Acoustic analysis of EFL speech production corpora according to the Speech Learning Model

Jon Lapresa Serrano

Degree in English Studies
Department of English Philology and Translation and Interpretation
Area of English

Supervisor: Maria Luisa García Lecumberri
Academic Year: 2018-2019

ABSTRACT

It is widely accepted that L1 influences the acquisition of L2 phonology during Second Language Acquisition. Models such as the Speech Learning Model (SLM) have tried to explain this process by defining bi-directional interactions between the phonetic systems of both languages. Flege (2007) called them assimilation and dissimilation, by which non-native sounds are classified as either "new", "identical" or "similar". "New" would be likely not to be assimilated to any L1 category and, therefore, a new L2 category would be created. "Identical", on the other hand, would be those sounds that the speaker assimilates, producing them in a native-like manner. Meanwhile, the ones classified as "similar" would be the most complicated sounds to both perceive and produce. The aim of this paper is to determine whether SLM can explain and predict FL (instead of L2) production mistakes by classifying them acoustically, i.e. using spectrograms. The data was provided by the supervisor of this paper, Dr. María Luisa García Lecumberri, and gathered at the University of Seville using the DiapixFL method as part of a larger corpus. This means that in order to prompt conversational speech between two subjects, a spot-the-difference task was used. In this case, native speakers of Spanish (more specifically, the Andalusian Spanish variety) were recorded. Afterwards, their production of English as a Foreign Language (EFL) was orthographically and phonetically transcribed and eventually classified by creating spectrograms. Finally, this classification of consonants and vowels was analysed in order to reach the following conclusion. The analysis of the results shows the limitations of SLM when dealing with phenomena such as hyper-correction and spelling pronunciation, as well as specific features of the variety of the study subjects. It would appear that, even though the acoustic analysis offered resourceful results, a more detailed research is preferable for upcoming occasions. Moreover, future collection of data is suggested, specifically in the Catalan and Galician regions. This addition could offer the chance to compare the production of English phonemes in native speakers of different varieties of Spanish.

Keywords: Speech Learning Model, English as a Foreign Language, Spanish speakers of English, acoustic analysis

INDEX

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CAH	Contrastive Analysis Hypothesis
CPH	Critical Period Hypothesis
EFL	English as a Foreign Language
fem	feminine
FL	Foreign Language
L1	First/native language
L2	Second language
masc	masculine
NLM	Native Language Magnet Model
OM	Ontogeny Model
OPM	Ontogeny and Phylogeny Model
PAM	Perceptual Assimilation Model
PSp	Peninsular Spanish
RP	Received Pronunciation
SLA	Second Language Acquisition
SLM	Speech Learning Model
S3B	Speaker 3B
S4B	Speaker 4B
U	Universals

1. INTRODUCTION

Since the mid-twentieth century, the acquisition of second language (L2) phonology and pronunciation errors have been explained based on the influence, also called transfer, of the speakers' native language (L1) (Eckman, 2004:515). Although many models and theories have been developed around that idea, not much research has been done relating it to the production of a foreign language (FL), i.e. a language typically taught in a class environment. When compared to L2, FL does not hold a considerable position inside the country or region the speaker is in (Richards, & Schmidt, 2010). As Freed (1995) explains, there are “contrasts which oppose the at home and immersion language learning environments; such contrasts often describe the former as "foreign language learning" and the latter as "second language acquisition." In the case of language learning which occurs in a study abroad context, this distinction becomes somewhat blurred.” (p. 4). An example they offer is an American student who is learning French in the USA (French as a foreign language) but afterwards decides to study temporarily in France (French as a second language or not?).

The literature review that is presented in this paper will furtherly explain the different second language acquisition (SLA) models that deal with L1 influence and only refer to L2, i.e. with native and abundant input (FL is mostly non-native and involves very little input). Furthermore, it will also offer some background information of the Critical Period Hypothesis (CPH) and the Contrastive Analysis Hypothesis (CAH), crucial for understanding the models. To finalise the review, and due to its relevance on the upcoming acoustic analysis, I will introduce the characteristics of Andalusian Spanish.

The relevance of this variety of Spanish is then related to the two subjects of the current study, which consists in an acoustic analysis of a set of recordings of Spanish native speakers from the city of Seville conversing in English. I will classify their production mistakes by using spectrograms directly drawn out of the recordings. By following one of the previously mentioned SLA models, more specifically the Speech Learning Model (SLM), I will try to predict these mistakes. Thus, the aim of this paper is to determine whether SLM can explain and predict FL production mistakes by classifying them acoustically.

2. LITERATURE REVIEW

2.1. SECOND LANGUAGE ACQUISITION (SLA) THEORIES

The influence of L1 on the acquisition of L2 phonology is currently widely accepted, even by those who doubt its influence in the language's syntax (Eckman, 2004:515). Since the middle of the 20th century, several models have emerged trying to explain this influence, e.g. the Ontogeny and Phylogeny Model (OPM), the Native Language Magnet Model (NLM), the Perceptual Assimilation Model (PAM) and the model chosen for the current analysis, the Speech Learning Model (SLM). But to better understand the development of all these models, I shall first explain two precedent hypotheses, precursors in trying to explain the influence (then called 'transfer') of L1: the Critical Period Hypothesis (CPH) and the Contrastive Analysis Hypothesis (CAH).

The Critical Period Hypothesis states that, once a certain age is reached (12-15), a speaker will be unable to produce a second language at a native-like level due to physiological changes that occur in the brain, often related to its plasticity (Lenneberg, 1967). As a result, in phonetic terms, the speaker will retain the accent of their L1. During the more than fifty years since it was formulated, this hypothesis has been a relevant subject in SLA. Some of the discussions have involved the existence of several critical periods, along with the so-called 'sensitive' but not 'critical' period, the lack of such or the differences between gradual and continual decline after puberty (Abello-Contesse, 2009).

The Contrastive Analysis Hypothesis, claims that the differences between the speaker's mother tongue (L1) and the foreign language (L2 or FL) could explain how the L2 is produced (Lado, 1957; in Eckman, 2004). L2 sounds which are different will be produced by transfering these from the L1. The pronunciation errors are explained and predicted by comparing the inventory distribution of phonemes and their allophones in the L1 and L2. An example of this is that of Stockwell and Bowen (1965), who compared English and Spanish. They categorised the sounds of both languages into three different groups: optional (a phoneme with an unpredictable distribution based on the surrounding phonological context), obligatory (allophones, being environmentally conditioned) and null (does not belong to the language). They

reached the conclusion that the most difficult sound to learn as a FL speaker would be an L2 allophone that is null in the mother tongue (Stockwell, & Bowen, 1965; in Eckman, 2004). All of the models we will review below follow, to some extent, this comparing method of the CAH.

2.1.1. Ontogeny and Phylogeny Model (OPM)

The first of the models is the Ontogeny and Phylogeny Model (OPM), whose name is directly inspired in two biological concepts. As Major (2001) himself explains, "ontogeny is the life cycle of a single organism, whereas phylogeny is the evolutionary development of groups of organisms, such as the origin and development of species" (p. 80). Originally, Major called it just the Ontogeny Model (OM) in 1986, later relating those two concepts to linguistic terms: ontogeny would be the life cycle of one's language and phylogeny would be that of languages and language types, including different phenomena such as change, loss, contact and variation. Major claims that the acquisition of a second language involves three different components: the native language (L1), the foreign language, i.e. language being acquired (L2/FL) and the universals (U). The model states that the development of these components is that as the L2 increases, the L1 decreases. The universals first increase but then decrease. This development can vary, though, depending on the similarity between L1 and L2; the more similar, the more influence/transfer from L1, decreasing more slowly.

2.1.2. Native Language Magnet Model (NLM)

The Native Language Magnet (NLM) model, as stated by Kuhl et al. (2005), focuses "on infant's native phonetic categories and how they could be structured trough ambient language experience" (p. 982). Officially established by Kuhl in 1993, the model describes three phases in the development of the child's speech perception:

Phase 1. They can distinguish every sound category of human speech due to their auditory processing mechanism (and not a speech-related one).

Phase 2. They experience a "'warping' perception" (Kuhl et al., 2005:982), distorting that previous perceptual sensitivity around the prototypes (the most often activated sound category representations). These begin to act like magnets, i.e. non-prototypes that are similar to prototypes are difficult to differentiate from them. As time is needed for this phase to take place, it usually occurs between the sixth and twelfth

month of the infant's life.

Phase 3. This distortion, called ‘perceptual magnet effect’, facilitates native but hinders foreign phonetic abilities. As a result, foreign language (FL) sounds that are similar to those of the L1 are more difficult to perceive than those that are more different.

2.1.3. Perceptual Assimilation Model (PAM)

Another SLA model is Perceptual Assimilation model (PAM), first developed in 1995 by Best. This theory, which yet again tries to explain the influence of L1 in L2 speech perception and, consequently, also production, states that the similarity between each of the two languages' sounds affects the difficulty of their acquisition. Just as stated above with the other models, the more different a sound is compared to that of the L1, the easier it will be to perceive and produce (Best, 1995; Best, & Tyler, 2007). Moreover, this model claims that languages differ in their articulatory gestures, so the perception of foreign language speech is also affected by the articulatory system of each language. PAM also predicts three different categories or “patterns of perceptual assimilation” for the L2 sound (Best, 1995:194):

- Assimilated. An L2 sound can vary in regards to its similarity to the L1 as a good, acceptable but not ideal, or a deviant exemplar. This category is the one predicted to be the most difficult to be perceived.

- Uncategorised. The L2 sound would not be an obvious example of any native category. In contrast with the previous pattern, this one would be the easiest one to be perceived.

- Non-assimilated. In this case, the L2 sound is perceived as a non-speech one.

This model would be ideal for the current analysis, due to the idea of similarity and its relevance within the patterns. Instead I chose SLM because I would not be able to conduct an analysis of the articulatory gestures, which PAM demands.

2.2. SPEECH LEARNING MODEL (SLM)

The model chosen for the current analysis, as stated above, is the Speech Learning Model (SLM), developed by Flege during the nineties. It departs from previous approaches because it presents a continuity in the mechanisms that are used in L1 acquisition and L2 learning later in life (Flege, 1996:42). The model explains the bi-directional interactions between the phonetic systems of L1 and L2, i.e. how they affect each other. These interactions are called assimilation and dissimilation (Flege, 2007:366-367). The former happens if the speaker is not able to create a new phonetic category for the FL sound that is differentiated from the most similar L1 sound. This can take place when the L2 sound is perceptually assimilated as a native category, when it is too similar to it or in both cases. The result of assimilation would be a non-native pronunciation of L2. Dissimilation, on the other hand, happens if speakers are indeed able to create the new phonetic category for the L2 sound. The model was developed in four postulates and seven hypotheses that derived from them (shown in Table I).

These interactions between phonetic systems lead to L2 sounds being classified in terms of the perception of L1 categories; they can be 'new', 'similar' or 'identical' (Flege, 1996:16). An identical L2 sound would be perceived and produced authentically due to the creation of a new category through a process called 'positive transfer' or the use of L1 (Weinreich, 1953, in Flege, 1996:17). A new sound, in case it is not too similar to that of L1, may also be correctly perceived and produced by the creation of a new category. Finally, a similar sound is the one that carries more problems. In this case, a new category is not created due to the L2 sound assimilating to an already existing one in the L1. As a result, the sound is neither perceived nor produced accurately.

Table 1. Postulates and hypotheses forming a speech learning model (SLM) of second language sound acquisition (Flege, 1995:239).

Postulates	
P1	The mechanisms and processes used in learning the L1 sound system, including category formation, remain intact over the life span, and can be applied to L2 learning.
P2	Language-specific aspects of speech sounds are specified in long-term memory representations called phonetic categories.
P3	Phonetic categories established in childhood for L1 sounds evolve over the life span to reflect the properties of all L1 or L2 phones identified as a realization of each category.
P4	Bilinguals strive to maintain contrast between L1 and L2 phonetic categories, which exist in a common phonological space.
Hypotheses	
H1	Sounds in the L1 and L2 are related perceptually to one another at a position-sensitive allophonic level, rather than at a more abstract phonemic level.
H2	A new phonetic category can be established for an L2 sound that differs phonetically from the closest L1 sound if bilinguals discern at least some of the phonetic differences between the L1 and L2 sounds.
H3	The greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned.
H4	The likelihood of phonetic differences between L1 and L2 sounds, and between l2 sounds that are noncontrastive in the L1, being discerned decreases as AOL increases.
H5	Category formation for an L2 sound may be blocked by the mechanism of equivalence classification. When this happens, a single phonetic category will be used to process perceptually linked L1 and 12 sounds (diaphones). Eventually, the diaphones will resemble one another in production.
H6	The phonetic category established for l2 sounds by a bilingual may differ from a monolingual's if: 1) the bilingual's category is "deflected" away from an L1 category to maintain phonetic contrast between categories in a common 11-l2 phonological space; or 2) the bilingual's representation is based on different features, or feature weights, than a monolingual's.
H7	The production of a sound eventually corresponds to the properties represented in its phonetic category representation.

2.3. SPEECH PERCEPTION AND PRODUCTION

The model followed, Speech Learning Model (SLM), is based on the idea that speech production is related to speech perception. The pronunciation of an L2 speaker is the result of the perception they have of that language, which, as already stated, refers to the perceptual system of the L1 (Major, 2001:52; see Gómez Lacabex, 2009 for a review). Even if such link is supported by many studies, it is common to find some that present unconvincing and statistically trivial results, leading some researchers claim that perception and production could dissociate during the beginning of L2 acquisition (Hanulíková, Dediu, Fang, Bašnaková, & Huettif, 2012, in Nagle, 2018:235). However, SLM does not state that every non-native production has a perceptual origin (Flege, 1999:1275). For example, sounds that are more complex to pronounce articulatorily and, at the same time, are uncommon in human languages are more limited by the age of the speaker in regards to articulatory learning.

2.4. ANDALUSIAN SPANISH

The autonomous community of Andalusia is commonly divided into two regions when talking about its phonology: Western Andalusia (Seville, Huelva, Cádiz and Córdoba) and Eastern Andalusia (Granada, Almería, Jaén and Málaga), known in Spanish as *Andalucía occidental* and *Andalucía oriental* (Hualde, 2014; Narbona, Cano, & Morillo, 1998; Torreira, 2012). In this section I will go through the different features that characterise the Spanish spoken not only in these two regions of Andalusia, but also in some areas of Extremadura and of the Region of Murcia. Hualde (2014) gives the name of "Southern Peninsular Spanish" (or *español peninsular meridional*) to this variety (p. 286), but being more popular and relevant, I will use "Andalusian Spanish" as an umbrella term to refer to the different varieties that share some of these features.

2.4.1. Consonantal features of Andalusian Spanish

The first consonantal feature, and probably the most noticeable one (Torreira, 2012:126), is that there is no contrast between /s/ and /θ/. For example, two words like *casa* 'house' and *caza* 'hunting' would be homophones, both pronounced either [kása] or [káθa]. This lack of contrast results in two phenomena: *seseo* (when the allophone is the alveolar) and *ceceo* (when it is the interdental). *Seseo* is one of the main differences between Castilian Spanish (often called European Spanish, or Peninsular Spanish) and Latin American Spanish, since every Spanish-speaking country from America is *seseante*. *Ceceo*, on the other hand, is often seen as a rural feature of Andalusian Spanish and is often stigmatised. In fact, in urban areas such as Seville or Granada, this lack of contrast is a receding feature, i.e. some speakers are starting to make a distinction between /s/ and /θ/. The speakers to be analysed in this paper are from Seville, a city where *seseo* is predominant even though it is surrounded by *ceceante* areas.

There is also a weak pronunciation of /x/, which is usually glottal. A word like *mejor* 'better' is [mehó] or [meɦó]. Notice that the final /r/ is not pronounced, this is due to another characteristic — the loss of lateral and vibrant consonants (liquids) in final position. There also exists a neutralisation or lack of contrast between these two consonants when they appear in front of another consonant. Therefore, when adding the article *el* 'the, masc' to *mejor* it would be [er mehó]. This is another stigmatised feature and is often avoided during formal contexts (Hualde, 2014:287).

Intervocalic /d/ is often dropped in unstressed syllables. This is yet another feature that is common in many Spanish varieties, usually in *-ado*, as in *cansado* [kansáo] or [kansáu] 'tired, masc'. What is special about Andalusian Spanish is the elision of /d/ in other endings such as *-ada* and *-ido*: *cansada* [kansá] 'tired, fem' or *partido* [partío] 'match'.

Some Andalusian Spanish speakers also weaken the voiceless postalveolar affricate, so a word like *chico* 'boy' would be pronounced as [ʃíko]. Hualde (2014) gives this process the name of "deaffrication of [tʃ]" (p. 146), i.e. the plosive part of the affricate is dropped and only the fricative is kept.

Another consonantal feature of Andalusian Spanish, although appearing in many other variations of Spanish, is the aspiration and loss of /s/ before a consonant and in final position. For example, *esta* 'this, fem.' [éhta] or [éta] and *los patos* 'the ducks' [lohpátoh] or [lopáto]. This phonological dialectal variation is one of the most studied

out of those related to the Spanish language (Hualde, 2014:157). In fact, the aspiration of /s/ before a consonant has different outcomes in Andalusian Spanish (Hualde, 2014:160):

- Preaspirated geminate consonants: /s/ + voiceless stop. For example: *hasta* 'until' [á^{h}t:a] or *vasca* 'Basque, fem' [bá^{h}k:a].

- Partially voiceless geminate consonants: /s/ + sonorant. For example: *asno* 'ass' [á$_{̥}$no] or muslo 'thigh' [mú$_{̥}$lo].

- Voiceless geminate consonants: /s/ + [β], [ð] or [ɣ]. For example: *los barcos* 'the ships' [lɔɸ:árkɔ], *los dientes* 'the teeth' [lɔθ:jéntɛ] and *los guantes* 'the gloves' [lɔxwántɛ].

- Metathesis of the aspiration: /s/ + [p], [t] or [k]. For example: *cesta* 'basket' [θét^{h}a], *basto* 'rough, masc' [bát^{h}o] and *pesca* 'fishing' [pék^{h}a]. These clusters are specific of Western Andalusia or *Andalucía occidental*, and they are not found in any other variety of Spanish (Torreira, 2012).

2.4.2. Vocalic features of Andalusian Spanish

The vowels in the varieties of Spanish are, generally speaking, quite stable, maintaining the vowels /i/, /e/, a/, /o/ and /u/ (Hualde, 2014:124). Regarding Andalusian Spanish, there is a variation related to the deletion of /s/ in final position, mentioned above. More specifically, Hualde (2014) states that the vowels /e/ and /o/ have the allophones [ɛ] and [ɔ] before this "silent /-s/" (p. 125). As a result, the plurals of nouns ending in /e/ or /o/ such as *hombres* 'men' or *hombros* 'shoulders' would be pronounced as [ɔ́mβɾɛ] and [ɔ́mβɾɔ]. Notice that this change also extends to the accented vowel (sometimes even to the previous one). This variation, however, is only active in the province of Granada and its surrounding areas, the above-mentioned Eastern Andalusia or *Andalucía oriental*. Again, since this paper is only focused on speakers of Sevilla, that is, Western Andalusia or *Andalucía occidental*, this feature is of no major importance. It is nevertheless important to remember it in case any of the speakers do show a vocalic variation before a final "silent /s/" when speaking English.

3. ACOUSTIC ANALYSIS

3.1. DATA

The recordings analysed in this paper were provided by my supervisor, Dr. Maria Luisa García Lecumberri. The data was gathered at the University of Seville as a continuation to a cross-language comparison of conversational production of English both as first and foreign language by Scottish and Basque speakers, respectively (García Lecumberri, Cooke, & Wester, 2017). In order to prompt a conversation between the speakers (each pair constituted of native speakers of the same language) a picture description task was used, based on Van Engen et al.'s (2010) Diapix. The task consists in a spot-the-difference exercise in which the participants can only see one of the images. In order to find the differences, they are expected to compare the drawings by describing in detail what they see. This oral comparison generates a spontaneous speech centered around one topic (that of the setting of the drawing, e.g. a beach). Moreover, the representation of specific objects will enhance the probability of speakers producing certain words (García Lecumberri, Cooke, & Wester, 2017). In this case, the United Kingdom's version of Diapix (DiapixUK) was adapted as DiapixFL (Baker, & Hazan, 2011, in García Lecumberri, Cooke, & Wester, 2017).

The speakers of the recordings are students in the third year of English Studies at the University of Seville. In fact, as young people (born around the nineties) who have been raised in Spain, it has been deduced that they all have studied English as a foreign language since Primary School. Even though all the speakers from the recordings that I first transcribed had a similar pronunciation and proficiency of the language, I chose two of them for the current analysis: Speaker 3B (S3B) and Speaker 4B (S4B). This decision was based on specific pronunciation errors that both speakers produced and the rest did not.

3.2. METHODOLOGY

First, I transcribed the recordings of three of the four pairs of female speakers orthographically in order to facilitate the searching of specific words. A sound file was created for each of the participants and, thus, a transcription as well. Secondly, after deciding on which speakers I would choose to analyse, I phonetically transcribed the pronunciation errors I had perceived during the previous transcription. Both transcriptions were made using Praat (Boersma, & Weenink, 2019). Finally, I obtained a spectrogram of each of the relevant pronunciation errors by using Praat's picture tool (see the figures of section 3.3). By using images from the computer software, it was easier for me to analyse, classify and, for the purpose of this paper, illustrate the speech of each participant more accurately.

This analysis and eventual classification, as mentioned above, was based on Flege's Speech Learning Model (SLM). Thus, first I divided the FL sounds into “new”, “identical” or “similar” for the case of native speakers of Andalusian Spanish (see Tables 2 & 3). Then, in the results section, I shall introduce some examples of each of these categories, as well as images of some of the utterances' pronunciation, showing both the spectrogram and the transcription with boundaries. I shall also present some examples of mistakes that are unpredicted or unexplained by SLM.

Table 2. Classification of English consonants for speakers of Andalusian Spanish according to SLM.

Categories	**English consonants**
New:	/ʒ /
Identical:	/m/, /n/, /ŋ/**, /θ/, /f/, /l/*, /ʃ/**, /w/**, /j/**, /h/
Similar:	/p/, /b/, /t/, /d/, /k/, /g/, /ð/**, /v/, /s/, /z/, dʒ/, /r/

Table 3. Classification of English consonants for speakers of Andalusian Spanish according to SLM.

Categories	**English vowels**
New:	/ə/, /ɜ/
Identical:	/i/*, /e/*, /u/*
Similar:	/ɪ/, /æ/, /ɑ/, /ɒ/, /ɔ/, /ʊ/, /ʌ/

* Semi-identical.

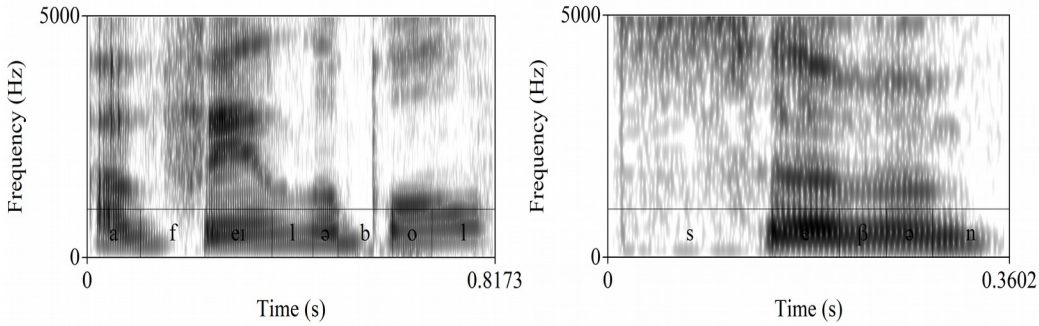
** In Spanish, these consonants are allophones, and not phonemes.

3. RESULTS

3.3.1. Consonants

The first example of similar sounds is that of /v/ and /f/, which S4B mistakens in, for example, the word *available*. The voiced labiodental fricative is not one of the nineteen consonants that constitute the Spanish language (Gómez González, & Sánchez Roura, 2016:38), therefore being usually perceived as similar sounds /b/ or /f/. S3B, in fact, has perceived this consonant as a word-final voiceless fricative in the word *five* and, as an voiced intervocalic bilabial fricative in the word *seven*, producing it as a [β] (see Figure 1). This former mistake might have happened because, in Spanish, the consonant [β] is an allophone of /b/ in all positions except after a pause or a nasal, as in *bebé* 'baby' [beβé] (Hualde, 2014:130). This consonant, classified as "new" above, was thus not assimilated as such by neither of the speaker.

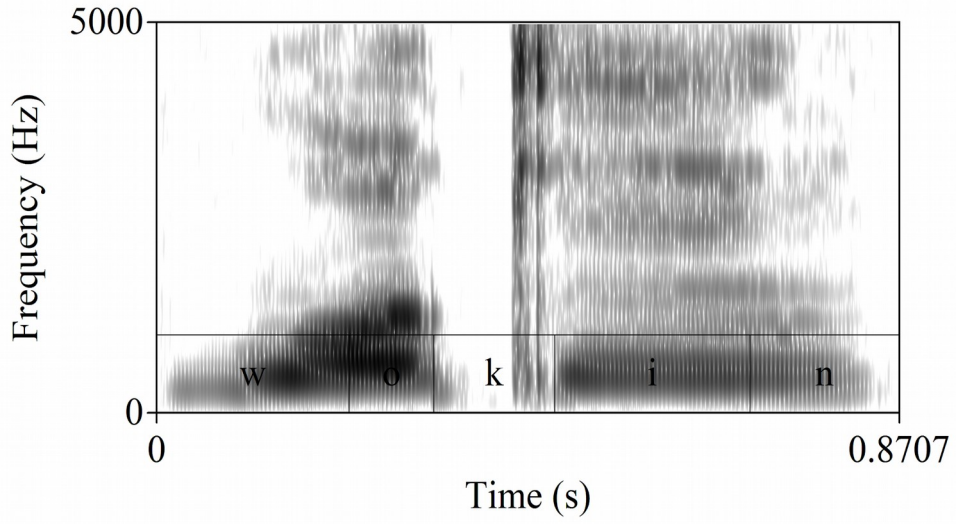
Figure 1. Utterance of the word *available* pronounced as [aféɪləbol] by S3B (left) and utterance of the word *seven* pronounced as [séβən] by S4B (right).



Another consonant that has not been produced correctly due to its similarity and, therefore, might have been predicted by SLM, is /z/. As will be seen below in the spectrogram to the right from Figure 3 and in the one from Figure 4, /z/ is pronounced as [s] both in *is* and *has*. In fact, as with /v/, I could not found any utterance of this voiced consonant, which the Spanish phonetic inventory does not contain (Gómez González, & Sánchez Roura, 2016:38).

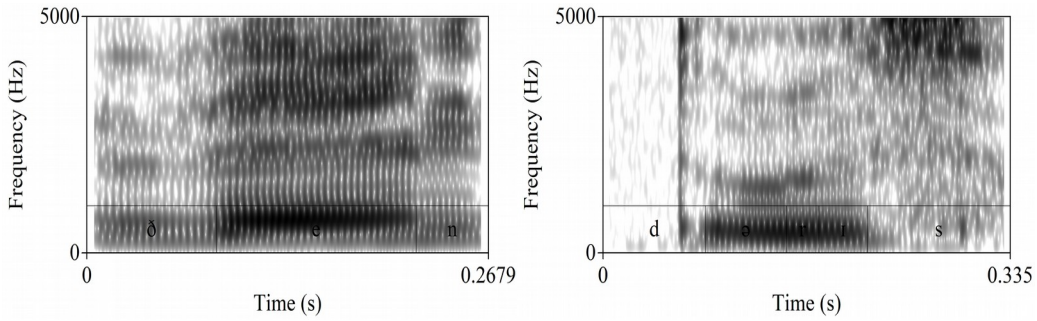
The third example I could find is that of /ŋ/ and /n/, identical consonants. Both speakers place the alveolar in word-ending positions in which the velar should be, e.g. *young*, *watching*, *walking*, etc. (see Figure 2). Similar to the first example seen above, [ŋ] is an allophone of Spanish /n/. But the allophone does appear both in coda position and inside of the word (Hualde, 2014:173). For example *manga* 'sleeve' [máŋga] or *tango* [táŋgo].

Figure 2. Utterance of the word *walking* pronounced as [wókin] by S4B.



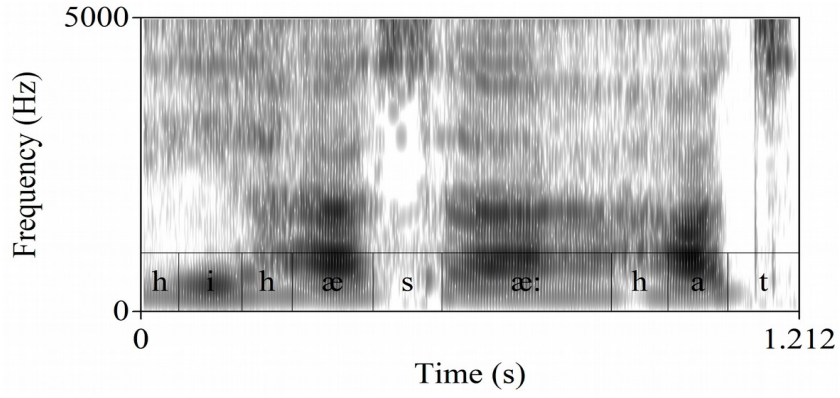
The next example is the voiced interdental fricative /ð/. This consonant appears in Spanish, more frequently as an approximant, as an allophone of /d/ in cases in which it is not positioned after a pause, a nasal or a lateral (Hualde, 2014:130), e.g. the word *dado*s 'dice' [dáðos]. Thus, a new consonantal category might have been established as dictated by the English language in common words such as *then* and *the* (S3B). In these cases, it has been noted that the previous sounds were either a nasal, a lateral or that there was a pause, therefore not being the distribution like in the L1. S4B, on the other hand, has perceived the consonant as being similar to /d/, thus producing it as such in, for example, the phrase "there is" (see Figure 3).

Figure 3. Utterance of the word *then* pronounced as [ðen] by S3B (left) and utterance of the phrase *there is* pronounced as [dər ɪs] by S4B (right).



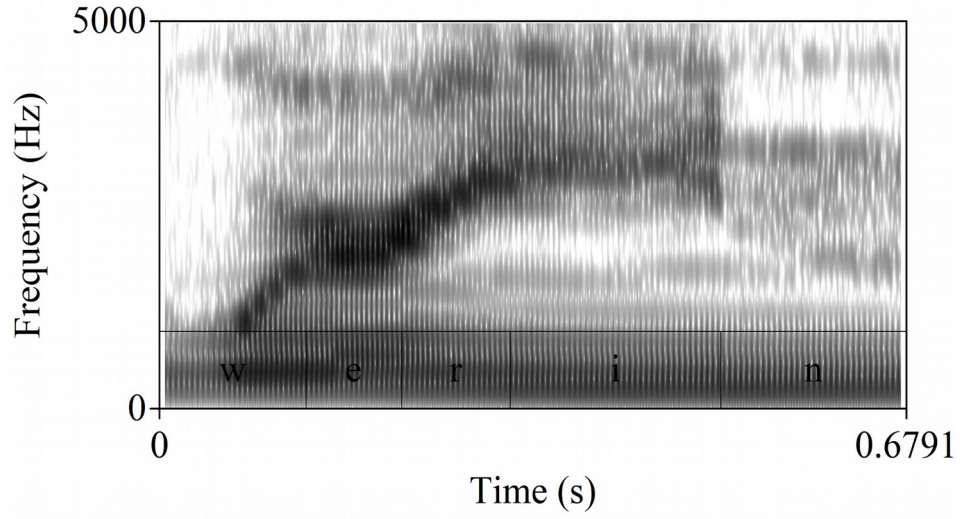
The only case of an identical consonant that I find worth mentioning is the voiceless glottal fricative /h/ which does not belong to the phonetic system of PSp. As a result, it is usually pronounced as /x/ (García Lecumberri, 2000, in Gómez González, & Sánchez Roura, 2016:203). However, in Andalusian Spanish, as mentioned above, [h] appears in every instance of what would be /x/ in PSp. Thus, both S3B and S4B have correctly perceived and, consequently, produced this consonant in every utterance of it (see Figure 4).

Figure 4. Utterance of the sentence *he has a hat* pronounced as [hi hæs æ: hat] by S4B.



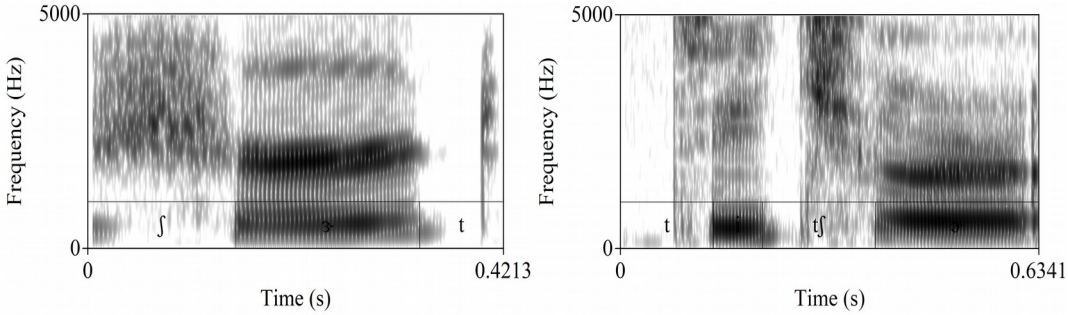
The establishment of a phonetic category for the consonant /w/ might be a clear example of what SLM would not predict. This sound does belong to Spanish in specific diphthongs such as [wa], [we], [wo] and [wi], e.g. *agua* 'water' [áɣwa], *huevo* 'egg' [wéβo] and *tuétano* 'marrow' [twétano]. It also appears in two triphthongs: [we̯i̯] and [wa̯i̯]. For example, *buey* 'ox' [bwe̯i̯] and *guay* 'great, cool' [gwa̯i̯]. The unpredictable aspect is that both consonants are articulately identical, but phonotactically different. Moreover, and unlike English, [w] is an allophone of /u/ in Spanish. In fact, some do not even consider it a Spanish phoneme per se (Gómez González, & Sánchez Roura, 2016:226). Some researchers even decide not to use /w/ as a phone to transcribe Spanish, using the non-syllabic form of /u/ instead, i.e. [u̯] (Hualde, 2014:41). Even so, both S3B and S4B produce the consonant /w/ as defined by English, probably due to their academic knowledge of the language's pronunciation. Native Spanish speakers of English usually insert [g] or [ɣ] before the /w/ (Gómez González, & Sánchez Roura, 2016:226), so knowing its proper pronunciation is elemental for university students of the language. Notice in Figure 5 how the second formant of the consonant rises, very similar to the movements away of the formants for short [u], a very characteristic feature of /w/ (Ladefoged, & Johnson, 2011:203).

Figure 5. Utterance of the word *wearing* pronounced as [wérin] by S3B.



Another example of the consonants that are neither explained nor predicted by SLM is the voiceless postalveolar fricative /ʃ/ pronounced as /tʃ/ in some words like *t-shirt*, *sheep*, *shoot* or *shark*. Either way, in some cases the pronunciation was correct (more times in the case of speaker 3B), both when pronouncing those same words or more common ones like *she* or *shop*. PSp does not include this consonant (Gómez González, & Sánchez Roura, 2016:38) but, as mentioned above, Andalusian Spanish might (that is why it is classified as "identical" in Table 2). This is due to the weakening of the voiceless postalveolar affricate, i.e. [tʃ] → [ʃ]. Neither speaker pronounces the word *chico* 'boy' [tʃíko] with the fricative [ʃíko], but they might well do so in more informal contexts. A way of explaining this mistake is the fact that speakers could be hyper-correcting their pronunciation of [ʃ] as [tʃ], just like they might be doing in Spanish with *chico*. Thus, we can say that hyper-correction is not explained nor predicted by SLM. In the spectrograms (Figure 6), although not being the same exact word, we can see that S3B pronounced the word 'shirt' correctly, while S4B introduces a voiceless alveolar stop /t/ after the first vowel (the vertical white area would be the obstruction of air before the burst of the stop, just like at the beginning of the word).

Figure 6. Utterance of the word *shirt* pronounced as [ʃɜ:t] by S3B (left) and utterance of the word *t-shirt* pronounced as [títʃɚ] by S4B (right).



Other examples of phenomena unexplained by SLM include the elision of different consonants:

1. /l/: *blue* [bu:] (S3B). Although this pronunciation mistake was one of the reasons I chose this speaker, I was not able to find a reason for its production. The initial consonant cluster /bl/ appears in both Spanish and English (Gómez González, & Sánchez Roura, 2016:21), and, moreover, the speaker does pronounce it correctly when speaking her native language in, for example, the word *blanca* 'white, fem.' [bláŋka]. It is also worth noting that she only makes the mistake with the word *blue*, and, on top of that, she does not make it in every utterance of the word.

2. final /d/ and /t/ after alveolars (/n/, /s/, /r/, /l/): *clouds* [klaʊ̯s], *called* [kɔ:l] or *shirt* [ʃɝ] (S3B and S4B). This feature might be related to the fact that the speakers' L1 is Andalusian Spanish, in which the elision or modification of final consonants is quite common, as seen in the characteristics above. Furthermore, these final clusters are not part of the Spanish phonetic inventory (Gómez González, & Sánchez Roura, 2016:22-23).

3. final /v/: *five* [faɪ] (S4B). As seen with S3B, this mistake might be due to the voiced fricative being perceived as a voiceless fricative /f/. In Spanish this consonant does not appear in final position (except in loanwords such as *golf* or *surf*). Thus the reason for this mistake, taking into account that in Andalusian Spanish final fricatives are usually dropped (as seen above with /s/).

4. final /k/: *like* [laɪ] (S4B). Same reason as the second example (although /k/ appears in Spanish, it never does so in final position).

5. final /s/: *tennis balls* [téni bɔ:l] (S3B). As already mentioned, one of the most characteristic feature of Andalusian Spanish (after *seseo* and *ceceo*) is the loss of /s/ in final position. This might have affected the production of only this compound noun, given the fact that the rest of final /s/ utterances by S3B are correct.

The following table summarises the results regarding consonants (see Table 4).

Table 4. Summary of results regarding consonants.

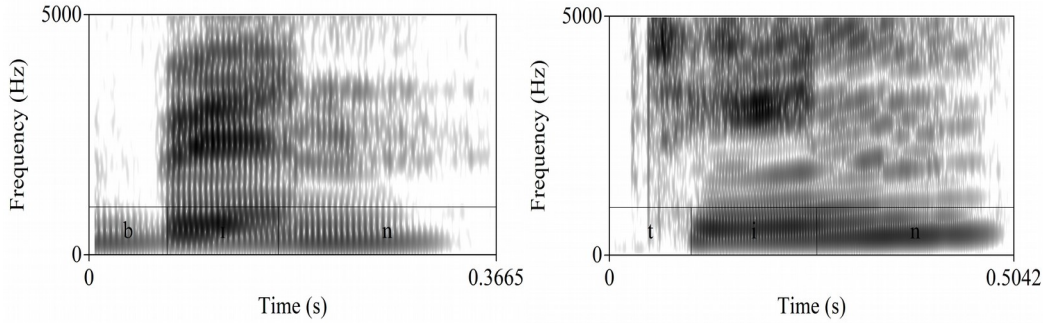
SLM classification of consonants		**Results S3B**	**Results S4B**
New:	∅		
Identical:	/h/	*he has a hat* [**h**i **h**æs æ: **h**at]	*how* [**h**aʊ̯]
	/ʃ/	*shark* [**ʃ**ark]	*she* [**ʃ**i]
	/ŋ/	*fishing* [fɪʃɪ**n**]	*walking* [wóki**n**]
Similar:	/v/	*five* [fai**f**] ∅	*available* [a**f**éɪləbol] *seven* [sé**β**ən]
	/ð/		*there is* [**d**ər ɪs]
	/z/	*has* [hæ**s**]	*is* [ɪ**s**]

Unpredicted results	**Results S3B**	**Results S4B**
/w/	*wearing* [**w**érin]	*with* [**w**iθ]
/ð/	*then* [**ð**en]	
/ʃ/ → /tʃ/	*shoes* [**tʃ**u:s]	*t-shirt* [tí**tʃ**ɚ]
/l/ → Ø	only in *blue* [bu:]	
/d/ → Ø /t/ → Ø	*clouds* [klaʊ̯s] *shirt* [ʃɝ]	*bird* [ber]
/v/ → Ø		*five* [faɪ]
/k/ → Ø		*like* [laɪ]
/s/ → Ø	*tennis balls* [téni bɔ:l]	

3.3.2. Vowels

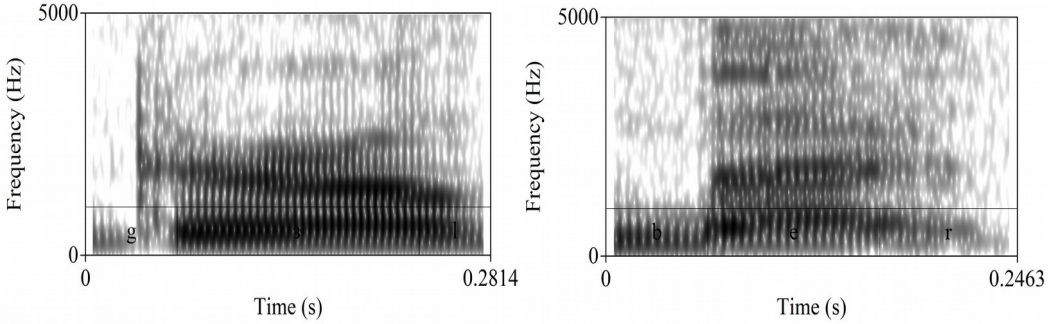
The first vocalic category that I took into consideration was that of /ɪ/. The Spanish language does not have this vowel in its inventory, often being confused with /i/. Both this fact and the English spelling of <i> as these two vowels lead to this erroneous assimilation (Escudero (2000), in Gómez González, & Sánchez Roura, 2016:95). Nevertheless, S3B accurately produces the vowel /ɪ/ in almost every instance of it, more noticeable in the words *bills*, *skin* and *bin*. The most obvious exception would be in the word *tin*, where she does pronounce the vowel as /i/ (see Figure 7). S4B, on the other hand, might have perceived the vowel as similar to /i/, and, therefore, produces it as such (the only exception I could find was in the word *pink*, where she does indeed produce the /ɪ/).

Figure 7. Utterance of the word *bin* pronounced as [bɪn] by S3B (left) and utterance of the word *tin* pronounced as [tin] by S3B (right).



The other two examples I found are very similar to that last one, both their reasons and the role of the speakers. The first one is related to the English diphthong /əʊ/, pronounced correctly by S3B in most cases (like in the word *go*) but more similar to /oʊ̯/ in some others (as produced by S4B). It could be said neither of these diphthongs are part of the Spanish phonetic inventory (even so, /oʊ̯/ does appear in Spanish words of Catalan or Galician/Portuguese origin (Hualde, 2014:66)) but, as already mentioned, vowels /o/ and /u/ are, unlike /ə/ and /ʊ/. The second and final example of a new category being established by S3B and not by S4B is the vowel /ɜ/. The former produces it accurately, e.g. in *girl*, although she does pronounce it as /e/ in other words such as *purple* or, in the case of S4B, *bird* and *skirt* (see Figure 8).

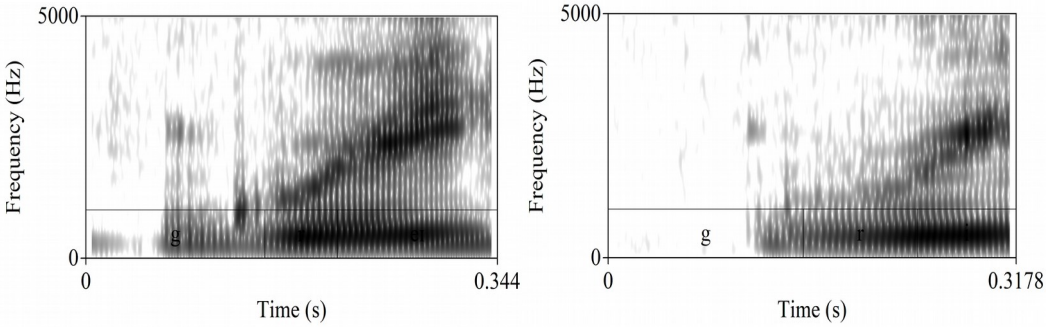
Figure 8. Utterance of the word *girl* pronounced as [gɜ-l] by S3B (left) and utterance of the word *bird* pronounced as [ber] by S4B (right).



In regards to identical vowels, it can be said that Spanish and English do share the vowels /i/, /e/ and /u/, even though in RP the close vowel is more centralised and the other two are more open and closer (Gómez González, & Sánchez Roura, 2016:89). Similarly, diphthongs /ei̯/, /ai̯/, /oi̯/ and /au̯/ could also be considered counterparts of English /eɪ/, /aɪ/, /ɔɪ/ and /aʊ/ (Gómez González, & Sánchez Roura, 2016:133). Both speakers produce short vowels instead of /i:/ and /u:/ in the majority of the cases, most probably because length in Spanish vowels lacks phonemic status. Therefore, the one vowel that we could identify as being identical in both languages is probably /e/, produced as such by both S3B and S4B in words like *next*, *seven* or *dress*. Even so, it might be interesting to note that the quality of this vowel is not the same in both languages; in fact, Gómez González and Sánchez Roura (2016) consider RP English /e/ to be more similar to Catalan and Galician e-sound (p. 102).

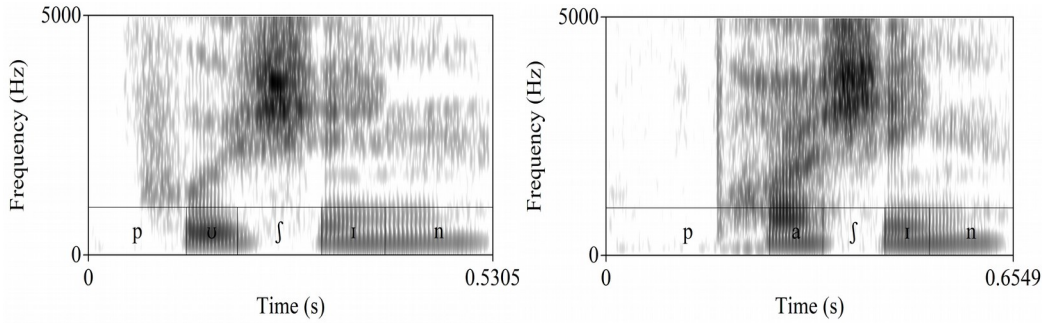
There is a mistake regarding the diphthong /eɪ/, produced mostly correctly by both speakers, that could not have been predicted by SLM. Happening only in one occasion, S3B pronounces the word *grey* as [gri] (see Figure 9). The most probable reason for this is a spelling pronunciation, as English <ey> can be pronounced as both /eɪ/ and /i:/ (compare the words *obey* and *key*). Let us not forget that the usual American spelling of the word *grey* is *gray* but, since /i:/ written as <ay> is a much more rare pronunciation (Gómez González and Sánchez Roura, 2016:93), it is more likely to be the former case.

Figure 9. Utterance of the word *grey* pronounced as [greɪ] by S3B (left) and utterance of the word *grey* pronounced as [gri] by S3B (right).



The next example, this time in relation to similar categories, is perceiving and, thus, producing, English /ʌ/ as a Spanish /a/. Both speakers do so in words such as *fun*, *truck*, *ducks* and *nothing*. What is interesting about this mistake is how it affects another one that would not have been predicted nor explained by SLM. The pronunciation error is simple: both S3B and S4B produce an /a/ in words like *butchers*, *bush* and *push*, i.e. those in which stressed <u> is pronounced /ʊ/. Yet again, the reason for this mistake is a spelling pronunciation. When learning English vowels, one common rule to identify the vowel /ʌ/ (non-existent in Spanish) is relating it to the vowel <u> in words like *cut*, *but* and *bus* (common monosyllabic words with CVC pattern). This "rule" leads to similar yet less common words being pronounced with /ʌ/ when, in fact, they are to be done so with /ʊ/, just like the examples above. Therefore, the speakers' lack of the establishment of a category for /ʌ/ has resulted in /ʊ/ being pronounced as /a/ (see Figure 10).

Figure 10. Utterance of the word *pushing* pronounced as [pʊʃɪn] (left) and [páʃɪn] (right) by S3B.



The following table summarises the results regarding vowels (see Table 5).

Table 5. Summary of results regarding vowels.

SLM classification of vowels		**Results S3B**	**Results S4B**
New:	/ɜ/	*girl* [gɜ˞l]	*bird* [b**e**r]
Identical:	/e/	*next* [n**e**kst]	*seven* [s**é**βən]
Similar:	/ɪ/	*bills* [b**ɪ**lls] *bin* [b**ɪ**n]	*pin* [p**i**n]
	/ʌ/	*fun* [f**a**n] *truck* [tr**a**k]	*nothing* [n**á**θin]
	/əʊ/	*go* [g**əʊ**]	*so* [s**oʊ̯**] *no* [n**oʊ̯**]

Unpredicted results	**Results S3B**	**Results S4B**
/eɪ/ → /i/	only once in *grey* [gr**i**]	
/ʊ/ → /a/	*push* [p**a**ʃ] *pushing* [p**á**ʃin]	*butchers* [b**á**tʃərs] *bush* [b**a**ʃ]

4. CONCLUSION AND DISCUSSION

The aim of this paper was to determine whether the Speech Learning Model could explain and predict production mistakes by classifying them using spectrograms. As has been seen in the last section, we can say that several errors cannot be explained by only applying SLM hypotheses. Some of those production mistakes were made due to hyper-corrections and spelling pronunciations, as well as specific or unique features of Andalusian Spanish. In fact, these results have led me to the conclusion that the production of S3B is more native-like than that of S4B, even though not being the main objective of the current analysis.

I would like to point out the fact that the acoustic analysis was chosen as a method after reading the Final Year Thesis of Maddi Morcillo (2015), whose topic was very similar to that of this paper (in her case the speakers were from the Basque Country). In her discussion section, Morcillo mentions how an acoustic analysis would help with the classification task, since the perception of a non-native might not be accurate enough (p. 20). In fact, with the help of spectrograms, it was indeed easier for me to prove that the sounds that I was classifying were those that I had previously perceived. A clear example of this is how, at first, I had perceived a correct /v/ in the word *available*, classifying it as a "new" category for S3B. But, after checking the spectrogram, I realised that it was actually a voiceless /f/, thus changing it to "similar" (see Figure 4). For this reason, I can say that the idea of an acoustic analysis using spectrograms was a resourceful one. Either way, for future research, a more detailed research might be needed, e.g. analysing VOT, specific formats, etc.

Related to further analysis, it might also be interesting to collect data from regions such as Galicia and Catalonia. Several results have hinted at the fact that Galician and Catalan speakers might actually perceive some sounds in a way that approximates native English speakers. In the Iberian Peninsula, the collection of DiapixFL data has already been done in regions such as the Basque Country and Seville, so it would be of no wonder if Galician and Catalan corpora were to be collected.

5. REFERENCES

Abello-Contesse, C. (2009), Age and the critical period hypothesis, *ELT Journal 63*(2) 170–172. Retrieved from https://academic.oup.com/eltj/article/63/2/170/441108

Baker, R., & Hazan, V. (2011). "DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs", *Behaviour Research Methods 43*(3), 761-770.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: theoretical and methodological issues*, (pp. 171-204). York Timonium, MD: York Press. Retrieved from https://www.researchgate.net/publication/229078334_A_direct_realist_view_of_crosslanugage_speech_perception

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: commonalities and complementarities. In O. -S. Bohn, & M. J. Munro, (Eds.), *Language Experience in Second Language Speech Learning: in honor of James Emil Flege*. (pp. 13-34). Amsterdam/Philadelphia: John Benjamins Publishing Company. Retrieved from https://www.researchgate.net/publication/258209938_Nonnative_and_second-language_speech_perception_Commonalities_and_complementarities

Boersma, P. & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.0.46, retrieved 3 January 2019 from http://www.fon.hum.uva.nl/praat/

Eckman, F. R. (2004). From Phonemic Differences to Constraint Rankings: Research on Second Language Phonology. *Studies in Second Language Acquisition 26*(4) 513-549.

Escudero, P. (2000). *Developmental patterns in the adult L2 acquisition of new contrasts: The acoustic cue weighting in the perception of Scottish tense/lax vowels by Spanish speakers* (MSc dissertation). University of Edinburgh, Scotland.

Flege, J. E. (1995). Second-language speech learning: theory, findings and problems. In W. Strange (ed.), *Speech Perception and Linguistic Experience. Issues in Cross-Language Research* (pp. 233-277). Timonium, MD: York Press.

Flege, J. E. (1996) English vowel production by Dutch talkers: More evidence for the "new" vs "similar" distinction". In A. James & J. Leather (Eds.), *Second Language Speech: Structure and Process* (pp. 11-52). Berlin: Walter de Gruyter.

Flege, J. E. (1999), "The relation between L2 production and perception". In J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville & A. C. Bailey (Eds.), *Proceedings of the XIVth International Congress of Phonetic Sciences*, (pp. 1273-1276). San Francisco: University of California.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole & J. I. Hualde (ed.), *Laboratory phonology 9* (pp. 352-381). Berlin/New York: Mouton de Gruyter.

Freed, B. (1995). Language Learning and Study Abroad. In B. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 3-33). Philadelphia: John Benjamins.

García Lecumberri, M. L. (2000). La pronunciación del inglés como lengua extranjera en escolares de diferentes edades. Paper presented at the Curso de Verano de la Universidad del País Vasco *El factor edad en la adquisición de lenguas extranjeras*. San Sebastián-Donostia (Spain), 10-12 July, 2000.

García Lecumberri, M. L., Cooke, M., & Wester, M. (2017). A bi-directional task-based corpus of learners' conversational speech. *International Journal of Learner*

Corpus Research 3(2) 175-195. DOI: http://dx.doi.org/10.1075/ijlcr.3.2.04gar

Gómez González, M. A., & Sánchez Roura, T. S. (2016). *English Pronunciation for Speakers of Spanish: From Theory to Practice.* Boston/Berlin: Walter de Gruyter.

Gómez Lacabex, E. (2009). *Percepcional and production of vowel reduction in second language acquisition.* University of the Basque Country: Leioa, Biscay.

Hanulíková, A., Dediu, D., Fang, Z., Bašnaková, J., & Huettig, F. (2012). Individual differences in the acquisition of a complex L2 phonology: A training study. *Language Learning, 62,* 79–109. 79–109. https://doi.org/10.1111/j.1467-9922.2012.00707.x

Hualde, J. I., (2014). *Los sonidos del español*. Cambridge: Cambridge University Press.

Key: The Cambridge School Dictionaries: Nivel inicial e intermedio. (2004). Cambridge University Press.

Kuhl, P. K. (1993). Early linguistic experience and phonetic perception: Implication for theories of developmental speech perception. *Journal of Phonetics, 21*, 125-139.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2007). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 363*(1493), 979–1000. http://dx.doi.org/10.1098/rstb.2007.2154

Ladefoged, P., & Johnson, K. (2011). *A Course in Phonetics* (6th ed.). Wadsworth: Cengage Learning.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.

Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York: Wiley & Sons.

Major, R. C. (1986). The Ontogeny Model: Evidence from L2 Acquisition of Spanish *r*.

Language Learning 36(4) 452-504.

Major, R. C. (2001). *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*. New Jersey: Lawrence Erlbaum.

Morcillo Berrueta, M. (2015). *Analysis of EFL speech production corpora according to the Speech Learning Model*. Retrieved from the University of the Basque Country's ADDI service: https://addi.ehu.es/handle/10810/21196

Nagle, C. L. (2018). Examining the Temporal Structure of the Perception-Production Link in Second Language Acquisition: A Longitudinal Study. *Language Learning 68*(1) 234-270.

Narbona, A., Cano R., & Morillo R. (1998). *El español hablado en Andalucía*. Barcelona: Ariel.

Richards, J. C., & Schmidt, R. (2010). *Longman Dictionary of Language Teaching and Applied Linguistics* (4th ed.). Harlow, England: Longman.

Stockwell, R. P., & Bowen, J. D. (1965). *The sounds of English and Spanish*. Chicago: University of Chicago Press.

Torreira, F. (2012). Investigating the nature of aspirated stops in Western Andalusian Spanish. *Journal of the International Phonetic Association 42*(1) 49–63.

Van Engen, K. J., Baese-Berk, M., Baker, A., Choi, R. E., Kim, M., & Bradlow, A. R. 2010. "The Wildcat Corpus of native-and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles", *Language and Speech 53*, 510–540.

Weinreich, U. (1953). *Languages in contact, findings and problems*. New York: Linguistic Circle of New York.