Universidad de Valladolid

FACULTAD de FILOSOFÍA Y LETRAS
DEPARTAMENTO de FILOLOGÍA INGLESA

Máster en Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanzas de Idiomas
Especialidad: Lenguas Extranjeras (inglés)

TRABAJO DE FIN DE MÁSTER

# Overcoming segmental difficulties in English pronunciation in Spanish 3-ESO bilingual students through the use of SpeechAce

Lidia Avellón Mayor

Tutor: Miguel Vilalta Nieto

Valladolid

2021

## ABSTRACT

The aim of this study is to account for the main problems with which secondary school students are faced in terms of English pronunciation, and to look into the effectiveness of 'SpeechAce' as a tool for English phoneme pronunciation improvement. Specifically, the data obtained were evaluated considering the challenging English phonemes, the mispronunciation rate of each English phoneme prior to the use of 'SpeechAce', and the improvement rate shown upon its use. The results obtained show that both consonant and vowel sounds present similar mispronunciation rates before the use of 'SpeechAce' as well as similar improvement rates after its use. Thereby, 'SpeechAce' has proved to be a useful tool to overcome secondary school students' segmental difficulties in English pronunciation.



Key words: 'SpeechAce', automatic speech recognition, pronunciation, segmental features, secondary education.



## RESUMEN

El objetivo de este Trabajo de Fin de Máster consiste en explicar los principales problemas de pronunciación inglesa que presentan los alumnos de Educación Secundaria Obligatoria e investigar la efectividad de 'SpeechAce' como herramienta para solventarlos. En concreto, se han analizado los fonemas ingleses más complicados, la tasa de pronunciación incorrecta de cada fonema antes de utilizar 'SpeechAce', y la tasa de mejora después de su uso. En términos generales, los resultados del estudio demuestran que los sonidos consonánticos y vocálicos presentan tasas similares de pronunciación incorrecta antes del uso de 'SpeechAce', además de tasas de mejora semejantes después de su uso. Por lo tanto, 'SpeechAce' ha demostrado ser una herramienta de utilidad a la hora de solucionar los problemas relacionados con los elementos segmentales de la pronunciación inglesa.



Palabras clave: 'SpeechAce', reconocimiento automático del habla, pronunciación, elementos segmentales, Educación Secundaria Obligatoria.

# TABLE OF CONTENTS

# 1. INTRODUCTION

Throughout the history of English language teaching in Spain, English pronunciation has been overshadowed in classrooms. Thus, Martínez Asís (2004) pointed out five main reasons why this situation is common occurrence in Spain. Firstly, vocabulary and grammar have been considered the most important areas by language theorists. Secondly, pronunciation was not studied until the 19th century. Thirdly, the vast majority of English language textbooks hardly include sections on pronunciation. Fourthly, teachers find it extremely harsh to teach or learn English pronunciation. Lastly, at university, future English language teachers are not generally taught how to teach English pronunciation.

Understandably, it is no surprise that the teachers surveyed by Martínez Asís (2004) claimed the existence of an English-level-related imbalance between oral and written skills among Spanish students, the focus being put on the latter ones. That is to say, teachers tend to follow the grammar-translation method. Accordingly, pronunciation has not been a priority for most English teachers at Obligatory Secondary Education.

Thus, the aim of this dissertation is to account for the main problems with which secondary school students are faced in terms of English pronunciation, and to look into the effectiveness of 'SpeechAce' —an English pronunciation training software— as a tool for English pronunciation improvement. In the present study, suprasegmental features have been relegated to the background in favour of segmental features, deemed the starting point for English pronunciation enhancement. More specifically, 3-ESO —Compulsory Secondary Education— students from Valladolid (Castile and León, Spain) enrolled in the bilingual section will be analysed regarding the English sounds with which they struggle the most before using 'SpeechAce' and the improvement, if any, shown upon its use.

This dissertation is divided into 10 main sections starting with the introduction. In the second section, an overview on English pronunciation in the Spanish curriculum will be provided. In the third section, the reasons why pronunciation is of importance will be explained. In the fourth section, a detailed account of frequent English vowels' and consonants' mispronunciations committed by L1 speakers of Spanish will be given. In the fifth section, automatic speech recognition will be introduced as an advantageous tool for English pronunciation teaching. In the sixth section, the methodology of the study will be presented,

with information about the participants, 'SpeechAce', the stages of the study and data collection, together with the research questions. In the seventh section, the data obtained will be presented and analysed. In the eighth section, the conclusion of this dissertation will be drawn. In the nineth section, all the bibliographical sources referred to throughout this dissertation will be provided. Finally, in the tenth section, the annexes of this dissertation will be attached.

## 2. ENGLISH PRONUNCIATION IN THE SPANISH CURRICULUM

The origins of the lack of attention paid to English pronunciation can be explained taking the "Orden EDU/362/2015" —the normative including the relations between contents and assessment criteria in ESO in Castile and León— as a reference.

Even though this normative claims to advocate for a communicative approach in the foreign language classroom, this statement is a far cry from real teaching practice. Undoubtedly, the grammatical aspects of the English language continue to be the protagonists, as thoroughly detailed in pages 363 to 375 of the normative mentioned above.

As opposed to grammar, only extremely vague and inaccurate descriptions of the pronunciation features to be covered in the classroom are provided. Moreover, these descriptions are not detailed for the English language specifically, but for foreign languages in general. For this reason, as Table 1 shows, no phonemes are specified in the contents of neither grade. The most specific guidelines given are "pronunciation of contracted forms" and "pronunciation of verb tense endings" —corresponding to the contents of second grade. Thus, this suggests that pronunciation has become to second fiddle and, in some cases, been consigned to oblivion. Besides, in terms of pronunciation, the academic requirements in the normative are rather low, as the assessment criteria for first, second and third grades of ESO are identical, while the ones for fourth grade are alarmingly similar to the criteria of the rest of the grades of ESO (Table 1). Consequently, the same demands are placed on first-grade students as on fourth-grade students.

| Grade | Contents | Assessment criteria |
|---|---|---|
| First | Identification of some phonetic symbols with the pronunciation of some frequently used phonemes.<br>Production of basic patterns of word and sentence rhythm, intonation and stress. | To pronounce and intonate simple messages clearly and intelligibly, even if foreign accents are sometimes evident or sporadic non-impeding mispronunciations are committed, and interlocutors may require having messages repeated from time to time. |
| Second | Pronunciation of special hardship phonemes.<br>Pronunciation of contracted forms.<br>Pronunciation of verb tense endings.<br>Production of basic patterns of word and sentence rhythm, intonation and stress. | To pronounce and intonate simple messages clearly and intelligibly, even if foreign accents are sometimes evident or sporadic non-impeding mispronunciations are committed, and interlocutors may require having messages repeated from time to time. |
| Third | Pronunciation of special hardship phonemes.<br>Production of different patterns of word and sentence rhythm, intonation and stress. | To pronounce and intonate simple messages clearly and intelligibly, even if foreign accents are sometimes evident or sporadic non-impeding mispronunciations are committed, and interlocutors may require having messages repeated from time to time. |
| Fourth | Deepening in the use of phonetic symbols.<br>Pronunciation of special hardship phonemes.<br>Autonomous production of different patterns of word and sentence rhythm, intonation and stress. | To pronounce and intonate simple messages clearly and intelligibly, even though interlocutors may need to have messages containing infrequent words or structures repeated, in the articulation of which non-impeding mistakes may be committed. |

Table 1: Contents and criteria for English pronunciation in ESO in Castile and León according to the "Orden EDU/362/2015."

These loose guidelines have a negative impact on English pronunciation teaching, causing teachers to underemphasise it and to focus on the most thoroughly detailed contents, namely grammar and written skills.

## 3. WHY IS PRONUNCIATION IMPORTANT?

A skill of uttermost importance, correct pronunciation not only enables L2 English speakers to be understood in an easier way, but also to understand the English language better and feel a massive sense of integration when speaking to L1 English speakers.

People can understand L2 English speakers having good pronunciation even if they make mistakes in other features of language. Nonetheless, they are not able to understand those presenting unintelligible pronunciation even if they have the knowledge of a wide variety of vocabulary and master English grammar (Yates & Zielinski, 2009). Thus, English-speaking listeners judge a speaker's English skills on the basis of his/her own pronunciation. Indeed, poor pronunciation is extremely difficult to listen to, with attention spans declining within seconds. Consequently, mispronunciation is liable to cause misunderstandings, as well as a breakdown in communication (Gilakjani & Sabouri, 2016). Moreover, in the view of Gilakjani (2012), if a speaker presents an acceptable pronunciation, listeners judge his/her overall language ability in a more effective way, even tolerating grammatical inaccuracy.

According to Gebhard (1996), as cited in Khaghaninejad & Maleki (2015), "there is a close link between pronunciation and listening comprehension since the ability to perceive and produce speech requires the knowledge of sounds, intonation, and stress pattern and how speech is organized" (pg. 2). Hence, teaching sounds can be useful for students to distinguish words from a listening comprehension. Furthermore, working on supra-segmental features can help students to understand longer utterances (Khaghaninejad & Maleki, 2015).

## 4. ENGLISH PHONEME PRONUNCIATION MISTAKES IN L1 SPEAKERS OF SPANISH

Spanish and English have a similar alphabet, but when the phonologies of the two languages are compared, numerous differences can be observed, namely in terms of phonemes. These differences influence the pronunciation of L1 speakers of Spanish learning English. Thus, Villarín & Helíodora (1998) distinguish three basic types of phoneme pronunciation mistakes.

The first type —hereafter, ‘Type 1’— involves mistakes consisting in the substitution of an English phoneme which does not exist in the Spanish phonology for the most approximate phoneme in Spanish. It is a matter of fact that, before L1 speakers of Spanish internalise the English phonological system, they tend to transfer Spanish pronunciation to their English speech (Gorman & Kester, 2001).

The second type —hereafter, ‘Type 2’— involves mistakes caused by distribution differences among phonemes shared by English and Spanish. That is to say, those mispronunciations produced by the incorrect placing of the mouth organs.

Lastly, the third type —hereafter, ‘Type 3’— involves mistakes rooted in the fact that, as opposed to Spanish, English is not a phonetic language. As a result, L1 speakers of Spanish may adopt Spanish grapho-phonemic correspondences when speaking in English.

In the following subsections, a brief comparative between the Spanish and the English vowel and consonant systems will be provided. Then, the most common vowel and consonant mispronunciations committed by L1 speakers of Spanish learning English will be classified attending to the types of mistakes described above.

## 4.1. Vowels

### 4.1.1. Simple vowels

With regards to the vowel system, the English language is extremely different from the Spanish language. As Figure 1 shows, the English vowel system —in black— presents twelve simple vowel phonemes —/ɑ:/, /ʌ/, /æ/, /e/, /ə/, /ɜ:/, /i:/, /ɪ/, /ɒ:/, /ɔ:/, /ʊ/, /u:/—, whereas the Spanish one —in grey— only five —/i/, /e/, /a/, /o/, /u/. Thus, in addition to lacking distinction between short and long vowels, the Spanish vowel system lacks the weak phoneme schwa (/ə/), which is present in almost every word in English of two or more syllables.

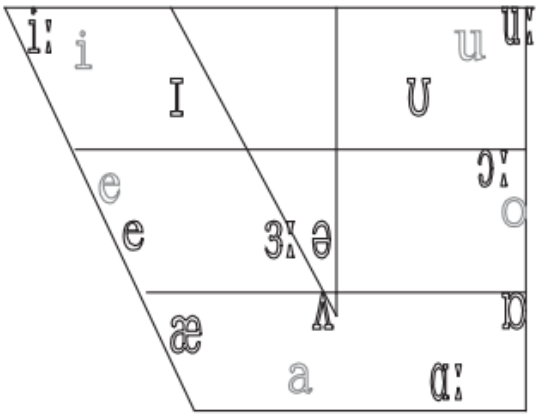


Figure 1: Diagram of English and Spanish vowels (Del Puerto & Lacabex, 2008).

As a result of the great array of simple vowel phonemes which the English language presents, L1 speakers of Spanish usually simplify them by pronouncing English words using the most similar vowel sounds available in the Spanish vowel system. This way, as presented in Table 2, Finch & Lira (1982) state that Spanish-speaking learners of English tend to pronounce the English /iː/ and /ɪ/ like the Spanish /i/ (Examples (1) and (2)); the English /e/ like the Spanish /e/ (Example (3)); the English /ɜː/ like the Spanish /e/; (Example (4)); the English /ɑː/, /ʌ/ and /æ/ like the Spanish /a/ (Examples (5), (6) and (7)); the English /ʊ/ and /uː/ like the Spanish /u/ (Examples (8) and (9)); the English /ɒ/ like the Spanish /o/ or /a/ (Examples (10) and (11)), and the English /ɔː/ like the Spanish /o/ (Example (12)). That is to say, 'Type 1' mistakes are common occurrence when it comes to the set of vowel sounds which have just been commented on.

| English vowels substituted | Spanish vowels |
|---|---|
| /ɑː/ | /a/ |
| /ʌ/ | /a/ |
| /æ/ | /a/ |
| /e/ | /e/ |
| /ɜː/ | /e/ |
| /iː/ | /i/ |
| /ɪ/ | /i/ |
| /ɒː/ | /a/ /o/ |
| /ɔː/ | /o/ |
| /ʊ/ | /u/ |
| /uː/ | /u/ |

Table 2: English vowels substituted by their most approximate vowels in Spanish.

(1) Green (/ˈgriːn/) → */ˈgrin/

(2) Sit (/ˈsɪt/) → */ˈsit/

(3) Men (/ˈmen/) → */ˈmen/

(4) Bird (/ˈbɜːd/) → */ˈberd/

(5) Car (/ˈcɑːr/) → */ˈcar/

(6) Sun (/ˈsʌn/) → */ˈsan/

(7) Bat (/ˈbæt/) → */ˈbat/

(8) Put (/ˈpʊt/) → */ˈput/

(9) Moon (/ˈmuːn/) → */ˈmun/

(10) Lot (/lɒt/) → */ˈlot/

(11) Quality (/ˈkwɒlɪti/) → */ˈkualiti/

(12) Door (/ˈdɔː/ → */ˈdor/

Furthermore, L1 speakers of Spanish are normally influenced by the spelling of English words when pronouncing —the false agreement between English spelling and Spanish grapho-phonemic correspondences. This is clearly the case of every English vowel sound —as Table 3 shows—, the most notorious being /ə/, /ɜː/, /ɔː/ and /ɒ/.

Firstly, as the English /ə/ can correspond to either the graphemes *a, e, o,* or *u*, Spanish speakers tend to pronounce the English /ə/ like the Spanish /a/, /e/, /o/ or /u/, respectively (Examples (13), (14), (15) and (16)).

Secondly, the English /ɜː/ can correspond to either the graphemes *u* or *o,* when in a stressed syllable. Hence, Spanish speakers usually pronounce the English /ɜː/ like the Spanish /u/ or /o/, respectively (Examples (17) and (18)). Note that /ɜː/ can also correspond to the grapheme *i*, albeit, as discussed above, Spanish speakers do not tend to pronounce it like the Spanish /i/, but /e/ (Example (4)).

Thirdly, the English /ɔː/ can correspond to either the graphemes *o* or *a*, when in a stressed syllable. Thus, they are usually pronounced by Spanish speakers like the Spanish /o/ or /a/, respectively (Examples (12) and (19)).

Lastly, the English /ɒ/ can correspond to either the graphemes *o* or *a* when in a stressed syllable. Thereby, as in the case of /ɔː/, Spanish speakers tend to pronounce the English /ɒ/ like the Spanish /o/ or /a/, respectively (Examples (10) and (11)).

Consequently, ‘Type 3’ mistakes are also frequent among L1 speakers of Spanish when articulating vowel sounds in English.

| English vowels | Associated graphemes | Spanish vowels produced |
|---|---|---|
| /ɑː/ | a | /a/ |
| /ʌ/ | o, u | /a/ |
| /æ/ | a | /a/ |
| /e/ | e | /e/ |
| /ɜː/ | i, o, u | /e/ |
| /ə/ | a, e, o, u | /a/, /e/, /o/, /u/ |
| /iː/ | i, ee | /i/ |
| /ɪ/ | i | /i/ |
| /ɒː/ | a, o | /a/, /o/ |
| /ɔː/ | o | /o/ |
| /ʊ/ | u | /u/ |
| /uː/ | oo, u | /u/ |

Table 3: English vowel grapheme association and Spanish vowels used.

(13) Attend (/əˈtend/) → */aˈtend/

(14) Centre (/ˈsentə/) → */ˈsenter/

(15) Doctor (/ˈdɒktə/) → */ˈdoktor/

(16) Future (/ˈfjuːtʃə/) → */ˈfiutur/

(17) Urban (/ˈɜːbən/) → */ˈurban/

(18) Word (/ˈwɜːd/) → */ˈword/

(19) War (/ˈwɔː/) → */ˈgwar/

### 4.1.2. Diphthongs

As for diphthongs, the reverse situation than that for simple vowels is found because English has fewer diphthongs than Spanish. Thus, the former has only 8 diphthongs —/ɪə/, /eə/, /ʊə/, /əʊ/, /aʊ/, /aɪ/, /eɪ/, and /ɔɪ/— whereas the latter has fourteen diphthongs —/ai̯/, /ei̯/, /oi̯/, /au̯/, /eu̯/, /ou̯/, /wa/, /we/, /wo/, /wi/, /ja/, /je/, /jo/, and /ju/ (Gómez González & Sánchez Roura, 2016).

On the one hand, English diphthongs can be classified into closing —/əʊ/, /aʊ/, /eɪ/, and /ɔɪ/— and centring — /ɪə/, /ʊə/, and /eə/— diphthongs (Figure 2), while, on the other hand, Spanish diphthongs can be classified into falling —/ai̯/, /ei̯/, /oi̯/, /au̯/, /eu̯/, and /ou̯/— and raising—/u̯a/, /u̯e/, /u̯o/, /i̯a/, /i̯e/, and /i̯o/— diphthongs (Figure 3).

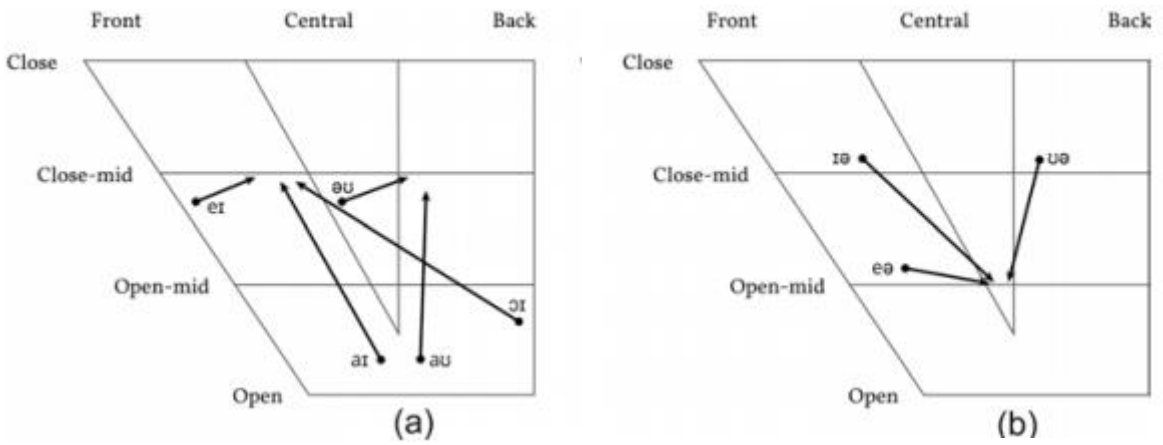


Figure 2: Diagrams of English closing (a) and centring (b) diphthongs (Gómez González & Sánchez Roura, 2016).

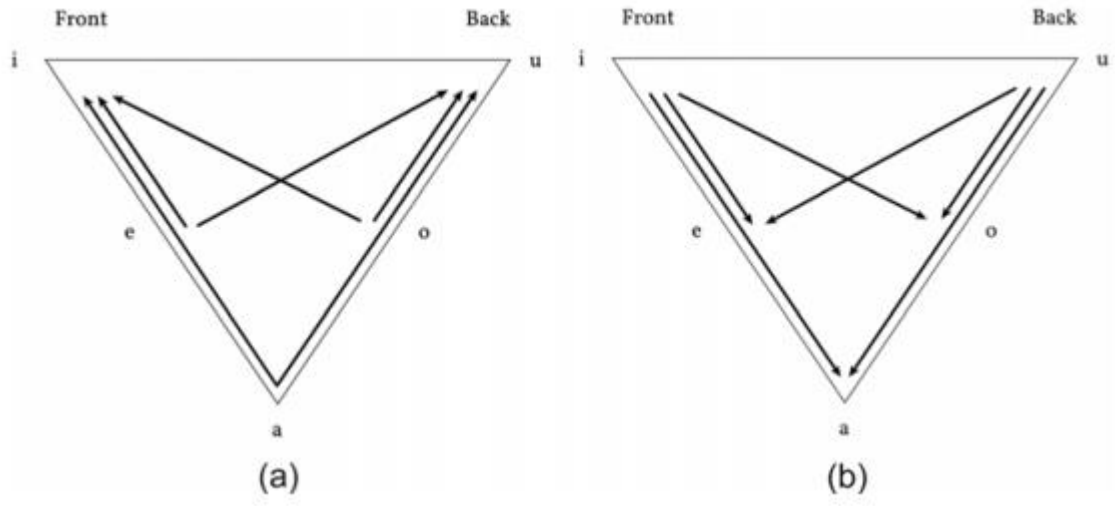


Figure 3: Diagrams of Spanish falling (a) and raising (b) diphthongs (Gómez González & Sánchez Roura, 2016).

Concerning the three English /ɪ/-diphthongs —/eɪ/, /aɪ/, and /ɔɪ/—, it can be stated that they approximate in quality to the Spanish ones —/ei/, /ai/, and /oi/— (Figure 4). English /aɪ/ and /aʊ/ begin in an area which is extremely similar to that of Spanish /a/ and the starting points of /a̯i̯/ and /a̯u̯/. English /eɪ/ and /ɔɪ/, nonetheless, start in a lower area than their Spanish equivalents /e̯i̯/ and /oi/. However, the key difference between the two languages lies in the second element: in Spanish it has an obvious /i/ quality, whereas in English the /ɪ/ sound is more relaxed and open, between Spanish /i/ and /e/ (Gómez González & Sánchez Roura, 2016). Yet,

no misunderstandings will occur if a Spanish /i/ style of vowel is articulated as the end point of these three diphthongs (Monroy Casas, 1980, 2012; Gallardo del Puerto, 2005; Estebas Vilaplana, 2009). Thereby, L1 speakers of Spanish tend to pronounce the English /eɪ/, /aɪ/, and /ɔɪ/ like the Spanish /ei/, /ai/, and /oi/ respectively (Examples (20), (21), and (22)), committing 'Type 1' mistakes.

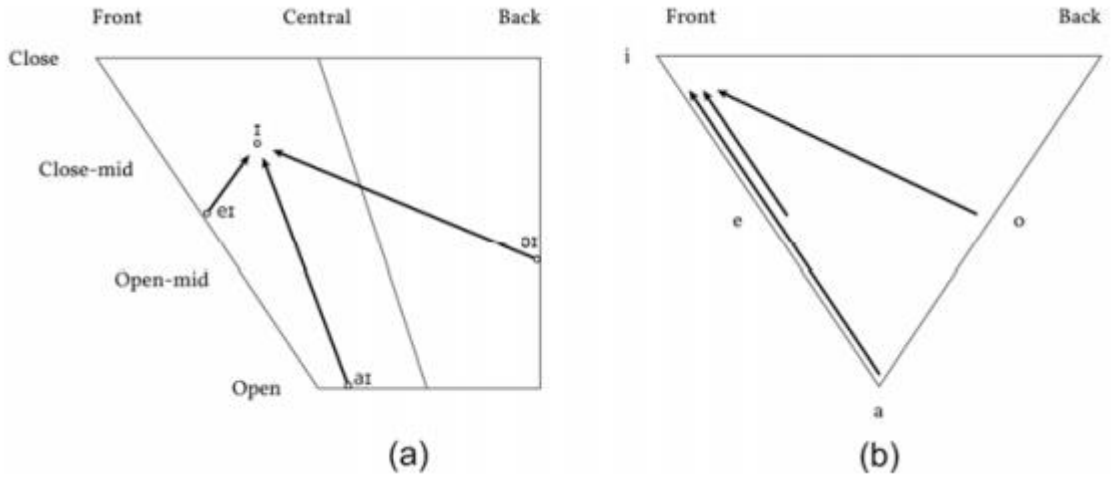

Figure 4: [ɪ/i] diphthongs in English (a) and Spanish (b) (Gómez González & Sánchez Roura, 2016).

(20) May (/ˈmeɪ/) → */ˈmei/

(21) Lie (/ˈlaɪ/) → */ˈlai/

(22) Boy (/ˈbɔɪ/) → */ˈboi/

In terms of the English /əʊ/ and /aʊ/, several issues should be considered. On the one hand, the pronunciation of /əʊ/ may be challenging for Spanish learners, since /ə/ does not form part of the Spanish vowel system. Thus, L1 speakers of Spanish tend to pronounce this diphthong with an initial /o/ followed by the Spanish /u/, thereby approximating the pronunciation of the Spanish diphthong /o̯u̯/ (Figure 5). Accordingly, L1 speakers of Spanish should try to start this diphthong with a /ə/-sound. The closest Spanish diphthong to the English /əʊ/ is /e̯u̯/ (Figure 5), but the first element of the English diphthong is more central and the second more relaxed (Gómez González & Sánchez Roura, 2016). On the other hand, the English /aʊ/ should have no problem for L1 speakers of Spanish, on the grounds of its similarity to the Spanish diphthong /a̯u̯/ (Figure 5), the only difference being that the starting point is more retracted and the endpoint less relaxed in Spanish (Gómez González & Sánchez Roura, 2016). Nonetheless, the pronunciation of a Spanish /u/ as the endpoint of these two diphthongs should not cause intelligibility problems (Monroy Casas, 1980, 2012; Gallardo del Puerto, 2005;

Estebas Vilaplana, 2009). Consequently, it is common occurrence for Spanish learners to pronounce the English /aʊ/ like the Spanish /au̯/. Thus, it can be said that L1 speakers of Spanish commit 'Type 1' mistakes when it comes to pronouncing the English /əʊ/ and /aʊ/, as they tend to produce them like the Spanish /ou̯/ and /au̯/, respectively (Examples (23) and (24)).

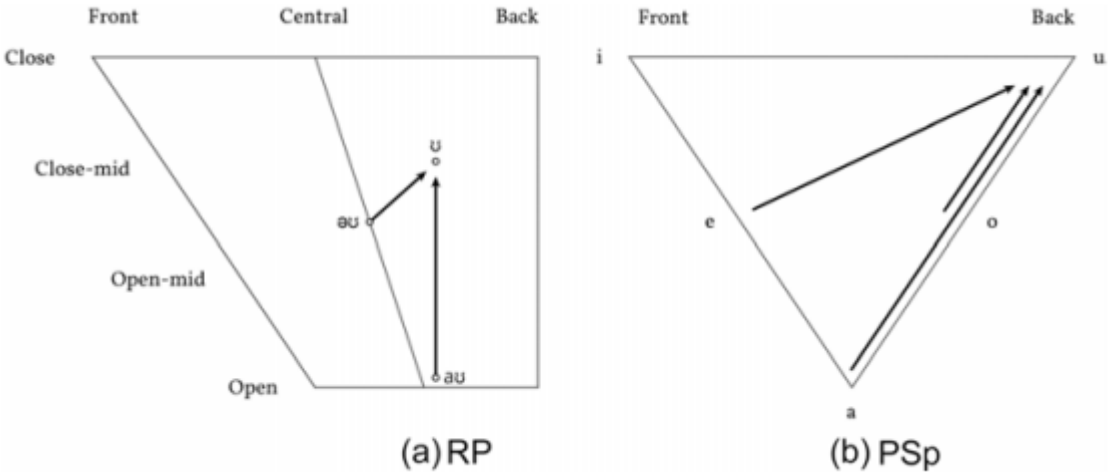


Figure 5: [ʊ/u] diphthongs in English (a) and Spanish (b) (Gómez González & Sánchez Roura, 2016).

(23) Low (/ˈləʊ/) → */ˈlou̯/

(24) Mouse (/ˈmaʊs/) → */ˈmau̯s/

As for centring diphthongs, the Spanish phonological repertoire does not present this type of diphthongs. Nevertheless, approximate realisations of English /ɪə/, /eə/ and /ʊə/ (Figure 2) may be found in the Spanish hiatus sequences /ía/, /ea/ and /úa/ (Gómez González & Sánchez Roura, 2016), the starting point being closer, and the endpoint, more prominent and more open and advanced in Spanish. Consequently, the glide is longer in Spanish than in English (Finch & Lira, 1982; Gallardo del Puerto, 2005). Therefore, L1 speakers of Spanish have a tendency to pronounce the English diphthongs /ɪə/, /eə/ and /ʊə/ like the Spanish hiatus /ía/, /ea/ and /úa/, respectively (Examples (25), (26), and (27)), thereby making 'Type 1' mistakes.

(25) Clear (/ˈklɪə/) → */ˈklíaɾ/

(26) Care (/ˈkeə/) → */ˈkear/

(27) Sure (/ˈʃʊə/) → */ˈsúar/

To wrap up, Table 5 shows a summary of 'Type 1' mistakes committed by Spanish speakers when trying to pronounce English diphthongs —i.e., the Spanish diphthongs and hiatuses which are usually transferred when speaking in English.

| English diphthong | Spanish sequence produced |
|---|---|
| /aɪ/ | /ai/ |
| /eɪ/ | /ei/ |
| /ɔɪ/ | /oi/ |
| /əʊ/ | /ou̯/ |
| /aʊ/ | /au̯/ |
| /ɪə/ | /ía/ |
| /eə/ | /ea/ |
| /ʊə/ | /úa/ |

Table 5: 'Type 1' mistakes by Spanish speakers when producing English diphthongs.

## 4.2. Consonants

Spanish and English consonant systems, in contrast to their respective vowel systems, do present some similarities, as some consonant sounds are shared by both languages. Nonetheless, English has a greater consonantal variability than Spanish: 24 consonant phonemes, compared to 19 in Spanish (Table 6), not including allophones or dialectical variations in either language (Gómez González & Sánchez Roura, 2016).

On the one hand, the English consonantal system consists of 6 plosives —/p/, /t/, /k/, /b/, /d/, and /ɡ/—, 9 fricatives —/f/, /v/, /θ/, /ð/, /s/, /z/, /ʃ/, /ʒ/, and /h/— , 2 affricates —/tʃ/ and /dʒ/—, 3 nasals —/m/, /n/, and /ŋ/—, 1 lateral —/l/—, and 3 approximants —/r/, /w/, and /j/.

On the other hand, the Spanish consonantal system consists of 6 plosives —/p/, /t/, /k/, /b/, /d/, and /g/—, 5 fricatives —/f/, /θ/, /s/, /j/, and /x/—, 1 affricate —/tʃ/—, 3 nasals —/m/, /n/, and /ɲ/—, 2 laterals —/l/ and /ʎ/—, and 2 vibrants —/ɾ/ and /r/— (Gómez González & Sánchez Roura, 2016).

This shows that only 11 English consonant phonemes have an equivalent in Spanish —/p/, /b/, /k/, /g/, /f/, /θ/, /s/, /tʃ/, /m/, /n/, and /l/. The other thirteen English consonant sounds —/t/, /d/, /dʒ/, /v/, /ð/, /z/, /ʃ/, /ʒ/, /h/, /ŋ/, /r/, /j/, and /w/— have no exact equivalent phonemes in Spanish. Accordingly, special attention should be paid to these sounds, on the grounds that they can project possible phonemic problems for English-as-a-foreign-language learners, albeit

certain may still be found in the phonetic or allophonic inventories of some dialects of Spanish (Gómez González & Sánchez Roura, 2016).

| | | Bilabial | Dental | Labiodental | Interdental | Alveolar | Palato-alveolar | Palatal | Velar | Labiovelar | Glottal |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Occlusive | S | /p/, /b/ | /t/, /d/ | | | | | | /k/, /g/ | | |
| | E | /p/, /b/ | | | | /t/, /d/ | | | /k/, /g/ | | |
| Fricative | S | | | /f/ | /θ/ | /s/ | | /ʝ/ | /x/ | | |
| | E | | | /f/, /v/ | /θ/, /ð/ | /s/, /z/ | /ʃ/, /ʒ/ | | | | /h/ |
| Affricate | S | | | | | | /tʃ/ | | | | |
| | E | | | | | | /tʃ/, /dʒ/ | | | | |
| Nasal | S | /m/ | | | | /n/ | | /ŋ/ | | | |
| | E | /m/ | | | | /n/ | | | /ŋ/ | | |
| Lateral | S | | | | | /l/ | | /ʎ/ | | | |
| | E | | | | | /l/ | | | | | |
| Approximant | S | | | | | | | | | | |
| | E | | | | | /r/ | | /j/ | | /w/ | |
| Vibrant | S | | | | | /ɾ/, /r/ | | | | | |
| | E | | | | | | | | | | |

Table 6: Comparison between Spanish (S) and English (E) consonants (González, 2016).

Thereby, Brunori (2016), Rolo Cruz (2019), and Coe (1987) remarked several issues with consonant sounds with which L1 Spanish speakers are faced.

With regards to occlusive sounds, L1 speakers of Spanish tend to struggle with /p/ and /b/ (Figures 6 and 7), /t/ and /d/ (Figures 8 and 9), and /k/ and /g/ (Figures 10 and 11), primarily because they are aspirated before a stressed vowel in English. Furthermore, both /t/ and /d/ are alveolar in English, whereas they are dentoalveolar and less tense in Spanish. This results in L1

speakers of Spanish pronouncing non-aspired /p/, /b/, /t/, /d/, /k/, and /g/, and dentoalveolar along with less tense /t/ and /d/ when speaking in English. That is to say, they make 'Type 2' mistakes.

Figure 6: English /p/ and /b/ (Judson, 2012).

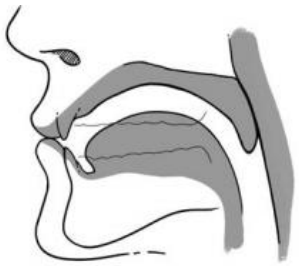

Figure 7: Spanish /p/ and /b/ (Clegg & Fails, 2017).

Figure 8: English /t/ and /d/ (Judson, 2012).

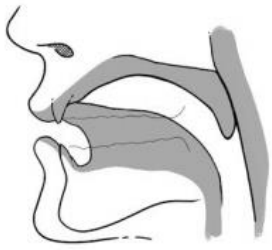

Figure 9: Spanish /t/ and /d/ (Clegg & Fails, 2017).

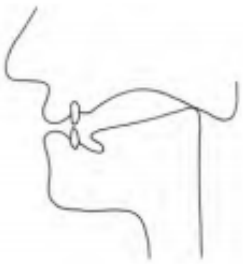

Figure 10: English /k/ and /g/ (Judson, 2012).

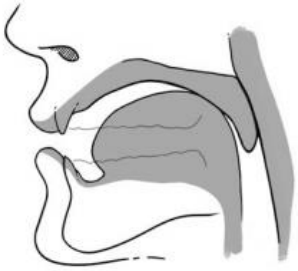

Figure 11: Spanish /k/ and /g/ (Clegg & Fails, 2017).

As for affricate sounds, L1 speakers of Spanish often pronounce the English /ʤ/ —unvoiced palato-alveolar affricate— (Figures 12 and 13), which is not present in the Spanish phonological system, as the Spanish /ʎ/ —voiced palatal lateral— (Figure 14) or as the Spanish /j/ —voiced palatal fricative— (Figure 15) when the former is located in initial or mid-word position (Examples (28) and (29)), and as the Spanish /tʃ/ — unvoiced palato-alveolar affricate– (Figure 16) when located in end-of-word position (Example (30)), being all of these the most approximate to the English sound available in Spanish. Moreover, they are influenced by the spelling too, as the English /ʤ/ sometimes corresponds to the letter "y", thus adopting Spanish

grapho-phonemic correspondences when pronouncing in this English consonant sound. In other words, L1 speakers of Spanish commit 'Type 1' and 'Type 3' mistakes when it comes to the English /ʤ/.

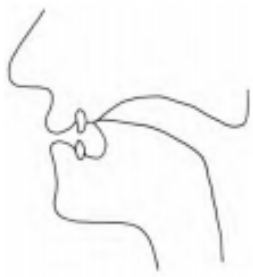

Figure 12: English /ʤ/ (1) (Judson, 2012).

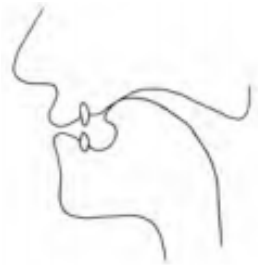

Figure 13: English /ʤ/ (2) (Judson, 2012).

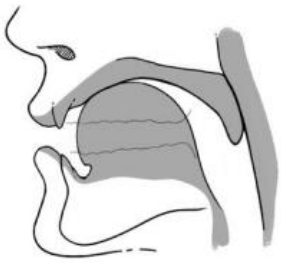

Figure 14: Spanish /ʎ/ (Clegg & Fails, 2017).

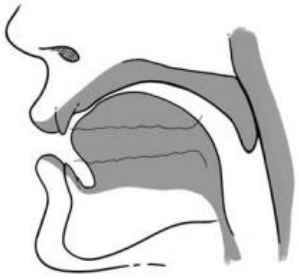

Figure 15: Spanish /j/ (Clegg & Fails, 2017).

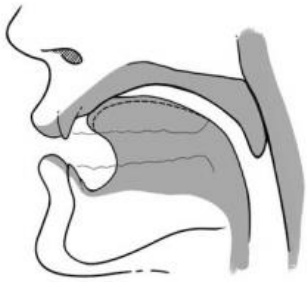

Figure 16: Spanish /tʃ/ (Clegg & Fails, 2017).

(28) Job (/ˈdʒɒb/) → */ˈʎob/ or */ˈjob/

(29) Adjust (/əˈdʒʌst/) → */adˈʎast/ or */adˈjast/

(30) Garage (/ˈgærɪdʒ/) → */ˈgaritʃ/

As for nasal sounds, L1 speakers of Spanish tend to pronounce the English /ŋ/ —voiced velar nasal— (Figure 17) using the Spanish /ŋ/ —voiced palatal nasal— (Figure 18) followed by /x/ —unvoiced velar fricative— (Figure 19), as they are influenced by the English spelling,

to which they apply Spanish grapho-phonemic correspondences (Example (31)). That is to say, L1 speakers of Spanish commit 'Type 3' mistakes when trying to pronounce the English /ŋ/.

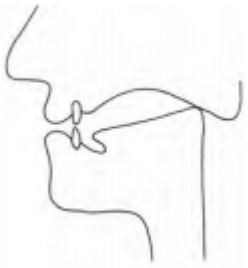

Figure 17: English /ŋ/ (Judson, 2012).

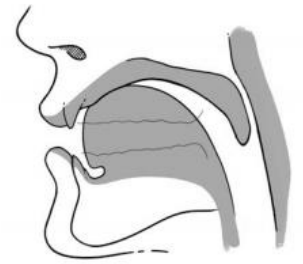

Figure 18: Spanish /ŋ/ (Clegg & Fails, 2017).

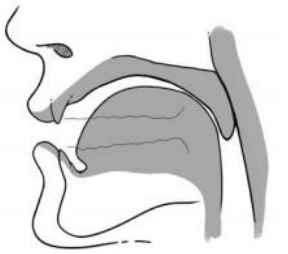

Figure 19: Spanish /x/ (Clegg & Fails, 2017).

(31) Going (/ˈgɔɪŋ/) → */ˈgoiŋx/

Concerning fricative sounds, L1 speakers of Spanish tend to struggle with /v/, /ð/, /z/, /ʃ/, /ʒ/, and /h/.

First, they usually pronounce the English /v/ —voiced labiodental fricative— (Figure 20) when located in initial or mid-word position as the Spanish /b/ —voiced bilabial occlusive— (Figure 7), since the latter sound is represented with the letter "v" in Spanish (Example (32)), thus adopting Spanish grapho-phonemic correspondences —'Type 3' mistakes. Nevertheless, when the English /v/ is located in end-of-word position, L1 speakers of Spanish usually pronounce it as the Spanish /f/ —unvoiced labiodental fricative— (Figure 21) (Example (33)), being the most approximate sound available in the Spanish phonological repertoire — 'Type 1' mistakes.

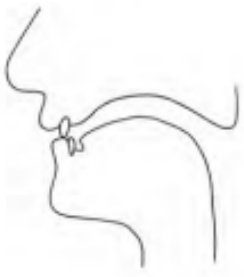

Figure 20: English /v/ (Judson, 2012).

Second, /ð/ —voiced interdental fricative— (Figure 22) is not present in the Spanish phonological system —solely as an allophone. Thereby, students tend to pronounce it as the Spanish /d/ (Figure 9) (Example (34)), which is the most approximate consonant sound available in Spanish of which L1 speakers of Spanish are aware. Therefore, they commit 'Type 1' mistakes.

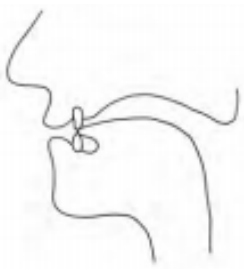

Figure 22: English /ð/ (Judson, 2012).

Third, /z/ —voiced alveolar fricative— (Figure 23) is also not present in the Spanish phonological system. When producing this sound, the vocal cords must vibrate. Nevertheless, the tendency among L1 speakers of Spanish is to produce it without making them vibrate, resulting in the production of the Spanish /s/ —unvoiced alveolar fricative— (Figure 24) (Example (35)). Moreover, it is common occurrence for Spanish students to pronounce the English /z/ —voiced alveolar fricative– as the Spanish /θ/ —unvoiced interdental fricative— (Figure 25) (Example (36)) or as the Spanish /s/ (Example (37)) on account of the influence of Spanish grapho-phonemic correspondences —i.e., the letter "z" corresponds to /θ/, and the letter "s" to /s/ in Spanish. That is to say, L1 speakers of Spanish are liable to commit 'Type 1' and 'Type 3' mistakes when it comes to the English /z/.

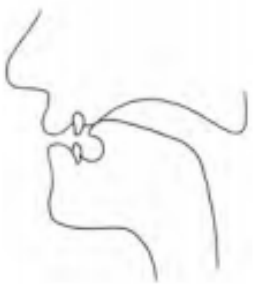

Figure 23: English /z/ (Judson, 2012).

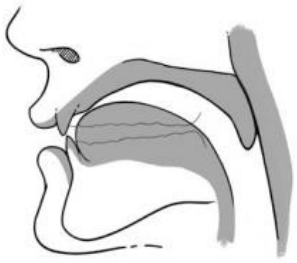

Figure 24: Spanish /s/ (Clegg & Fails, 2017).

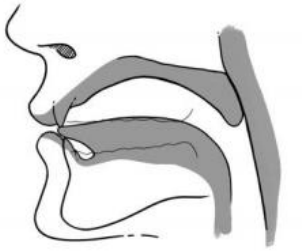

Figure 25: Spanish /θ/ (Clegg & Fails, 2017).

Fourth, neither /ʃ/ —unvoiced palate-alveolar fricative— nor /ʒ/ —voiced palate-alveolar fricative— (Figure 26) being present in the Spanish phonological system, L1 speakers of Spanish tend to pronounce them using the Spanish /s/ —unvoiced alveolar fricative— (Figure 24) (Examples (38) and (39)) when speaking in English, as it is the most approximate consonant available in Spanish and they usually correspond to the letter "s" —i.e., Spanish students adopt Spanish grapho-phonemic correspondences in these cases. Thus, L1 speakers of Spanish tend to commit 'Type 1' and 'Type 3' mistakes concerning the English /ʃ/ and /ʒ/.

Figure 26: English /ʃ/ and /ʒ/ (Judson, 2012).

Finally, L1 speakers of Spanish tend to pronounce the English /h/ —unvoiced glottal fricative— (Figure 27) as the Spanish /x/ —unvoiced velar fricative— (Figure 19) (Example (40)), the former not included in the Spanish phonological system and the latter being the most

approximate consonant sound available in Spanish. Hence, 'Type 1' mistakes are made by L1 speakers of Spanish in terms of the production of the English /h/.

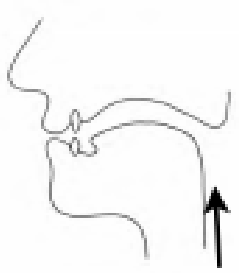

Figure 27: English /h/ (Judson, 2012).

(32) Vital (/ˈvaɪtəl/) → */ˈbaital/
(33) Live (/ˈlɪv/) → */ˈlif/
(34) The (/ˈðə/) → */ˈde/
(35) Lazy (/ˈleɪzi/) → */ˈleisi/
(36) Zoo (/ˈzuː/) → */ˈθu/
(37) Nose (/ˈnəʊz/) → */ˈnous/
(38) She (/ˈʃiː/) → */ˈsi/
(39) Vision (/ˈvɪʒən/) → */ˈbision/
(40) He (/ˈhiː/) → */ˈxi/

In terms of approximant sounds, L1 speakers of Spanish normally struggle to pronounce the English /r/, /j/, and /w/.

In first place, L1 speakers of Spanish have a tendency to pronounce the English /r/ —voiced alveolar approximant— (Figure 28) as the Spanish /r/ —voiced alveolar vibrant— (Figure 29) in initial position (Example (41)) or in mid-word position if it corresponds to a double "r" spelling (Example (42)), or as the Spanish /ɾ/ —voiced alveolar vibrant— (Figure 30) in mid-word position corresponding to a simple "r" spelling (Example (43)) or in end-of-word position (Example (44)). This shows that they are highly influenced by Spanish grapho-phonemic correspondences along with the fact that L1 speakers of Spanish commit 'Type 2' and 'Type 3' mistakes concerning the English /r/.

Figure 28: English /r/ (Judson, 2012).

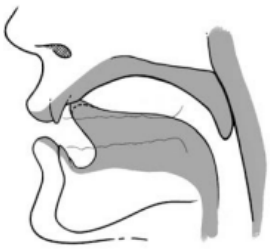

Figure 29: Spanish /r/ (Clegg & Fails, 2017).

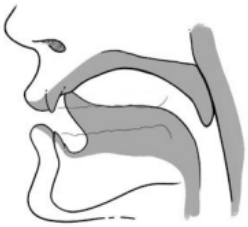

Figure 30: Spanish /ɾ/ (Clegg & Fails, 2017).

In second place, L1 speakers of Spanish tend to pronounce the English /j/ —voiced palatal approximant— (Figure 31) as the Spanish /j/ —voiced palatal fricative— (Figure 15) or /ʎ/ — voiced palatal lateral— (Figure 14) (Example (45)), being these sounds the most approximate ones to the English sound and learners being influenced by English spelling, to which they apply Spanish grapho-phonemic correspondences. Thereby, they commit 'Type 2' and 'Type 3' mistakes.

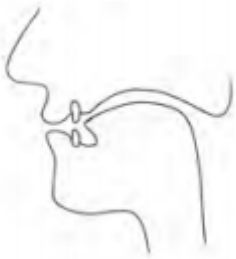

Figure 31: English /j/ (Judson, 2012).

In last place, L1 speakers of Spanish usually pronounce the English /w/ —voiced labiovelar approximant— (Figure 32) as the Spanish /g/ —voiced velar occlusive— (Figure 11)

followed by /u/ (Example (46)), being the most approximate sound available in the Spanish phonological repertoire. Thus, they commit 'Type 1' mistakes.

Figure 32: English /w/ (Judson, 2012).

(41) Rice (/ˈraɪs/) → */ˈrais/

(42) Sorry (/ˈsɒri/) → */ˈsori/

(43) Camera (/ˈkæmərə/) → */ˈkameɾa/

(44) Ear (/ˈɪə/) → */ˈíaɾ/

(45) You (/ˈjuː/) → */ˈju/ or */ˈʎu/

(46) Water (/ˈwɔːtə/) → */ˈgwateɾ/

| Type of consonant | Consonant | Type(s) of mistake |
|---|---|---|
| Occlusive | /p/ | 2 |
| | /b/ | 2 |
| | /t/ | 2 |
| | /d/ | 2 |
| | /k/ | 2 |
| | /g/ | 2 |
| Affricate | /dʒ/ | 1 and 3 |
| Nasal | /ŋ/ | 3 |
| Fricative | /v/ | 1 and 3 |
| | /ð/ | 1 |
| | /z/ | 1 and 3 |
| | /ʃ/ | 1 and 3 |
| | /ʒ/ | 1 and 3 |
| | /h/ | 1 |
| Approximant | /r/ | 2 and 3 |
| | /j/ | 2 and 3 |
| | /w/ | 1 |

Table 7: Problematic English consonant sounds for SS and types of mistake.

*SS= Spanish speakers

| Type of consonant | English consonant | Spanish sound(s) transferred |
|---|---|---|
| Affricate | /dʒ/ | /ʎ/, /j/, or /tʃ/ |
| Fricative | /v/ | /b/ or /f/ |
| | /ð/ | /d/ |
| | /z/ | /θ/ or /s/ |
| | /ʃ/ | /s/ |
| | /ʒ/ | /s/ |
| | /h/ | /x/ |
| Approximant | /w/ | /gu/ |

Table 8: Frequent SCS transferred into English by SS.

*SCS= Spanish consonant sounds

*SS= Spanish speakers

# 5. AUTOMATIC SPEECH RECOGNITION FOR TEACHING PRONUNCIATION

According to Techopedia (2017), "automatic speech recognition (ASR) is the use of computer hardware and software-based techniques to identify and process human voice […]."

With regard to pronunciation assessment, the most advanced systems incorporating ASR technology can offer feedback in terms of segmental —word-level— or suprasegmental —sentence-level— features (Elimat & AbuSeileek, 2014). Automatic feedback can vary from rejecting poorly pronounced utterances and acknowledging good ones to determining certain mistakes either in phonemic quality or sentence accent (Bunnel, Yarrington, & Poliknoff, 2000; Eskenazi & Pelton, 2002).

## 5.1. How does ASR work?

ASR computer programmes and software designed for pronunciation teaching follow a sequence of five phases (Neri, Cucchiarini, and Strik, 2003):

1. Speech recognition: The ASR engine translates the incoming speech signal into a sequence of words based on internal phonetic and syntactic patterns. Nonetheless, the key pedagogical benefit that ASR can offer for teaching oral skills in the EFL is the provision of an assessment of pronunciation quality.

2. Scoring: This phase makes it possible to provide a global assessment of pronunciation quality in the form of a score. The ASR system examines the spoken utterance that has been

earlier known. The analysis can be done based on a comparison between a student's utterance and a native's utterance. The main advantage of automatic scoring for pronunciation instruction is that it provides the learner instant information about the quality of their pronunciation.

3. Error detection: The system can ubicate the errors in the utterance and show the student where they make mistakes. Referring to any problematic phoneme within a word can be helpful to raise awareness in the learner of that difficulty and hence useful for them to focus and practice more on that area.

4. Error diagnosis: The ASR technology recognizes the precise type of mistake committed by the student and suggests how to improve it.

5. Feedback presentation: This phase presents the global score as a graded bar, or as a number on a given scale. This phase is essential, since the student will only be able to benefit from all the information obtained through ASR if shown in a meaningful way.

## 5.2. Advantages of ASR in the literature

A great array of studies on ASR have proven several benefits of its use when teaching pronunciation in an EFL classroom. In this case, the focus is put on six relevant pieces of research on the matter.

In first place, Mitra, Tooley, Inamdar, and Dixond (2003) analysed the role of the SR-based feedback in enhancing learners' pronunciation of problematic sounds. The research revealed that ASR can provide reliable feedback on pronunciation improvement over time.

In second place, in another study, Graff (2006) explored the role of SR in improving students' English pronunciation. The results indicated that subjects who practiced pronunciation skills with "Rosetta Stone" software, unlike those who practiced pronunciation with the traditional teaching, did experience a statistically considerable progress in the quality of their pronunciation.

In third place, Shams (2006) focused on researching the use of computerized pronunciation practice as a tool in the reduction of foreign language anxiety and to investigate

the use of ASR in improving students' pronunciation. The statistical analyses suggested that there was no connection between method of practice and the reduction in anxiety. Nevertheless, the results of the research revealed that the participants who practiced using ASR experienced a statistically meaningful improvement in the quality of their pronunciation whereas those who practiced with the cassettes did not.

In fourth place, Neri, Cucchiarini, and Strik (2008) tested the efficacy of ASR-based feedback for enhancing pronunciation. The results indicated that the group getting ASR-based feedback made the greatest improvement, but the groups' mean improvements did not vary substantially. However, the group receiving ASR-based feedback showed a considerably larger enhancement than the no-feedback group in the segmental quality of the challenging phonemes targeted.

In fifth place, Hinks and Edlund (2009) investigated the impact of SR-based visual feedback in enhancing pitch. The experimental group showed a higher improvement than the control group, suggesting that the feedback was effective. These positive findings imply that the feedback could be positively used in a system for practicing oral presentations.

Finally, a study was conducted by Lai, Tsai, and Yu (2009) to examine the efficacy of using a multimedia English learning (MEL) system, based on ASR for teaching learners to boost their English phonetic awareness and pronunciation. The results revealed that the students' pronunciation skills in the experimental group progressed more than did their classmates in the control group.

In short, the literature reveals that ASR can provide reliable and effective feedback on pronunciation improvement, students experience progress in the quality of their pronunciation, the segmental quality of the challenging phonemes is enhanced, and learners' English phonetic awareness is boosted.

# 6. METHODOLOGY

## 6.1. Participants

8 3-ESO bilingual students from a secondary school in Valladolid (Castile and León, Spain) participated in this study and their oral production was analysed.

All the participants of this study were aged 15, with a 31-hour weekly timetable. In total, 12 hours of their timetable are devoted to English, of which 4 hours are dedicated to the subject of English as a Foreign Language. Their bilingual-section subjects are Music, Technology, Arts, and Physical Education. All of them have been enrolled in the bilingual section since the first year of primary education.

## 6.2. 'SpeechAce'

'SpeechAce' (https://app.SpeechAce.co/placement/) is an ASR software that offers a 16-unit English pronunciation practice, as shown in Figure 33, —Level 1, Healthcare, Hospitality-Vietnam, Beginner, Vowels 1, Consonants 1, Fluency, Fluency Speaking Practice, Vowels 2, Consonants 2, Simple Sentences, Basic Sentences, Beginner Sentences, Intermediate Sentences, Citizenship, and Poem— each of which consisting of several lessons. An example of how lessons are presented is provided in Figure 34.

In the case of the present study, the focus is set on the ones dealing specifically with segmental pronunciation features —i.e., Vowels 1 (/ɪ/ and /iː/, /e/ and /æ/, /ə/ and /ʌ/, /ɔː/ and /ɑː/, and /ʊ/ and /uː/), Vowels 2 (/ər/; /ɪr/, /er/ and /ɔːr/; /ʊr/ and /ɑːr/; /ɑɪ/, /eɪ/, and /ɔɪ/; and /aʊ/ and /əʊ/), Consonants 1 (/b/ and /p/, /d/ and /t/, /g/ and /k/, /dʒ/ and /tʃ/, /v/ and /f/, and /ð/ and /θ/), and Consonants 2 (/s/ and /z/, /ʃ/ and /ʒ/; /m/, /n/ and /ŋ/; /l/ and /r/; and /h/, /w/ and /j/).

This tool provides an array of words and sentences in each of the lessons available. The students record their own voice in an attempt of imitating the native model presented by the software. Afterwards, they get their feedback, which is rated to a percentage up to 100% (Figure 35). Additionally, every time the students pronounce words or sentences in a wrong manner, the software provides what exactly is wrong in detail, highlighting the correct syllable, the phoneme that the learners pronounce inaccurately and the score —good/bad and how it actually sounds like— (Figure 36). In addition, it is of relevance to remark the fact that the same word or phrase can be repeated as many times as needed until the highest result is achieved (Σαμουλαδά, 2019).


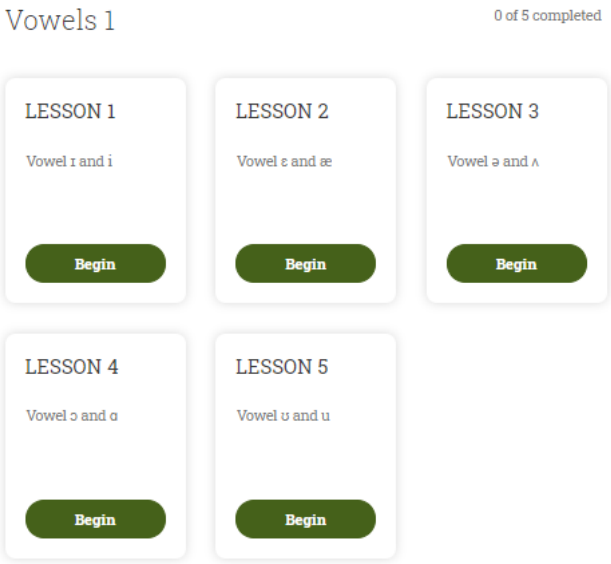


Figure 34: Overview of the lessons available in Vowels 1.

Figure 33: Overview of the units available on ‘SpeechAce’.

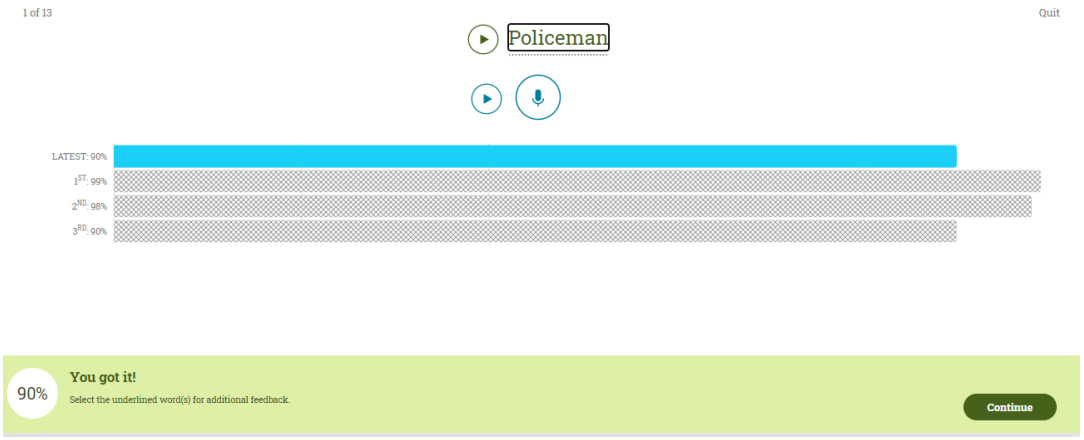


Figure 35: Overview of the ASR interface (Vowels 1, Lesson 1).

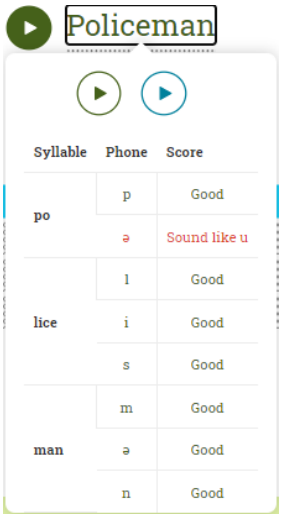


Figure 36: Detailed feedback given by ‘SpeechAce’.

## 6.3. Stages of the study

The present study was divided into three stages.

In the first stage, upon parental consent (Annex 1), a recording was made of the participants pronouncing a list of words containing instances of every single simple vowel, diphthong, and consonant sound available in the English phonological repertoire (Annex 2). This was done so in order to detect which English sounds were the most challenging for the participants prior to the use of the ASR software.

In the second stage, the participants were required to sign in on ‘‘SpeechAce’’, and work on four specific units of the software —Vowels 1, Vowels 2, Consonants 1, and Consonants 2— for 2 weeks. In order to motivate the participants of the study, an English pronunciation contest was conducted taking into account the punctuation which they were awarded by the software. Thus, they would be more liable to insist even on the pronunciation of the phonemes they struggled with.

In the last stage, the participants were recorded reading aloud the same word list as they did 2 weeks before so as to detect English pronunciation improvements, if any, upon the use of ‘SpeechAce’.

## 6.4. Data classification

The data obtained from both recordings of the 8 students participating in the present study were classified in an Excel database. The data were classified in 10 spread sheets as they appear in the corresponding database that can be found in the following link: https://bit.ly/3xyHqmL The first spread sheet contains the data from participant "A", the second contains the data from participant "D", the third contains the data from participant "G", the fourth contains the data from participant "I", the fifth contains the data from participant "L", the sixth contains the data from participant "M", the seventh contains the data from participant "N", the eighth contains the data from participant "P", the ninth contains the average data from all the participants, and the tenth contains the specific mispronunciations committed by the participants of the study. All these letters are used as a code corresponding to the initial letter of the participants' names.

The data presented in the spread sheets were classified into different variables. From spread sheet 1 to 8, a first set involves the number of mispronunciations of each phoneme before and after the use of 'SpeechAce', followed by the total number of instances of each phoneme which can be found in the word list which they were required to read aloud. On Table 9, an instance of this first set, retrieved from participant "A"'s spread sheet, is presented.

| Type of sound | Subcategory | Phoneme | Before | After | Total |
|---|---|---|---|---|---|
| Consonants | Occlusive | /p/ | 1 | 0 | 4 |
| | | /b/ | 1 | 0 | 1 |
| | | /t/ | 0 | 0 | 4 |
| | | /d/ | 0 | 0 | 3 |
| | | /k/ | 0 | 0 | 3 |
| | | /g/ | 1 | 0 | 2 |

Table 9: Variables in the first set (spread sheets 1-8).

Then, a second set involves the mispronunciation rate of each phoneme before and after the use of 'SpeechAce', followed by the improvement rate of each phoneme and type of phoneme which each participant has shown after the use of the ASR software. On table 10, an example of this second set of variables, retrieved from participant "A"'s spread sheet, is provided.

| Type of sound | Subcategory | Before | After | Improvement |
|---|---|---|---|---|
| Consonants | Occlusive | 25% | 0% | 25% |
| | | 100% | 0% | 100% |
| | | 0% | 0% | 0% |
| | | 0% | 0% | 0% |
| | | 0% | 0% | 0% |
| | | 50% | 0% | 50% |
| | | 58% | 0% | 58% |

Table 10: Variables in the second set (spread sheets 1-8).

In spread sheet 9, 2 sets of data can be found. In the first one, the average mispronunciation rate of each phoneme and type of phoneme before and after the use of 'SpeechAce' is presented, followed by the average improvement rate of each phoneme and type of phoneme. Besides, the overall percentage of improvement of all the participants is given below. On Table 11, an example of this first set of variables is provided.

| Type of sound | Subcategory | Phoneme | Before | After | Improvement |
|---|---|---|---|---|---|
| Consonants | Occlusive | /p/ | 59% | 25% | 34% |
| | | /b/ | 88% | 13% | 75% |
| | | /t/ | 50% | 13% | 38% |
| | | /d/ | 8% | 4% | 4% |
| | | /k/ | 50% | 17% | 33% |
| | | /g/ | 69% | 25% | 44% |
| | | Average | 67% | 16% | 51% |

Table 11: Variables in the first set (spread sheet 9).

In the second set of variables, the average mispronunciation rate of consonant and vowel sounds in general before and after the use of the software is provided, followed by the average improvement rate. On Table 12, an example of this second set of variables is shown.

| Consonants | | |
|---|---|---|
| Before | After | Improvement |
| 67% | 29% | 38% |

Table 12: Variables in the second set (spread sheet 9).

Finally, in spread sheet 10, the specific types of mistakes and mispronunciations committed by the participants are indicated. On Table 13, an instance of this last set of variables is presented.

| Type of sound | Subcategory | Phoneme | Type of mistakes | Mispronunciation |
|---|---|---|---|---|
| Consonants | Occlusive | /p/ | 2 | Spanish /p/ |
| | | /b/ | 2 | Spanish /b/ |
| | | /t/ | 2 | Spanish /t/ |
| | | /d/ | 2 | Spanish /d/ |
| | | /k/ | 2 | Spanish /k/ |
| | | /g/ | 2 | Spanish /g/ |

Table 13: Variables in the last set.

## 6.5. Research questions

The data elicited and classified as presented before have been so in order to account for English pronunciation improvements, if any, which the use of 'SpeechAce' may bring about. The main objectives were to detect the participants' weak points concerning segmental features of English pronunciation before using 'SpeechAce', to find out the pronunciation of which sounds is enhanced upon the use of the latter, and to discover the degree of improvement which the participants show after using the software. These aims were articulated into the following research questions which are the ones that guide the present study and, in particular, the data analysis that appears in the subsequent section.

The following research questions go from the more general ones, dealing with the overall improvement presented by the participants, the most challenging phonemes, and the most frequent type of pronunciation mistakes, to the more specific ones, dealing with the concrete types of sounds which, on the one hand, proved to be the most problematic and, on the other hand, the ones which experimented a higher improvement rate upon the use of the ASR software.

- Research question 1: What percentage of overall pronunciation improvement do the participants show after the use of 'SpeechAce'?

- Research question 2: With which phonemes did the participants struggle?

- Research question 3: In which ways were those phonemes mispronounced?

- Research question 4: Which were the most problematic sounds —vowels or consonants— before the use of ‘SpeechAce’?

- Research question 5: Which kind of sound —vowels or consonants— presents a higher improvement rate after the use of ‘SpeechAce’?

- Research question 6: Which type of vowel sound —simple or diphthongs— were the most problematic before the use of ‘SpeechAce’?

- Research question 7: Which type of vowel sound —simple or diphthongs— presents a higher improvement rate after the use of ‘SpeechAce’?

- Research question 8: Which type of consonant sounds were the most problematic before the use of ‘SpeechAce’?

- Research question 9: Which type of consonant sounds presents a higher improvement rate after the use of ‘SpeechAce’?

# 7. DATA ANALYSIS

In this section, the results obtained from the present study are examined as a means to give an answer to the research questions established. Thereby, the data analysis was structured in view of these research questions.

Research question 1 is concerned with the overall English pronunciation improvement which the participants showed upon the use of ‘SpeechAce’. The result appears in Table 13.

| | |
|---|---|
| Overall improvement of the participants | 38% |

Table 14: Overall pronunciation improvement.

As Table 14 shows, the participants achieved a 38-% overall English pronunciation improvement after using the software for 2 weeks — i.e., the pronunciation of the words in the word list in the second recording shows a 38% improvement if compared to the first recording.

Thus, ‘SpeechAce’ has proved to be a favourable tool when willing to enhance students’ English pronunciation in terms of segmental features.

Research questions 2 and 3 are concerned with the phonemes with which the participants struggled when pronouncing the words in the word list and the ways in which they were mispronounced (Table 15).

| Type of sound | Subcategory | Phoneme | Type of mistakes | Mispronunciation |
|---|---|---|---|---|
| Consonants | Occlusives | /p/ | 2 | Spanish /p/ |
| | | /b/ | 2 | Spanish /b/ |
| | | /t/ | 2 | Spanish /t/ |
| | | /d/ | 2 | Spanish /d/ |
| | | /k/ | 2 | Spanish /k/ |
| | | /g/ | 2 | Spanish /g/ |
| | Affricates | /dʒ/ | 1 | Spanish /j/ |
| | Fricatives | /v/ | 3 | Spanish /b/ |
| | | /ð/ | 1 | Spanish /d/ |
| | | /z/ | 1 | Spanish /s/ |
| | | /ʃ/ | 1 | Spanish /s/ |
| | | /ʒ/ | 3 | Spanish /s/ |
| | | /h/ | 1 | Spanish /x/ |
| | Approximants | /ɾ/ | 1 | Spanish /ɾ/ |
| | | /j/ | 3 | Spanish /j/ |
| | | /w/ | 1 | Spanish /gu/ |
| Vowels | Simple | /ʌ/ | 1 | Spanish /a/ |
| | | /aː/ | 1 | Spanish /a/ |
| | | /æ/ | 1 | Spanish /a/ |
| | | /e/ | 1 | Spanish /e/ |
| | | /ə/ | 3 | Influenced by spelling |
| | | /ɜː/ | 1 | Spanish /e/ |
| | | /ɔː/ | 1 | Spanish /o/ |
| | | /ʊ/ | 1 | Spanish /u/ |
| | | /uː/ | 1 | Spanish /u/ |
| | Diphthongs | /aɪ/ | 1 | Spanish /ai/ |
| | | /aʊ/ | 1 | Spanish /au̯/ |
| | | /eɪ/ | 1 | Spanish /ei/ |
| | | /əʊ/ | 1 | Spanish /ou̯/ |
| | | /ɔɪ/ | 1 | Spanish /oi/ |
| | | /eə/ | 1 | Spanish /ea/ |
| | | /ɪə/ | 1 | Spanish /ía/ |
| | | /ʊə/ | 1 | Spanish /úa/ |

Table 15: Problematic phonemes and their mispronunciations.

As it can be observed on Table 15, the participants mispronounced, in the case of consonant sounds, occlusives —/p/, /b/, /t/, /d/, /k/, and /g/—, one affricate —/dʒ/—, fricatives —/v/, /ð/, /z/, /ʃ/, /ʒ/, and /h/—, and approximants —/r/, /j/, and /w/— and, in the case of vowel sounds, every simple vowel —/ʌ/, /a:/, /æ/, /e/, /ə/, /ɜ:/, /ɔ:/, /ʊ/, and /u:/— and diphthong —/aɪ/, /aʊ/, /eɪ/, /əʊ/, /ɔɪ/, /eə/, /ɪə/, and /ʊə/. Thus, these phonemes will be the ones at stake in the present study.

Prominence must be given to the fact that the majority of the mispronunciations correspond to 'Type 1' mistakes —consisting in the substitution of an English phoneme which does not exist in the Spanish phonology for the most approximate phoneme in Spanish. This is the case of /dʒ/, /ð/, /z/, /ʃ/, /h/, /r/, /w/, and all vowel sounds —except for /ə/. Regarding 'Type 2' mistakes —caused by distribution differences among phonemes shared by English and Spanish—, only the mispronunciations related to occlusives —/p/, /b/, /t/, /d/, /k/, and /g/— fall into this category. Finally, 'Type 3' mistakes — rooted in the fact that, as opposed to Spanish, English is not a phonetic language— can only be observed in the mispronunciations of /v/, /ʒ/, /j/, and /ə/.

Research questions 4 and 5 deal with the most problematic sounds —vowels or consonants— before the use of 'SpeechAce' and the ones which showed a higher improvement rate upon its use. The results appear on Table 16.

| | Before | After | Improvement |
|---|---|---|---|
| Consonants | 67% | 29% | 38% |
| Vowels | 63% | 26% | 37% |

Table 16: Incorrectness rate before and after, and improvement presented (consonants vs. vowels).

As it can be seen in Table 16, there is no significant difference between consonant and vowel sounds in terms of the incorrectness rate shown prior to the use of 'SpeechAce' (67% vs. 63%) and the improvement rate presented upon its use (38% vs. 37%). That is to say, both types of sounds present a similar degree of difficulty for the participants, which can be overcome to the same extent through the use of 'SpeechAce'.

Research questions 6 and 7 are concerned with the most problematic type of vowel sound —simple vowels or diphthongs— before the use of the software and the one presenting a higher improvement rate after its use. This information appears in Table 17.

| | Before | After | Improvement |
|---|---|---|---|
| Simple | 76% | 32% | 44% |
| Diphthongs | 51% | 20% | 31% |

Table 17: Incorrectness rate before and after, and improvement presented (simple vowels vs. diphthongs).

As Table 17 shows, if compared to diphthongs, simple vowels present both a higher mispronunciation rate prior to the use of 'SpeechAce' (76% vs. 51%) as well as a higher improvement rate upon its use (44% vs. 31%). Accordingly, the participants found it more difficult to articulate English simple vowels than diphthongs, possibly due to the fact that the latter are more similar to their Spanish equivalents than the former. Plus, the use of 'SpeechAce' seems to be more effective when it comes to English simple vowels' than to diphthongs' pronunciation enhancement.

Finally, research questions 8 and 9 are concerned with the most problematic type of consonant sounds prior to the use of the 'SpeechAce' and the one presenting a higher improvement rate upon its use. The results are provided in the table below.

| | Before | After | Improvement |
|---|---|---|---|
| Occlusive | 67% | 16% | 51% |
| Affricate | 88% | 63% | 25% |
| Fricative | 81% | 37% | 44% |
| Approximant | 38% | 7% | 31% |

Table 18: Incorrectness rate before and after, and improvement presented per type of consonant.

As shown in Table 18, addressing first to research question 8, affricates (88%) were the most challenging consonants for the participants before the use of the ASR software, followed by fricatives (81%), occlusives (67%), and approximants (38%). Then, addressing to research question 9, occlusives (51%) are the most enhanced consonant sounds after its use, followed by fricatives (37%), approximants (31%), and affricates (25%).

# 8. CONCLUSION

This dissertation presents 'SpeechAce' as a tool to overcome segmental difficulties in English pronunciation in secondary school students. In order to prove its efficacy and get an insight into the main difficulties with which they are faced in terms of segmental features' pronunciation, a study was conducted with 8 3-ESO students enrolled in the bilingual section in a secondary school located in Valladolid. The participants were required to use 'SpeechAce' for 2 weeks. Furthermore, 2 recordings —prior to and upon the use of the ASR software— were made of them reading aloud a word list containing instances of every English phoneme. Data from both recordings were classified in terms of four different variable sets: a first set involving the number of mispronunciations of each phoneme before and after the use of 'SpeechAce' by each participant; a second set involving the mispronunciation rate of each phoneme before and after the use of 'SpeechAce', followed by the rate of improvement of each phoneme and type of phoneme by each participant; a third set involving the average mispronunciation rate of each phoneme and type of phoneme before and after the use of 'SpeechAce', followed by the average improvement rate of each phoneme and type of phoneme; a fourth set involving the average mispronunciation rate of consonant and vowel sounds in general before and after the use of the software, followed by their average improvement rate; and a fifth set involving the specific types of mistakes and mispronunciations committed by the participants. From this, several conclusions were drawn after having provided answer to the previously established research questions.

First, 'Type 1' mistakes are the most common mispronunciations among the participants of the study, which is not surprising due to the fact that it is common occurrence for L1 speakers of Spanish to substitute an English phoneme which does not exist in the Spanish phonological repertoire for the most approximate phoneme available in Spanish. Moreover, consonants and vowels were similarly challenging for the participants, albeit it is true that the former presented a slightly higher mispronunciation rate than the latter. Plus, on the one hand, simple vowels have proved to be more problematic than diphthongs in terms of vowel sounds and, on the other hand, fricatives were the most demanding consonant sounds for the participants.

Second, upon the use of 'SpeechAce', consonants have shown a greater improvement rate than vowels, simple vowels a higher one than diphthongs, and occlusives —closely followed by fricatives— are the consonant sounds which present the highest improvement rate.

Consequently, ‘SpeechAce’ has proved to be a useful tool for secondary school students to enhance their pronunciation of English phonemes, with a view to overcome those difficulties related with the most challenging phonemes for them —as the results of the study show.

The present study could be expanded by increasing the number of participants in order to get more robust results. Besides, making more recordings over time could lead to interesting conclusions concerning the effectiveness of the ASR software in the long run. Moreover, it would be appealing to take into consideration both segmental and suprasegmental features of pronunciation so as to provide a more complete picture of the participants’ difficulties in terms of English pronunciation and ‘SpeechAce’’s effectiveness.

## 9. BIBLIOGRAPHY


'SpeechAce'. (n.d.). Retrieved 19 June 2021, from https://app.'SpeechAce'.co/placement/

Adey, M. (2015). La enseñanza de la pronunciación de la lengua inglesa a alumnos del cuarto curso de Primaria (Bachelor's tesis, Universidad Internacional de la Rioja).

Brunori, R. (2016). English pronunciation for Spanish (and other non-native) speakers. Retrieved from http://www.englishpronunciationmadrid.com/

Bunnell, H. T., Yarrington, D. M., & Polikoff, J. B. (2000). STAR: articulation training for young children. In *Sixth International Conference on Spoken Language Processing*.

Clegg, J. H., & Fails, W. C. (2017). *Manual de fonética y fonología españolas*. Routledge.

Coe, S (1987). Speakers of Spanish and Catalan. In M. Swan (Ed.) *Learner English: A Teacher's Guide to Interference and Other Problems, Volume 1* (90-113). Cambridge. Cambridge University Press.

Del Puerto, F. G. (2004). La adquisición de la pronunciación del inglés como tercera lengua. (Doctoral dissertation, Universidad del País Vasco-Euskal Herriko Unibertsitatea).

Del Puerto, F. G., & Lacabex, E. G. (2008). La enseñanza de las vocales inglesas a los hablantes de español. *PULSO. Revista de Educación*, (31), 37-66.

Elimat, A. K., & AbuSeileek, A. F. (2014). Automatic speech recognition technology as an effective means for teaching pronunciation. *JALT CALL Journal*, *10*(1), 21-47.

Eskenazi, M., & Pelton, G. (2002). Pinpointing pronunciation errors in children's speech: examining the role of the speech recognizer. In *ISCA Tutorial and Research Workshop (ITRW) on Pronunciation Modeling and Lexicon Adaptation for Spoken Language Technology*.

Estebas-Vilaplana, E. (2009). *Teach yourself English pronunciation*. La Coruña: Netbiblo.

Finch, D. F., & Lira, H. O. (1982). *A course in English phonetics for Spanish speakers*. Heinemann Educational Books.

Gilakjani, A. P. (2012). A study of factors affecting EFL learners' English pronunciation learning and the strategies for instruction. *International Journal of Humanities and Social Science*, *2*(3), 119-128.

Gilakjani, A. P., & Sabouri, N. B. (2016). How can EFL teachers help EFL learners improve their English pronunciation? *Journal of Language Teaching and Research*, *7*(5), 967-972.

Gómez González, M. D. L. Á., & Sánchez Roura, T. (2016). *English pronunciation for speakers of Spanish: From theory to practice*. De Gruyter.

González Panero, L. (2016). La pronunciación de las consonantes inglesas y su dificultad para el hispanohablante: una investigación de campo.

Gorman, K. B., & Kester, E. S. (2001). Spanish Influenced English: typical phonological patterns in the English language learner. In *Seminar offered by Bilingiustics at ASHA*.

Graff, M. M. (2006). *A study of Rosetta Stone's effectiveness on improving English pronunciation*. California State University, Dominguez Hills.

Hincks, R., & Edlund, J. (2009). Promoting increased pitch variation in oral presentations with transient visual feedback. *Language Learning & Technology*, *13*(3), 32-50.

Judson, J. (2012). *The Sound of English: A Practical Course in British English Pronunciation*. London: Pronunciation Studio.

Khaghaninejad, M. S., & Maleki, A. (2015). The effect of explicit pronunciation instruction on listening comprehension: Evidence from Iranian English learners. *Theory and Practice in Language Studies*, *5*(6), 1249-1256.

Martínez Asís, F. (2004). Estudio De Una Intervención Pedagógica Para La Enseñanza De La Pronunciación Inglesa En 4º Curso De La E.S.O. (Doctoral dissertation, Universidad de Murcia).

Mateo Plana, S., & González Pastor, E. (2016). La pronunciación del inglés: dificultades para los hispanohablantes y propuesta de actividades.

Mitra, S., Tooley, J., Inamdar, P., & Dixon, P. (2003). Improving English pronunciation: An automated instructional approach. *Information Technologies & International Development*, *1*(1), pp-75.

Monroy Casas, R. (1980). *Aspectos fonéticos de las vocales españolas*. Madrid: Paraninfo

Monroy Casas, R. (2012). *La pronunciación del inglés británico simplificada*. Murcia: Editum.

Neri, A., Cucchiarini, C., & Strik, H. (2003). Automatic speech recognition for second language learning: How and why it actually works. In *Proc. ICPhS* (pp. 1157-1160).

Neri, A., Cucchiarini, C., & Strik, H. (2008). The effectiveness of computer-based speech corrective feedback for improving segmental quality in L2 Dutch. *ReCALL: The Journal of EUROCALL*, *20*(2), 225.

ORDEN EDU/362/2015, de 4 de mayo, por la que se establece el currículo y se regula la implantación, evaluación y desarrollo de la educación secundaria obligatoria en la Comunidad de Castilla y León.

Rolo Cruz, R. J. (2019). La Enseñanza y el Aprendizaje de la Pronunciación del Inglés en Enseñanza Secundaria.

Shams, A. N. (2005). The use of computerized pronunciation practice in the reduction of foreign language classroom anxiety (Doctoral dissertation, Florida State University).

Techopedia. (2017). Automatic Speech Recognition (ASR). Retrieved 2 June 2021, from https://www.techopedia.com/definition/6044/automatic-speech-recognition-asr

Villarín, C., & Helíodora, M. (1998). Consideraciones para la Enseñanza de la Pronunciación Inglesa a Hablantes Nativos de Español.

Yates, L., & Zielinski, B. (2009). *Give it a go: Teaching pronunciation to adults*. Macquarie University.

Yen-Shou, L., Hung-Hsu, T., & Pao-Ta, Y. (2009). A multimedia English learning system using HMMs to improve phonemic awareness for English learning. *Educational Technology & Society*, *12*(3), 266-282.

Σαμουλαδά, Ζ. (2019). Βελτίωση της προφοράς της αγγλικής γλώσσας μέσω διαδικτυακών εργαλείων.

# 10. ANNEXES

## ANNEX 1

### Letter addressed to parents

Estimadas familias:

Mi nombre es Lidia Avellón Mayor, estudiante del Máster en Profesor de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanzas de Idiomas (especialidad en inglés como lengua extranjera) de la Universidad de Valladolid.
Me dirijo a ustedes para informarles sobre el estudio que voy a realizar junto al profesor Miguel Vilalta Nieto (Universidad de Valladolid) para mi Trabajo de Fin de Máster, que versa sobre los beneficios de "Englace" (una aplicación móvil que ayuda a mejorar la pronunciación en inglés). Dicho estudio consta de tres fases:

1) La grabación de audio de los alumnos para analizar su pronunciación en inglés antes del uso de la aplicación "Englace".
2) El uso de la aplicación "Englace" por parte de los alumnos.
3) La grabación de audio de los alumnos para analizar su pronunciación en inglés después del uso de la aplicación "Englace".

El uso de esta aplicación se realizaría en horas no lectivas y en formato de concurso de pronunciación inglesa.
Para respetar la privacidad de los alumnos, las grabaciones serán anónimas y tendrán como único fin el ser objeto de estudio para mi Trabajo de Fin de Máster.

Agradezco su colaboración de antemano.

Reciban un cordial saludo,

Lidia Avellón Mayor

## Parental consent document

D./Dña. _____, mayor de edad, titular del DNI _____, padre, madre o tutor legal de _____, por el presente documento manifiesto mi consentimiento para la grabación de audio en el marco de la investigación llevada a cabo por Lidia Avellón Mayor para el Trabajo de Fin de Estudios del Máster de Profesor de Educación Secundaria Obligatoria, Formación Profesional y Enseñanza de Idiomas de la Universidad de Valladolid.

Las grabaciones se utilizarán únicamente con fines de investigación. Los resultados del estudio son susceptibles de divulgación o publicación, pero en tal caso se anonimizarán debidamente los datos utilizados de modo que los sujetos de investigación no serán identificados o identificables.

Nombre y firma del PADRE / MADRE / TUTOR

Lugar y fecha:

## ANNEX 2

## Study's word list

**Vowel sounds**

- Ship
- Sheep
- Put
- You
- Hear
- Name
- Ten
- Letter
- Girl
- Saw
- Poor
- Toy
- No
- Cat
- Sun
- Car
- Hot
- Where
- My
- How

**Consonant sounds**

- Pot
- Bed
- Tap
- Door
- Chair
- Jam
- Come
- Gum
- Four
- Very
- Think
- The
- Son
- Zoo
- She
- Vision
- Man
- Nose
- Angry
- Happy
- Late
- Red
- Want
- Yes