drawing of the *esh* as compared with */s/*. In this way, the degree of accuracy in the drawing of phonetic symbols increases rapidly.

At the end of the video, the lecturer reminds the students of the names (which are a bit difficult) of a couple of symbols. In the same way as lecturers round up their sessions with a brief summary or by adding a final reminder of the main points presented during their lectures, this is also incorporated to the last few seconds in the video. In doing so, the lecturer reminds the student of the aims and issues of the video, without losing track of its purpose.

The degree of naturalness together with a clear intention on the lecturer’s side to keep the momentum flowing is observed through expressions like ‘Let me make sure that you know the word I’m referring to’ followed by the immediate action of orthographically writing the word referred to.

3.3 Parametric diagrams with whiteboard and lecturer’s presence

Also useful to develop the core competency [2] mentioned above are the videos illustrating the parametric diagrams. Such diagrams have been designed to be viewed by both students and lecturers in the practical lab sessions, where volunteers are invited to draw examples following the methodology proposed in the video itself. Thus, a segmented word in transcription form was written on the board for its analysis in terms of one of the following parameters: 1) presence or absence of vocal fold vibration; 2) velum activity; and 3) degrees of articulatory stricture for consonants: stop,

narrow approximation, and wide approximation. The parameters were individually analyzed in separate videos. During the recording, the lecturer stood in front of the board, making sure he/she pointed at the screen behind without completely turning around, or if he/she really needed to do so for writing on it, then he/she would make sure to remain silent.

Since a parametric diagram may consist in a stylized representation of the sounds in a given word, the lecturer always offered a very clear description of what the different formats of lines represent in each of the individual videos. For instance, in the video illustrating vocal fold activity, following Ashby (1995), a wiggly line corresponds to a stylized representation of vocal fold vibration in its corresponding diagram; in the video containing velum activity, following Ashby and Maidment (2005), one horizontal line followed by a slanted rising line represent a closing movement of the velum. After the lecturer finishes his/her description, an image of the whiteboard with the parametric diagram is frozen on the screen so that students have a chance to check the complete diagram at a single glance.

3.4 How to use WASP® with screencast

One of the videos is an online tutorial in which the programme WASP® (Huckvale 2007) is introduced in order for students to acquire core competency number [3], namely, program-specific computer competency (speech analysis software). This software allows students to both record and analyse their own productions of words and utterances presented in class. This video is one of the examples of screencast that we intend to use in our lab sessions; it is a digital recording of the computer screen output, enhanced with audio narration, which contains a step-by-step guide to the information the students need to start using the software. Thus, after informing the students about the installation process, they receive clear instructions concerning the functions of the menu bar. Then the default speech file that comes with the software is analyzed by dissecting every cue in the spectrogram or in the waveform that might help the students to recognize phonetic issues related to the acoustic manifestation of English vowels or consonants, whether they are associated with the parameters voice, place or manner (for consonants), voice, height, location or lip rounding (for vowels), or periodicity/apperiodicity for both types of sounds. At the end of the video the lecturer, as usual, invites the students to record their own renditions.

4 Conclusion

The videos described in the present paper offer the following advantages:

1. They include the answers to a selection of exercises proposed either in the lab sessions, or in the mock exams.

2. Since the videos produced with the support of Powerpoint® presentations were designed following a script read aloud by the lecturers, they contain a step by step description of the spectrograms, which students can take as their model for the description of other spectrograms.
3. Each video focuses on a specific piece of content in its own acoustic and articulatory context. This allows for an integrated learning, since students learn simultaneously how to use WASP and the acoustic cues of the phenomenon they are analyzing.
4. The videos are user friendly and lead the student smoothly towards a detailed learning, without feeling overwhelmed by excessive information.
5. The videos about symbols for transcription display real mistakes which are also recurrently made by students. After watching the video, there is a better chance that students might eventually reduce the number of errors in their transcriptions.
6. As a consequence of 5, the time formerly spent in the correction of transcription errors can now be invested in other practical exercises.
7. The videos can be watched, stopped, rewind, fastforward and re-played depending on the student's individual needs.
8. Students can monitor their own progress following the unfolding of the exercise, as they can stop the video and proceed onto making predictions of what comes next.

References

1. Huckvale, M. 2007. <http://www.phon.ucl.ac.uk/resources/sfs/wasp.htm> Last access 23rd October 2014
2. Ashby, P. 1995. *Speech Sounds*. London: Routledge.
3. Ashby, M. and Maidment, J. 2005. *Introducing Phonetic Science*. Cambridge: C.U.P.
4. Llisterri, J. 1998. "Algunas reflexiones en torno a la corrección fonética en E/LE". Instituto Cervantes – Universitat Autònoma de Barcelona.
5. http://bustreaming.ulpgc.es/bustreaming_reproducirpublicacion&idpublicacion=26423
6. <https://www.youtube.com/watch?v=mbG8W-QDT-c&feature=youtu.be>
7. http://bustreaming.ulpgc.es/bustreaming_reproducirpublicacion&idpublicacion=26487