Plugin for automatisation of phonetic-phonological analysis and obtaining analytical feedback for Spanish learners

Plugin para la automatización del análisis fonético-fonológico y la obtención de retroalimentación analítica para estudiantes de español

Tamara Couto-Fernández[1], Albina Sarymsakova[2], Nelly Condori-Fernández[3] and Patricia Martín-Rodilla[4]

[1][3][4] *University of A Coruña, Faculty of Computer Science, Camiño do Lagar de Castro, 6, A Coruña, 15008, Spain*
[2] *University of A Coruña, Faculty of Philology, Campus da Zapateira, A Coruña, 15008, Spain*

Abstract

We present in this article the Plugin for phonetic-phonological analysis in Spanish (PAFe), which consists of a series of scripts (a code written with a programming language (Python) that, implement three different intonation comparison algorithms of an ELE (Spanish as a foreign language) student and a native speaker of Spanish), allowing, in turn, three different types of analysis: global, tonal tendency and intersyllabic. In addition, PAFe has a database to keep a history of different types of data (user profile, pronunciation exercises and audios) and a graphical interface to include reports on pronunciation evolution in Praat, a tool for acoustic analysis. PAFe is a software solution that offers new functionalities of Praat and allows the following: (i) to perform a comparative analysis between the intonational patterns of an ELE student and a native speaker; (ii) to report the evolution of the acquisition of such patterns in Spanish thanks to the history of the stored data. In this way, automated feedback is provided to both students and teachers.

Keywords

Praat, intonation analysis, ICT, Python.

1. Introduction

The present work is framed in the area of natural language processing, specifically, in the comparative-contrastive analysis of intonation for the didactic purposes provided by our original tool PAFe. Despite the existence of some tools, such as the Oplustil and Toledo [11] proposal, or the study by Strik, Truong, Wet and Cucchiarini [8], which offer results of phonetic-phonological similarity or detect errors made in pronunciation.

Nonetheless, no tool provides both facilities at the same time, nor offers to monitor the evolution of the students.

For this reason, we have decided to develop a system that complements language teaching, in particular, one that can be used remotely or in hybrid modalities.

Our tool offers the functionality to perform an instant comparative analysis of a student's pronunciation, taking as a reference the speech of a native speaker, and observing the evolution of this through data stored in history.

For the development of our plugin, several technologies have been used to support the work done, such as Praat, Python and PostgreSQL.

2. Methodology

We start designing our work based on the following essential principles of intonation analysis:

1. We annotate the syllables of each speech act in a Praat textgrid (Boersma and Weenink [1]; we identify pitch values of all vowels in the syllables (voiced or voiced consonants are measured as well), using the Praat Script developed by Mateo Ruiz [9, 10], which extracts the absolute values in Hz, relativises them and draws the standardised melody graph;

2. we discriminate relevant frequency values between tonal segments from irrelevant values; according to Cantero Serena [2, 3], Font-Rotchés and Cantero Serena [6, 7], less than 10% difference between segments is considered imperceptible.

Once we have obtained the relevant data from the intonation analysis, we move on to the PAFe architecture.

Our project develops an extension to an existing desktop application for acoustic speech analysis: Praat. Therefore, we start from a developed architecture to which a new module (PAFe) is coupled (Figure 1) consisting of Praat scripts, Python code and a PostgreSQL database.
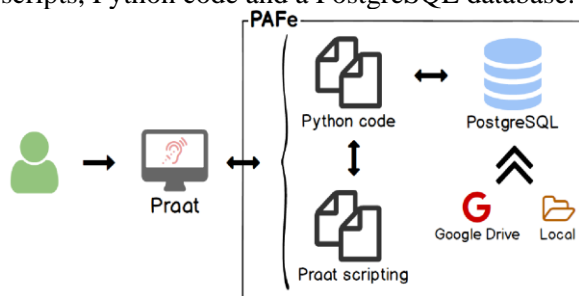

Figure 1: Overall architecture of PAFe

Praat, through its scripting, allows command line calls to other systems, as described by Dragos-PaulPop [5], thus making it possible to extend the application through the use of other languages and technologies, external to Praat. This new module (PAFe) communicates with the original system by employing new Praat scripts that are associated with the application's menu items (see Figure 2), from which these files are executed. Sometimes, the new module dispenses

with calls to Praat and generates information windows directly from Python code files. The intermediary between Praat and the data managed in the database is Python.

We employ natural language processing and audio processing techniques in our tool, taking as our main source the human voice recordings of native speakers and students. Praat allows us to extract quantitative information at the prosodic level from the audios.

Subsequently, the native/student comparative algorithms in terms of prosodic aspects that are presented and implemented by the tool can offer comparative information between two native/student audios to provide feedback in Spanish language learning. These algorithms are an original contribution implemented in the tool since there was no algorithmic proposal of this type for Spanish until now.

We have developed the PAFe Plugin following an iterative and incremental methodology based on agile technologies and scrum development methodology, based on the work of Schwaber and Sutherland [12]. Figure 2: Example of User Interface visualising the new functionalities added in Praat

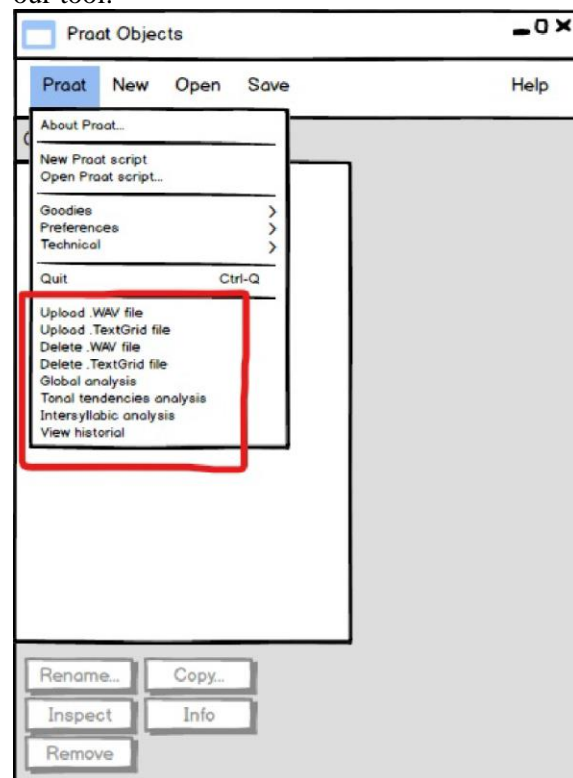In the following, we describe the development of our tool.

Figure 2: Example of User Interface visualising the new functionalities added in Praat

In the following, we describe the development of our tool.

3. Solution: PAFe Plugin

Our PAFe tool, in its final version, allows comparative analysis by providing similarity results and intonation graphs based on pitch values[2] and tonal tendency in each defined segment and, finally, visualisation of a student's progress over time. We highlight the following operations made possible by our plugin:

1. The application allows the creation of different profiles to facilitate the process of managing the data uploaded by users.

a) First of all, the teacher is registered.

b) A pupil is then assigned to the teacher previously registered. This step avoids confusion if there is more than one user of the same computer or laptop.

c) Finally, the profile of a native Spanish speaker is recorded to upload the data that will serve as a reference for the programme;

2. PAFe enables the management of WAV and TextGrid files[3]: our programme includes both storage and deletion of audio files and annotations;

3. It also allows for different types of acoustic analysis (global analysis, tonal tendency analysis and intersyllabic analysis): the algorithm that performs the global analysis consists of dividing the previously saved audios of learners and native speakers of Spanish into about 1000 intervals (discarding silences) to obtain very precise comparative values. However, this type of analysis does not provide feedback about possible deviations in tone but provides generic data on the percentage similarity of the native speaker's and learner's audio. As far as tonal tendency analysis is concerned, the programme works with .TextGrid annotations and the previously saved .WAV audio files. In this case, the utterances are divided by words and, to obtain the similarity locally, it is indicated whether the pitch of each word has been reproduced correctly or not and, in case it has not been reproduced correctly, the percentage of deviation is indicated; the percentage of pitch similarity and the average difference between two audios are also obtained. Finally, the intersyllabic acoustic analysis is a comparative analysis, syllable by syllable, of the similarity between the tone realisation of a learner and that of a native speaker; in this case, for each

syllable, the difference in pronunciation concerning the reference audio is indicated, as well as the percentage of similarity of tone and the average difference between two audios is obtained. According to the results obtained through this last type of analysis, both the similarity and the difference between the reference audio and the learners' audio are shown more accurately. Finally, we can see the evolution of our students' results through the option to view the history.

Finally, we show a flowchart (Figure 3) that provides information about the behaviour of our plugin, exposing the functionalities and their interrelation, as well as presenting the operators that interact with the application and their restrictions.
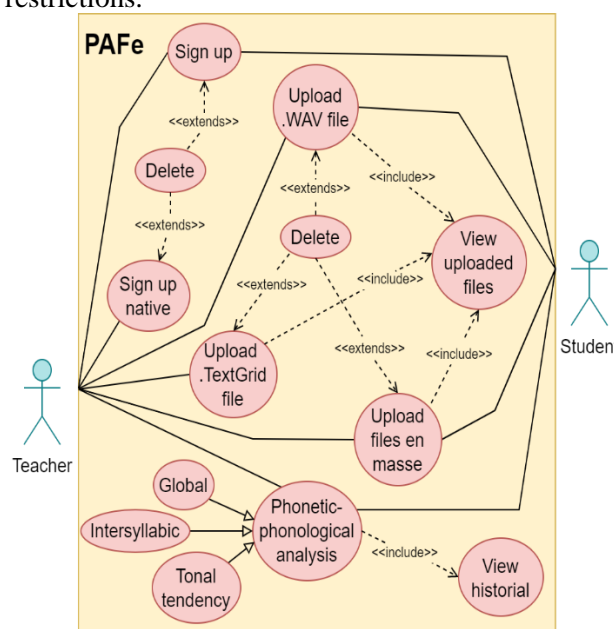

Figure 3: Use case diagram (PAFe functionalities and main actors)

4. Illustrative example of intersyllabic analysis

In this section, we show how one type of comparative analysis is carried out. To perform the intersyllabic analysis, it is necessary to fill in a form (Figure 4) with the data that characterise the audio of the learner we want to compare.

---

[2] Tone frequency in Hz

[3] File with tags segmenting associated audio
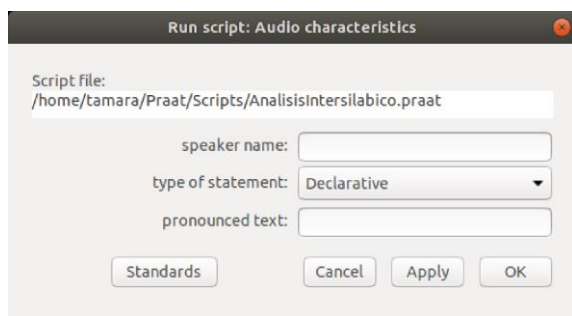
Figures 4: Form for conducting an intersyllabic analysis

The audios of that student that meet these properties are then filtered out and display a window with a drop-down menu for the selection of the audio to be analysed. Once the audio is selected, the corresponding. TextGrid file is selected in the same way.

Each type of analysis returns different results. For the intersyllabic analysis, we show a similarity result per syllable and the average percentage difference (Figure 5). Finally, we obtain a graph with the tonal differentiation curves in each syllable for each audio (Figure 6).
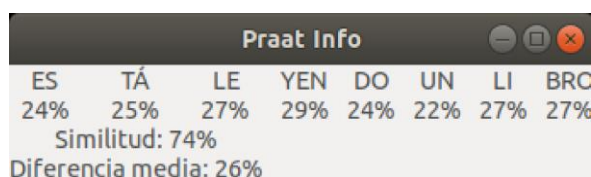
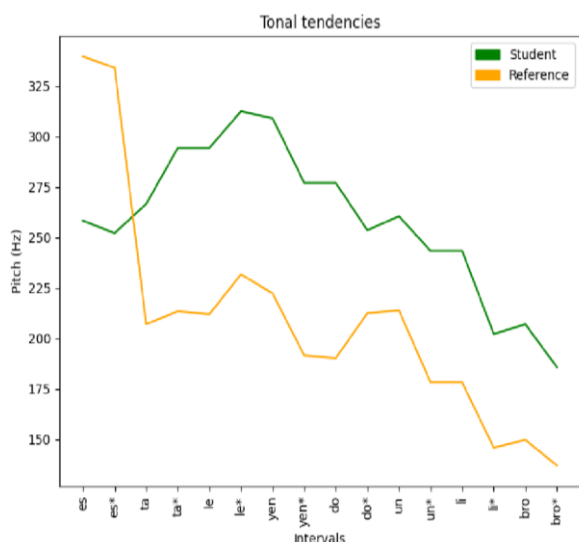**Figures 5:** Intersyllabic analysis information

Figure 6: Graph showing the tonal curves of each audio for each syllable (the X-axis represents the syllable division of an utterance and the Y-axis the pitch values).

5. Conclusions

In conclusion, we highlight the following key issues that we have addressed in this paper:

1. The PAFe tool allows different types of comparative-contrastive analysis of the intonation (global, tonal tendencies and intersyllabic) of EFL learners and native speakers of Spanish; Among them, we consider the intersyllabic as the most accurate since the results of tonal difference appear syllable by syllable and show the tonal deviations of the students, and the global as the most efficient in terms of response time since it does not require the uploading of TextGrids, and the segmentation is done in an automated way, as shown by the empirical data of the Couto Fernández [4] work.

2. This application has several functions; apart from performing the intonational analysis, it allows to store the audios, the . TextGrid files and the results of the analysis (the history) of each utterance according to the profile of the speaker (student or native speaker of Spanish).

3. PAFe has been developed to achieve the following didactic objectives: to facilitate the work of teachers with regard to the identification and correction of intonation deviations (we have carried out an empirical analysis with teachers of Spanish as a foreign language, where we measured the degree of satisfaction with PAFe, with positive results, as indicated in the work Couto Fernández [4]; to store the results of the analyses carried out for future improvement; to serve as a self-evaluation and self-correction tool for ELE students, given that the tool itself allows them to upload .WAV and . TextGrid files, run the analyses and obtain the results without constant help from teachers.

As a future line of research, we highlight the need to measure this degree of feedback to students empirically.

As far as we know, it is the only existing solution both under Praat and outside Praat that allows this type of analysis and offers feedback to the student in the Spanish language. We highlight that as feedback and self-evaluation, our tool offers the percentage of similarity and difference of pitch values so that the student can correct his pronunciation. Also, as future lines of work, we plan to improve the graphical environment of the plugin and open to the student, as an end user, the possibility of its use via the web.

86

References

[1] P. Boersma, D. Weenink, Praat: doing phonetics by computer, 2019. URL: http://www.praat.org/.

[2] F. J. Cantero Serena, Teoría y análisis de la entonación, volume 54, 2002.

[3] F. J. Cantero Serena, Análisis prosódico del habla: más allá de la melodía, Comunicación Social: Lingüística, Medios Masivos, Arte, Etnología, Folclor y otras ciencias afines 2 (2019) 485-498.

[4] T. Couto Fernández, Una herramienta de análisis del habla de audio para proporcionar retroalimentación automática a los estudiantes en la pronunciación en español. UDC. A Coruña.

[5] Dragos-Paul Pop, Adam Altar, Designing an MVC Model for Rapid Web Application Development, Procedia Engineering 69 (2014) 1172-1179. DOI: 10.1016/j.proeng.2014.03.106

[6] D. Font Rotchés, F. J. Cantero Serena, La melodía del habla: acento, ritmo y entonación, Eufonía: didáctica de la música (2008) 19-39.

[7] D. Font Rotchés, F. J. Cantero Serena, Melodic Analysis of Speech Method applied to Spanish and Catalan, Phonica 5 (2009) 33-47.

[8] H. Strik, K. Truong, F. Wet, C. Cucchiarini, Comparing different approaches for automatic pronunciation error detection, Speech Communication 51 (2009) 845–852. DOI: 10.1016/j.specom.2009.05.007

[9] M. Mateo Ruiz, Protocolo para la extracción de los datos tonales y curva estándar en análisis melódico del habla, Phonica 6 (2010) 49-90.

[10] M. Mateo Ruiz, Scripts en Praat para la extracción de datos tonales y curva estándar, Phonica 6 (2010) 91-111.

[11] P. Oplustil, G. Toledo, Uso de una herramienta didáctica para la práctica de la entonación en hablantes no nativos de español, Sintagma: Revista de lingüística 31 (2019) 37–50.

[12] K. Schwaber, J. Sutherland, La guía definitiva de scrum: Las reglas del juego, 2020.