

Chapter 10

Phonic adaptation of Spanish Anglicisms in Mexico and Spain: A corpus data analysis

Linda Bäumler

University of Vienna

In this chapter, I discuss the phonic adaptation of Anglicisms by Spanish natives from Mexico and Spain. Since phonemes of the source language (in this case, English) which do not exist in the recipient language (in this case, Spanish) are of special interest in loanword phonology, I based my study on the realization of the English voiced alveopalatal affricate /dʒ/. Previous research describes mainly two different ways of pronouncing this foreign phoneme: the realization of grapheme-phoneme correspondence as in *manager* ['manaxer] on the one hand, and the adaptation of the foreign phoneme as [j] as in ['manajer] on the other. However, the comparative oral corpus data analyzed in this study also reveals many cases of phoneme importation, corresponding to pronouncing [dʒ]. Moreover, the data suggest that regarding the perception process, the most striking feature of the foreign affricate is [+VOICE], since more speakers realize a voiced palatal sound than an unvoiced postalveolar affricate when imitating the foreign sound. In addition, statistical analysis using a generalized linear mixed-effects model revealed that, at least in this data set, a speaker's country of origin does not affect the pronunciation of the loanword. This suggests that in a globalized world, geographic proximity to the US might not be a relevant factor influencing the realization of the phoneme. However, the level of exposure and affinity a speaker has to the English language and American culture is the most relevant factor actually influencing the pronunciation of Anglicisms in this study.



1 Introduction

1.1 Loanword phonology in Spanish-English language contact

For loanword phonology, those phonemes of the source language that do not exist in the phonological system of the recipient language are of special interest. Since Spanish incorporates an increasing amount of English loanwords due to globalization (Cabanillas & Martínez 2012: 95; Oncins-Martínez 2009: 116; Gómez Capuz 2001: 16; Schweickard 1991), this provides us with vast material to examine how Spanish natives pronounce these unknown phonemes in the respective Anglicisms. In general, speakers tend to either adapt the foreign phoneme to the Spanish phoneme system or to imitate the English model.

In the case of imitation, the literature describes several different possibilities: First, the unknown English sound can be replaced by a native Spanish sound that is close to the foreign sound (Kang 2011: 8; Gómez Capuz 2001: 27, 52; Pratt 1980: 155). An example of this would be the realization of *brunch* as [brantʃ̞] instead of [bɾantʃ̞]. Second, sounds of the source language (English) that exist in the phonological system of the recipient language (Spanish), but only in a specific phonological environment, are used in new phonological environments. Haugen (1950: 216) refers to this phenomenon as *phonemic redistribution*. An example of this is the realization of the consonant [tʃ] in word-final position, as in the Anglicism *sandwich*. Note that in Standard Spanish, the affricate is not used in word-final position (Gómez Capuz 2001: 49). Third, the foreign sound is imported to the borrowing language, a process that Haugen (1950: 216) refers to as *phonemic importation* (e.g., realization of [ʌ] in *brunch*). For Spanish, Gómez Capuz (2001: 52) states that even though the influence of English is constantly rising, phonemes are at present not imported and speakers would prefer the first two ways of imitation. However, a prominent example of an English phoneme that got imported is the fricative [ʃ], as in *show*. On the other hand, if speakers tend not to imitate the foreign English sound, they might realize Spanish grapheme-phoneme correspondences, and therefore realize *brunch* as [bruntʃ̞] rather than [bɾantʃ̞], as in English, for instance.

In the literature, different factors that influence the pronunciation of loanwords are discussed. Various authors argue that the age of the loanword affects its pronunciation (Rodríguez González 2017: 127; Haspelmath 2009: 42; Gómez Capuz 2001: 16; Haugen 1950: 216). Imitation of the foreign sound is more likely to occur when speakers are pronouncing younger Anglicisms than when pronouncing older ones. In addition, more frequent loanwords are more likely to undergo phonic adaptation (Rodríguez González 2018: 354). Moreover, the way

in which loanwords enter the language has been mentioned as a relevant factor in this process. Pratt (1980: 16) distinguishes between *eye loans*, that is, loanwords that entered the recipient language through the written channel, and *ear loans*, which were introduced orally into the recipient language. In the former case, the graphic and therefore visual code plays a major role, while in the latter, perception influences the integration of the loanword. Furthermore, it has been suggested that speakers' competence in English and their level of education affect the pronunciation of Anglicisms (Rodríguez González 2017: 102, 123, 127; Gómez Capuz 2001: 42). Speakers with good language skills and/or a high level of education are claimed to be more likely to imitate the foreign English sound. Moreover, the high prestige of the English language might prompt the imitation of the latter (Pustka 2021: 508; Gómez Capuz 2001: 42; Meisenburg 1992: 51). In addition, due to geographic proximity to the US, a higher impact of English on Hispanoamerican varieties of Spanish than on those of Spain has been suggested (Pustka 2021: 504; Rodríguez González 2017: 127; Oncins-Martínez 2009: 116).

To identify and analyze phonic adaptation of loanwords and the respective variables that influence the process, huge oral corpora¹ are needed. Nevertheless, Spanish Anglicisms have mostly been studied by means of written corpora (Rodríguez González 2017: 100; Gómez Capuz 2001: 6). Therefore, the pronunciation of Anglicisms has only occasionally been the subject of studies. Pratt (1980) works among other things with television recordings, but discusses the phonic realization of Anglicisms only briefly. Lorenzo (1997) examines the integration of vowels but analyzes mainly written data. Gómez Capuz (2001) analyzes the phonic assimilation of Anglicisms by means of the oral corpus Val.Es.Co (Briz & Grupo Val.Es.Co. 1995) and provides the reader with Anglicisms used in spontaneous speech. Although this gives insights into many different observable realizations, the corpus does not allow for systematic comparison between many speakers. LaCharité & Paradis (2005) analyze English and French loanwords in different oral corpora, among which is also Mexican Spanish. Data were gathered from written documents and spontaneous speech, and the pronunciations of the loanwords were then verified with a minimum of three L1 speakers (LaCharité & Paradis 2005: 228). Even though their corpus contains a huge amount of loanwords, it lacks realizations of many different speakers that could be compared. Rodríguez González (2017) analyzes the adaptation of consonants and vowels that are unknown to the Spanish phoneme system and bases his analysis on his personal observations of the spoken language during different decades, especially on the occasion of the preparation of dictionaries of Anglicisms (Rodríguez

¹For the definition of the term *corpus* used by the author, see Section 2.1.1.

González 2017: 100). At the time of writing, to the best of my knowledge, no large-scale corpus-phonological data set exists for Spanish Anglicisms. However, this is indispensable to investigate the pronunciation of Anglicisms systematically. Such corpus data not only allow for detection of different realizations and comparative studies, but also for revealing the different social and linguistic factors that influence the pronunciation of the loanwords.

1.2 The English phoneme /d͡ʒ/

When comparing the phonological systems of English and Spanish, many consonants that do not exist in Spanish but do in English stand out. Among these is the English voiced alveopalatal affricate /d͡ʒ/ that appears in words such as *jersey* and *jazz*. Standard Spanish² does not possess an alveopalatal affricate that would be voiced. Instead, it has the voiceless variant that appears for example in *chico* [ˈt͡ʃiko] ‘small boy’ and shares the same place and manner of articulation, but unlike /d͡ʒ/ is not voiced.

The literature attests the adaptation of the English voiced affricate as [j] by Spanish natives in Spain (Rodríguez González 2017: 107; Gómez Capuz 2001: 7).³ Gómez Capuz (2001: 29) assumes that when speakers perceive the loanword, the sonority of the phoneme plays a major role, and Spanish natives therefore tend to realize [j] instead of the voiceless affricate [t͡ʃ], which except for [+VOICE], shares the same features as [d͡ʒ]. This is sometimes also reflected in the spelling of the respective loanwords: *yins* (< jeans), *yip* (< jeep) or *yonqui* (< junkie) (Rodríguez González 2017: 107).

Generally, Rodríguez González (2017: 107) states that mostly speakers who are not familiar with English realize the Spanish grapheme-phoneme correspondence and therefore pronounce [x] instead of [j], as in *jeans* [xeans] (versus [jins]), *manager* [maˈnaxer] (versus [ˈmanajer]), or *challenger* [t͡ʃalenxer] (versus [t͡ʃalenjer]). Gómez Capuz (2001: 21) claims that the higher the age of the loanword, the more likely [x] and therefore grapheme-phoneme correspondence is realized.

²Note that the voiced alveopalatal affricate is known in certain Spanish varieties – in Paraguay, some areas of Ecuador, and Argentina, in the realization of <ll>, <y> or <(h)i> (Real Academia Española & Asociación de Academias de la Lengua Española 2011: 226).

³Note that Gómez Capuz (2001: 29) refers to a fricative and transcribes the allophone by means of the *Alfabeto de la Revista de Filología Española* with the symbol [y]. Rodríguez González (2017: 107) also transcribes [y]. However, I decided in this work to transcribe [j] and refer to an approximant realization, since previous research revealed that in Spanish, realization of [j] is far more frequent than [j̞] (Martínez Celdrán 2015).

In some cases, the realization is also found as [g], for example ['bangi] for *bungee* (Rodríguez González 2017: 108). Rodríguez González (2017: 107) documents cases where the English phoneme is realized as the voiceless variant [tʃ] in intermediate or word-final position (as in *pidgin*, *backstage*, *porridge*). Additionally, Cassano (1973: 211–213) documents one case each of [dʒ] in New Mexico and Argentina in the initial position in *ginger-beer*. In general, he claims that the development of the voiced alveopalatal affricate is independent of foreign influence and characteristic of several Hispanoamerican varieties of Spanish. He supposes that English loanwords might play a reinforcing role in the integration of the voiced alveopalatal affricate.

This study aims to investigate the different realizations of the phoneme in question and to analyze the factors that influence whether speakers imitate the English phoneme. For this study, it was of particular interest to know whether geographical proximity is still a relevant factor in the realization of Anglicisms, as suggested previously (Pustka 2021: 504; Rodríguez González 2017: 127; Oncins-Martínez 2009: 116). However, I question if through the interplay of modern communication tools (e.g., the Internet and social media) and American cultural imperialism (Gray 2007, Hamm & Smandych 2005) – the worldwide diffusion of culture, trademarks, values, media, and language from the US – the effect of greater geographic proximity might no longer play a crucial role.

2 Methodology

In the following section, I will describe the applied methodology: the corpus that served as database and the analysis I carried out, namely the audiovisual coding, the calculation of interrater reliability, and conducting the statistical analysis.

2.1 Corpus

2.1.1 Participants

For this study, the corpus designed for the dissertation project of the author (Bäumler 2023) served as the database. In the following, I will refer to the corpus as the *English Influence on Spanish (EIS)* corpus. Note that corpus phonology traditionally includes naturalistic and experimental data (Eychenne 2021: 324, Chaudron 2007: 736). The definition of *corpus* applied in this study is therefore not limited to spontaneous speech, since experimental data, such as reading a word list, are needed to enable comparison between speakers. The EIS corpus was designed to investigate possible factors influencing the adaptation of English loanwords

in Spanish. Therefore, it contains speech samples of 71 natives from two Spanish-speaking countries, namely Mexico and Spain. These two countries were chosen in order to test whether geographical proximity is still a relevant factor in the realization of Anglicisms (cf. Section 1.2). Informants were recruited primarily in the capital of each country, namely Madrid and Mexico City, but also in rural areas in the surroundings of the capitals, namely in Pedrezuela, Spain, that counts 5,892 habitants (INE 2019) and Chiconcuac, Mexico, which has 19,656 habitants (INEGI 2009). These two rural areas were chosen as they seem relatively comparable with regard to their social composition and their infrastructure as they are equally distant from the respective capital (both can be reached in 45 minutes by car). Due to the rather short distance between the urban and the rural areas, it can be assumed that the rural speakers use the same variety as the speakers in the respective capitals.

For reasons of research economy and to guarantee the speakers' confidence in the investigator (Pustka et al. 2018: 13; Milroy & Gordon 2003: 32), recruitment took place using a snowball technique, ensuring an equal distribution between females and males and across three different generations. This led to 25 speakers in each of the cities and to 10 (Pedrezuela) and 11 (Chiconcuac) speakers in the respective rural areas. The youngest speaker was 21 years old; the oldest was 92. The mean age was 48 years. The boxplot in Figure 1 shows the distribution of age.

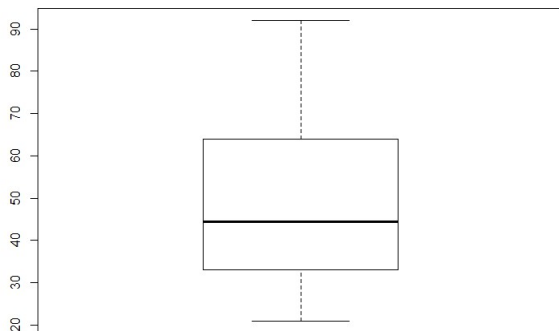


Figure 1: Distribution of age of the informants.

2.1.2 Task

In the framework of the project, speakers were asked to complete various tasks, including reading a word list. Since Anglicisms are only found rarely in spon-

taneous speech,⁴ this makes it possible to compare the pronunciation between many speakers. The word list consists of 102 items, among which there are 77 Anglicisms, the other 25 items being fillers and therefore not Anglicisms. Among the 77 Anglicisms are 29 Anglicisms that are morphologically and/or orthographically integrated or semantic loans (e.g., *administración*) and therefore also function as fillers, since they distract the informants. The word list provides seven Anglicisms that contain the voiced alveopalatal affricate in the source language (English), namely *jersey*, *ginger-ale*,⁵ *jet-lag*, *manager*, *jazz*, *gentleman*, and *digital*. Informants were asked to read the list aloud without familiarizing themselves with the words in advance. Recordings were carried out with the digital audio recorder ZOOM H4n Handy at a sampling rate of 44.1 kHz and a 16-bit depth. To further improve audio quality, the condenser microphone AKG C 520 was used.

2.1.3 Internal variables

Besides the speech material, the EIS corpus provides us with vast metadata. As previous research considers word frequency and age of the loanword as possible influencing factors in adaptation of loanwords, these variables were included in the analysis. Frequency values of each loanword were extracted from the corpora *American Spanish Web 2011* (esamTenTen11) and *European Spanish Web 2011* (eseuTenTen11), both accessible via the software Sketch Engine (Kilgarriff et al. 2014). Frequency values were divided by the total number of tokens included in the respective corpora to enable comparison between the two. In the analysis, frequency values from the American corpus were included only for speakers from Mexico, whereas values from the European corpus were only included for speakers from Spain in order to deal with differences in frequency between Mexico and Spain. Table 1 gives an overview of the frequency values for each word from the two corpora: *digital* is the most frequent, followed by *jazz* in both corpora. *Jet-lag* and *ginger-ale*, on the other hand, represent the less-frequent ones.

The age of the loanwords – also assumed as a relevant factor when it comes to the adaptation of loanwords – was defined as the time the loanword has existed in the target language (Spanish). To approximate this value, I extracted the year the loanword first appeared in a Spanish dictionary in the *Nuevo tesoro lexicográfico de la lengua Española* (NTLLE) (Real Academia Española 2019), which unifies an extensive selection of different dictionaries. I am aware of the limitations of

⁴In the interviews which were done in the framework of the project, only 141 occurrences of Anglicisms (78 types) were found in total among all speakers.

⁵Note that in this study, the analysis was limited to the phoneme in the first syllable to ensure that for each word, one phoneme was analyzed.

Table 1: Absolute and relative frequency of each loanword according to the American Spanish Web and European Spanish Web corpus.

	Freq. in European corpus	%	Freq. in American corpus	%
<i>digital</i>	263,013	130.10	652,311	87.26
<i>jazz</i>	37,182	18.39	96,871	12.30
<i>manager</i>	26,020	12.87	69,935	9.35
<i>jersey</i>	15,356	7.60	27,222	3.64
<i>gentleman</i>	858	0.42	1,622	0.22
<i>jet-lag</i>	196	0.10	230	0.03
<i>ginger-ale</i>	4	<0.01	12	<0.01

this method, since loanwords have to be more or less known by the community already before entering a dictionary. Nevertheless, this method was evaluated as the only possibility to obtain the data. Table 2 gives an overview of the year the different loanwords appeared in a dictionary in the NTLLE for the first time.

Table 2: Information on the first appearance of each loanword.

	Date of first appearance	Calculated age (years)	Dictionary of first appearance
<i>gentleman</i>	1895	124	<i>Zerolo</i>
<i>jersey</i>	1917	102	<i>Alemany y Bolufer</i>
<i>digital</i>	1983	36	<i>DRAE</i>
<i>jet-lag</i> ⁶	1984	35	<i>DRAE</i>
<i>manager</i>	1984	35	<i>DRAE</i>
<i>ginger-ale</i>	2014	5	<i>DRAE</i>
<i>jazz</i>	2014	5	<i>DRAE</i>

2.1.4 External variables

In addition, a large amount of extra-linguistic metadata on the informants was collected in the framework of the EIS project. In this analysis, I included the common sociolinguistic variables *age* and *sex*. To account for geographic differences between Mexico and Spain, the provenance of the speakers was also included. This makes it possible to test the hypothesis of previous research, which suggests a higher impact of English on Spanish in Hispanoamerica than in Spain (cf.

⁶First appearance as *jet*.

Section 1.1). Moreover, I included the informants' environments in the analysis to test the hypothesis that residents of rural areas are more commonly associated with conservative speech forms than those in cities (Sandøy 2014: 229; Chambers & Trudgill 1998: 46; Caballero Fernández-Rufete 1994: 134). However, this study aims to test whether this hypothesis is still true in a globalized world. Certainly, results in this regard have to be treated with caution, as Mexican immigrants from rural areas (Fussell 2004: 937) might have returned from the US and may use Anglicisms to a particular extent. Since contact with migrants who used to live in the US and therefore might incorporate non-adapted Anglicisms in their speech may also influence the realization of Anglicisms of speakers who have not lived in the US, the (binary) variable of speakers' contact with former migrants was also included in the study.

As only four speakers in the studied population speak an additional first language besides Spanish (in Mexico, one speaker indicates Totonaco and one Nahuatl as first languages; in Spain, one speaker indicates Valencian and one French as first languages), this variable was not included in the statistical analysis. However, the number of foreign languages spoken by the speakers was included.

Moreover, I included the variable *months spent abroad*, assuming that this variable might also influence the pronunciation of Anglicisms, since people who have lived in other (not only English-speaking) countries might have had more contact with foreign languages and cultures and might as a result show a more open attitude toward loanwords.

In addition, a variable called *Language Exposure and Affinity Score (LEAS)* that describes the affiliation of the speakers with the English language and American culture was included. To establish this score, speakers indicated on a five-point Likert scale how often they are in contact with the English language and rated their attitude toward the language and American⁷ culture. They were asked how often they read in English, watch films in English, make use of the Internet in English, and are in contact with the English language in their daily lives. They also rated their attitude toward the English language and American culture and indicated how many times they had visited an English-speaking country.⁸ In consequence, up to four points were added to the score per question. Furthermore, I included the variables *education* and *English level* in the analysis. Note that the latter – just as the *LEAS* – is based on self-evaluation, since speakers

⁷See Section 1.2 for an explanation as to why I chose to concentrate on American culture.

⁸Since this last information was not measured by means of a Likert scale, it was included in the score as follows: no points were added if the informant had never been to an English-speaking country, two points if they had been to one English-speaking country, and four points if they had been to two or more English-speaking countries.

rated their competence as *zero*, *basic*, *intermediate*, or *advanced*. This approach naturally represents a limitation, but was considered the most effective way to obtain the data.

2.2 Analysis

In the following, I will describe the different steps of data analysis, namely audio-visual coding, the calculation of interrater reliability, and the statistical analysis.

2.2.1 Audiovisual coding

To enable the analysis, data were first processed in PRAAT (Boersma & Weenink 2021) before being segmented and orthographically transcribed. Subsequently, three linguists with phonetic background (the author of the article, one MA student, and another person who had a master's degree in linguistics) transcribed all 568 Anglicisms phonetically. Two of the transcribers speak German as their first language and English and Spanish as foreign languages. The other speaks English as their first language. All of them studied Spanish and/or English linguistics and therefore display extended knowledge in either Spanish or English linguistics or both.

The three coders transcribed the data individually before the transcriptions were compared to grant objectivity. In doing so, each coder transcribed the data audiovisually, aligning the phonetic transcription with the acoustic signal.

2.2.2 Interrater reliability

To ensure the reliability of the analysis, the transcription of the phoneme in question by the three coders was compared and a Fleiss' kappa (Fleiss et al. 2003) was calculated. For calculation, the R package DescTools (Signorell 2021) was used. Fleiss' kappa was elaborated for the calculation of interrater reliability of more than two coders. Landis & Koch (1977) proposed the interpretation of kappa values as shown in Table 3.

Landis & Koch (1977) mention in their paper that these divisions are clearly arbitrary and should only be seen as useful benchmarks. The arbitrary nature of such benchmarks has been criticized, with observers pointing out that the effects of prevalence and bias on kappa must be considered when judging its magnitude (Sim & Wright 2005: 264). It has to be taken into account that kappa can be influenced by the number of categories in the measurement scale: The more categories available, the greater the potential for disagreement among raters, resulting in a lower kappa with many categories than with few. With these limitations

Table 3: Interpretation of kappa value (Landis & Koch 1977: 165).

Kappa statistic	Interpretation
<0.00	Poor agreement
0.00–0.20	Slight agreement
0.21–0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–1.00	Almost perfect agreement

in mind, the interpretation provided by Landis & Koch (1977) will only be used as orientation in this study. As phonetic transcription opens up many different possible categories, it can be assumed that calculated kappa values in phonetic studies could be even higher if categories were further reduced.

2.2.3 Statistical analysis

After calculating the interrater reliability, the transcriptions in which two or all three transcribers matched were selected. In cases where the transcriptions of all three transcribers gave different results, the author relistened to the phoneme and decided on one solution.

Data were then processed in R (R Core Team 2021). Analysis of the data took place in two steps. First, the transcribed data were analyzed descriptively to get an overview of the used phoneme adaptations and their respective frequencies. In a second step, statistical modelling was applied to infer the effects of the variables on the realization of the phoneme in question. For this, the R package lme4 (Bates et al. 2015) for fitting linear mixed-effects models was used. For statistical modelling, dichotomous coding of the transcribed data took place.

Therefore, I coded all realizations of grapheme-phoneme correspondences as *Spanish*, while all imitations of the English model were coded as *English*. The first group therefore included realizations as [ˈmanaxer] or [dixiˈtal]. In the latter group, I included all cases where no grapheme-phoneme correspondences were realized and therefore imitation of the English model can be assumed. Therefore, cases such as [ˈmæniðʒə] that were pronounced source language-like (*phoneme importation*), and cases like [ˈmanajer], which represent instances where the speakers imitate the English model but replace it with a phoneme that is native to their L1 (*phoneme adaptation*), were included in this group. Figure 2 gives an overview of the dichotomous coding.

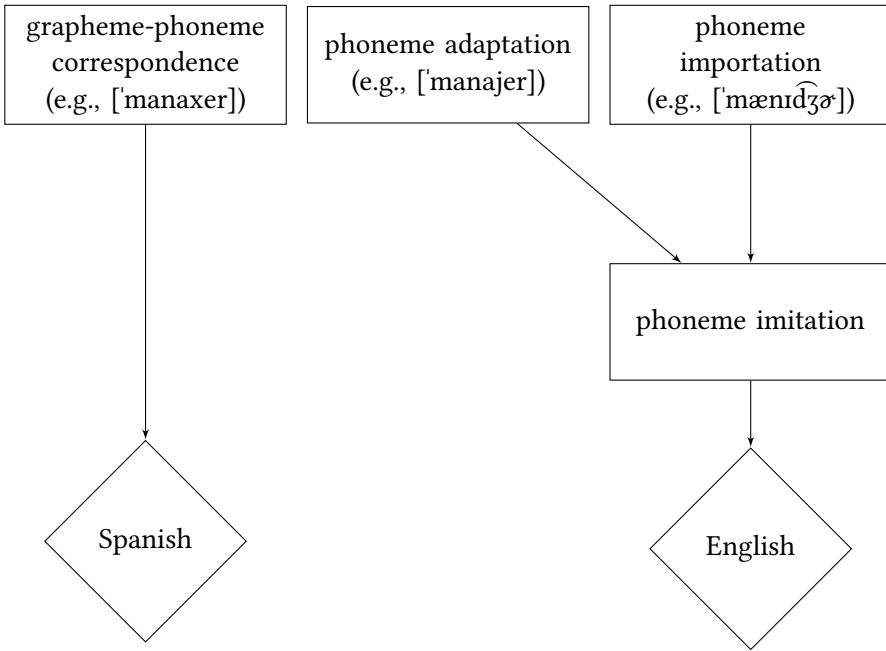


Figure 2: Dichotomous coding of the variable.

Prior to the calculation of the generalized mixed model, a directed acyclic graph (DAG) was developed (Figure 3) using the R package dagitty (Textor et al. 2016). In this graph, all assumed causal relationships between the dependent variable and independent variables (cf. Sections 2.1.3 and 2.1.4) and between each of the independent variables are illustrated, resulting in a very dense network of potential causal implications. As a next step, I aimed to identify among these relationships only those who actually (i.e., in terms of statistical significance) and directly (i.e., not confounding by another independent variable) infer the pronunciation. For this purpose, the generalized linear mixed model was built, employing a bottom-up and top-down approach.

I did not include the phoneme realizations of *jersey* in the statistical modelling, since it can be assumed that Mexicans and Spaniards do not refer to the same item with this lexeme (see Section 4 for further explanation). Before setting up the model, the variable *word frequency* was log-transformed to account for the differences between the different word frequencies. In addition, I scaled this variable and the other continuous variables to ensure that all continuous variables are on the same scale. For the categorical variables *area*, *country of origin*, and *sex*, the reference level was set as follows: city (*area*), Spain (*country of origin*),

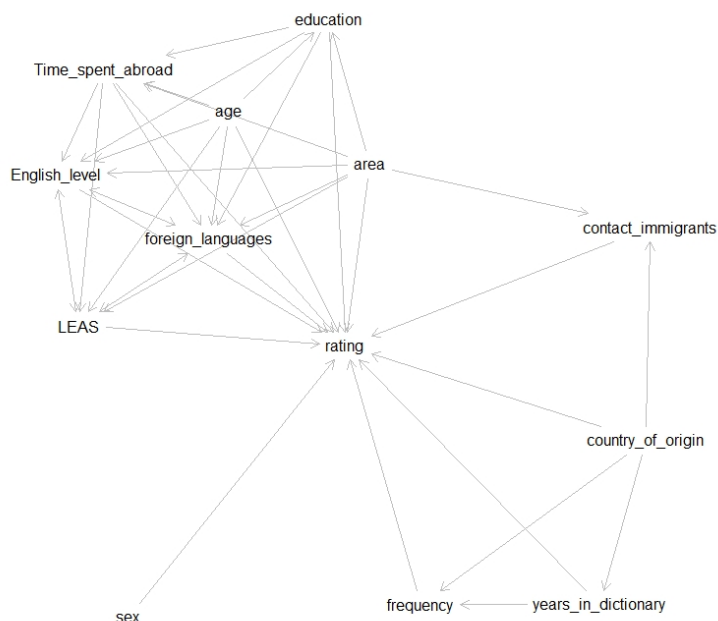


Figure 3: Directed acyclic graph of potential predictors of English phoneme realization.

female (*sex*). Further, I added the random effects *speaker*, *word*, *following vowel*, and *position* to the model to account for group-level variations. The random effect *speaker* therefore accounts for variations in pronunciation between the informants, whereas *word* accounts for any word-level differences in pronunciation. The random effect *following vowel*, on the other hand, accounts for differences in pronunciation due to the quality of the following vowel, which can be either a low, medium, or high vowel (/a/, /e/, or /i/, respectively). With the inclusion of this random effect, I therefore aim to expose whether the phonetic environment influences the realization of the phoneme. Finally, the random effect *position* accounts for variations in pronunciation between the different positions in which we find the phoneme in question, namely word initially and word internally (in both cases in onset position).

3 Results

3.1 Fleiss' kappa

The calculation of Fleiss' kappa revealed an interrater reliability of 0.71. According to Landis & Koch (1977), this represents "substantial agreement." Since phonetic transcriptions naturally show many different categories, the Fleiss' kappa values indicate a sufficient interrater reliability.

3.2 Realization of the phoneme in question

For the phoneme in question, all in all, nine different realizations among the speakers were found. Table 4 gives an overview of the different realizations of <ge>, <gi>, and <j> in all words among all speakers and their respective frequencies.

Table 4: Realizations of the phoneme in question in all words among all speakers.

Realization	Frequency in %
$\widehat{d}ʒ$	32.2
$\widehat{tʃ}$	5.8
ʃ	7.0
j	16.3
g	2.0
ç	9.3
x	16.3
χ	4.8
h	0.4
other	3.6
error	2.2

The analysis shows that in most cases (32 percent), speakers import the English phoneme. In the respective corpus, importation of $\widehat{d}ʒ$ is therefore not rare. In total, speakers realize grapheme-phoneme correspondences in 30.8 percent of cases. Note that the different allophones [x] (16.3 percent), [ç] (9.3 percent), and [χ] (4.8 percent) constitute this group. While the latter, which represents a posterior pronunciation of [x], is typical of the center and north of the Iberian Peninsular (Hualde 2014: 149), the voiceless palatal fricative [ç] appears in the speech

of Mexicans in front of the fronted vowels [i] and [u], as in *digital*. Even though the realization of the glottal [h] does not represent a standard Spanish grapheme-phoneme correspondence, it was included in this group, since it can more likely be interpreted as allophonic variation than as imitation of the English model.

In between the importation of the foreign phoneme and the realization of the Spanish grapheme-phoneme correspondence, different allophones that aim to imitate the English phoneme can be observed. Regarding the place of articulation, they range from postalveolar ([tʃ]), 5.8 percent, place of articulation shared with the English phoneme) to palatal pronunciations ([ʃ], 7 percent and [j], 16.3 percent) and velar realizations ([g], 2 percent). Further, regarding the manner of articulation, speakers realize affricate ([tʃ]), plosive ([ʃ] and [g]), and approximant ([j]) pronunciation. Regarding the voicing, 25.3% of the speakers realize a voiced phone, whereas 5.8% realize an unvoiced alternative.

The analysis of each word individually shows that the realization varies considerably between words. Table 5 gives an overview of the frequencies in percent of the realizations for each word.

Table 5: Realizations of the phoneme in question in the different words among all speakers in percent.

Realization	<i>ginger-ale</i>	<i>jet-lag</i>	<i>gentleman</i>	<i>jazz</i>	<i>manager</i>	<i>jersey</i>	<i>digital</i>
ḋʒ	60.6	54.9	36.6	36.6	19.7	16.9	0
tʃ	5.6	9.9	5.6	5.6	8.5	5.6	0
ʃ	0	8.5	7.0	19.7	1.4	12.7	0
j	5.6	8.5	12.7	25.4	52.1	9.9	0
g	4.2	0	7.0	0	2.8	0	0
ç	2.8	2.8	1.4	0	0	0	57.7
x	12.7	4.2	15.5	8.5	8.5	36.6	28.2
χ	1.4	4.2	1.4	2.8	0	9.9	14.1
h	0	0	0	0	0	2.8	0
other	4.2	4.2	8.5	0	7.0	1.4	0
error	2.8	2.8	4.2	1.4	0	4.2	0

The results show that *digital* displays the most cases of realization of grapheme-phoneme correspondences, followed by *jersey*, with 100 percent and 49.3 percent, respectively. In the first case, no imitation of the English phoneme was observed. In all other words, cases of phoneme imitation outweigh those of grapheme-

phoneme correspondences. That is, speakers are more likely to imitate the English model than to realize grapheme-phoneme correspondences. Within this group, where realizations of phonic imitation outweigh realizations of grapheme-phoneme correspondences, differences regarding the frequencies of phonic importation can be observed. *Ginger-ale* and *jet-lag* show the most cases of phonic importation, with 60.6 percent and 54.9 percent, respectively. *Gentleman* and *jazz* both show phoneme importation in 36.6 percent of cases. When pronouncing *manager*, on the other hand, only 19.7 percent of the speakers import the English phoneme. Nonetheless, the latter shows a strikingly elevated number of cases of the palatal approximant, with 52.1 percent of the speakers adapting the foreign sound to the respective realization.

3.3 Effects influencing the imitation of the English phoneme

To infer the effect of the selected variables on the phonic realization, a general linear mixed model was used, as described in Section 2.2.3. To summarize, for this model, the realization of Spanish grapheme-phoneme correspondence (or the imitation of the English realization) was set as the dependent variable. The results of this analysis are shown in the forest plot (Figure 4). Plots were created using the package *ggplot2* (Wickham 2016).

In this plot, for every independent variable, the effect on the realization is depicted by a horizontal bar, corresponding to the 95 percent (thin bar) and the 50 percent (thicker bar) credibility interval of the predicted effect. The x-axis represents the probability of the type of loanword realization; whereas positive values correspond to a higher probability of an imitation of the English phoneme, negative values correspond to a higher probability of realization of grapheme-phoneme correspondence. The mean effect of each predictor is marked by a dot. When interpreting these plots, it is important to consider that if the 95 percent credibility interval overlaps the zero-line (dashed vertical line), the model does not yield a distinct direction of the effect for this particular predictor, and therefore should be seen as statistically insignificant.

The model reveals a significant effect of the *LEAS* on the realization of loanwords. The higher the exposure and affinity to the English language and American culture (assessed by *LEAS*), the more likely it was the English phoneme was imitated, and the less likely the realization was grapheme-phoneme correspondent. Moreover, the plot shows an effect of word frequency on the loanword: the more frequent the loanword, the higher the probability that speakers pronounce the phoneme in question grapheme-phoneme correspondent. Furthermore, the type of area also impacted the realization, as being situated in cities or urban

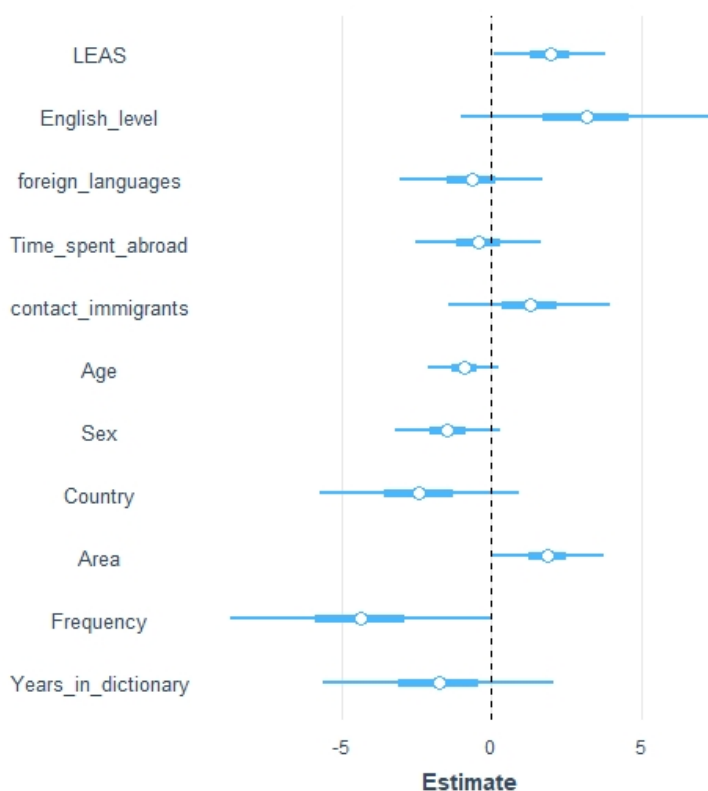


Figure 4: Forest plot of predictors of English phoneme realization.

areas was associated with a higher probability of an imitation of the English model. Yet further analysis showed that these two predictors interact with *LEAS*, thus distorting their effects. For this reason, it must be assumed that the effects observed in those measured are in fact not there.

Similarly, the analysis revealed that *age*, *foreign languages*, and *education* are confounds of *LEAS*. The effects found for these variables when considered separately in the model vanished in the process of the bottom-up analysis, taking into account interactions and confounding between the different independent variables. All these initially observed effects vanished with the integration of *LEAS* into the model, therefore indicating that the abovementioned variables may also influence the realization, but only indirectly, since they are collinear with *LEAS* and data are better explained by the latter. For instance, when analyzing *age* separately, it may seem as if a direct influence on the realization of loanwords exists;

however, this influence is actually explained by differences in *LEAS*, which decreases with increasing age. Respectively, we can assume a causal relationship between *education* and the realization insofar as speakers with a higher level of education may have higher scores in English media consumption. The variable *education* was not included in the final model since it interacts with *LEAS* and influences its values in the final model.

The final model does not reveal any direct influence of *English level* on pronunciation. However, bottom-up analysis showed that when only including *LEAS* and *English level* in the model, the effects both variables show in univariate analysis persist, although the effect of *LEAS* is higher than the effect of *English level*. It is therefore possible to assume a bidirectional causal relationship between the two variables, not only as speakers with higher language skills may have higher scores in English media consumption, but also since English media consumption may also improve language skills.

Furthermore, in the data analyzed, there is no indication of an effect of the country of origin on the preferred type of loanword realization – no significant difference between Spaniards and Mexicans was found. In addition, contact with former migrants who used to live in the US and the age of the loanwords did not influence the pronunciation of the Anglicisms.

The analysis of the random effects yielded no remarkable differences between the speakers. Standard deviation between the speakers is 1.69. The analysis of the random effect *word*, on the other hand, revealed greater differences in the realization between words. The forest plot in Figure 5 shows the group-level intercepts per word. The x-axis indicates the realization in relation to the population mean, while negative values indicate the realization of grapheme-phoneme correspondences and higher values the imitation of the English sound. However, these posterior intercepts have to be interpreted with caution, since they are drawn to the mean by the *pooling* effect during regression modelling, which prevents the model from overfitting. Figure 5 shows a certain tendency toward the probability of grapheme-phoneme correspondent realization being higher for *jet-lag* and *digital*, whereas the English phoneme in *manager* is more likely to be imitated. However, the depicted confidence intervals underline the uncertainty of the results and do not allow for specific interpretation.

Moreover, the analysis of the random effect *following vowel* revealed that in cases where the phoneme in question is followed by /i/, speakers are more likely to realize grapheme-phoneme correspondences than when followed by /a/ or /e/, as shown in Figure 6.

The analysis of the random effects showed that *following vowel* influences *word*. As such, it is not simply a case of the different words showing dif-

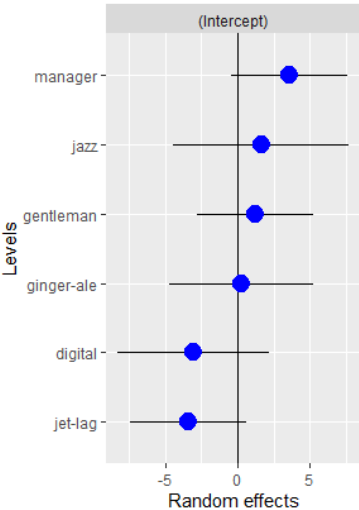


Figure 5: Forest plot of group-level intercepts per word.

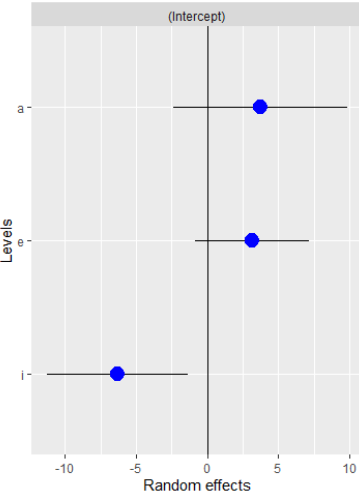


Figure 6: Forest plot of group-level intercepts per following vowel.

ferent realizations; the following vowels are also an important factor. Since the analysis of the random effect *position*, meaning if the phoneme in question is located word initially or internally, did not yield any remarkable differences, I did not include it in the final model.

4 Discussion

In this section, I will discuss the results described above; I will begin with the different phones, before discussing the results of the statistical modelling.

First of all, the analysis showed many cases of phoneme importation that have barely been reported in previous literature (cf. Section 1.2). Since native Spanish speakers can nowadays be constantly exposed to the model sounds via English media, it can be assumed that this promotes the importation of foreign phonemes. Additionally, since some Spanish varieties, such as those found in Paraguay and some areas of Ecuador and Argentina (Real Academia Española & Asociación de Academias de la Lengua Española 2011: 226), display the voiced alveopalatal affricate [dʒ] in the realization of <ll>, <y>, and <(h)i>, research on these varieties is needed to analyze whether the realization of the respective phoneme in English loanwords is even higher in the respective geographical areas.

Moreover, in cases where speakers imitate the foreign sound, but do not realize the voiced postalveolar affricate [dʒ], the analysis revealed that speakers realize phones of different places of articulation. Despite the realization of [tʃ], in all other cases, speakers shift the articulation of the phone further to the back. Since only 5.8 percent of the speakers imitate the English phoneme with the voiceless postalveolar affricate [tʃ], this leads to the assumption that the place of articulation is not decisive in imitation. The palatal pronunciation, which is realized by 23.3 percent and therefore by the majority of the speakers who imitate the English model but do not pronounce [dʒ], is nonetheless the closest to the postalveolar articulation point. Regarding the manner of articulation, results showed that only 5.8 percent of the speakers realize an affricate (other than the English variant), which leads to the assumption that the manner of articulation is not striking when it comes to the perception and subsequently to the realization of the phoneme. Concerning the voicing of the phone, the finding that 25.3 percent of the speakers realize a voiced phone (other than [dʒ]) in contrast to 5.8 percent who realize a unvoiced one suggests that the voicing plays a major role in the perception and realization of the English model. It obviously plays such a big role that speakers would rather realize the voiced palatal approximant [j], which is posterior in means of place of articulation, than the unvoiced variant of

the English affricate, which not only shares the same place of articulation, but also the same manner of articulation.

These results suggest that the voicing of the English affricate is the most striking feature when Spanish natives perceive the foreign sound. This has also been shown by Gómez Capuz (2001: 29). It can be assumed that it is more important in the imitation process than the place of articulation, even though the latter plays a major role in comparison to the manner of articulation, which seems rather unimportant in the selection of the closest sound in the native system.

Furthermore, the results show varying frequencies of the presented phones in the different words. This finding underlines that “each word has its own history.”⁹ Regarding the Anglicism *digital*, the first appearance used in a technical way dates from 1983, but the word was already used in Spanish in the nineteenth century to name a plant (> lat. DIGITALIS). Before the loanword entered the Spanish language via English, it was therefore already established and is not perceived as an Anglicism. It can consequently rather be seen as a semantic loan. Hence, we can assume that the new meaning coming from the Anglicism was linked to the already existing Spanish pronunciation. Concerning the Anglicism *jersey*, a considerable difference between Mexicans and Spaniards can be found; while 88 percent of the Spaniards realize grapheme-phoneme correspondences, only 12 percent of Mexicans do so. This geographical difference can be explained as the Anglicism is common in Spain to name the respective garment. It first appeared in a dictionary with this meaning in 1917, while it was already known in the Spanish lexicon for a much longer period (at least since 1705) as the denomination of the isle in UK waters. This suggests that the new meaning was linked to the already existing Spanish pronunciation of the English phoneme. Since Mexicans are more likely to use *suéter* for the respective garment, they do not link the same object to the Anglicism as Spaniards. One can assume that Mexicans more likely refer to the respective garment used in American football. Remember that *jersey* was not included in the statistical analysis, since speakers in each country use the term to refer to different items.

As shown above, *ginger-ale* and *jet-lag* showed the most cases of phonic importation, with 60.6 percent and 54.9 percent respectively. The analysis of the respective entries in the NTLLE shows that these Anglicisms range among the “younger” ones, with *ginger-ale* being included in 2014 and *jet-lag* in 1984. Further, it was shown that *gentleman* and *jazz* both showed phoneme importation

⁹According to Pustka (2007: 40), the quote “chaque mot a son histoire” can be assigned to Gilliéron: “Cette phrase ne se trouve pas explicitement dans l’œuvre de Gilliéron, mais lui est attribuée par de nombreux auteurs” (Pustka 2007: 40).

in 36.6 percent of cases. The analysis of the NTLLE showed that *gentleman* was first included in a dictionary in 1895, while the latter was not included until 2014. We therefore find an older and a very young loanword in this group. The analysis of *manager* showed lower rates of phonemic importation, while revealing high rates of adaptation to the palatal approximant. It can be assumed that the intervocalic position of the phone favors the articulation of the approximant instead of the affricate, which is more likely to be produced word initially. The first dictionary entry of *manager* dates from 1984, the same year as for *jet-lag*, which shows high rates of phonemic importation. At first glance, these findings already suggest that the date of entrance alone cannot be a reliable factor for predicting the adaptation of the foreign phoneme. Moreover, it has to be kept in mind that the year of first incorporation in a dictionary can only approximate the time, the loanword exists in the recipient language: First, loanwords are often already known for a mostly indeterminable time period in the speech community before they get incorporated in the dictionary. Second, the decision of incorporation of new items can also be influenced by purist tendencies of the language academies.

Instead, it can be assumed that the way of entrance (*eye loans* versus *ear loans*) and transparent orthography play a non-negligible role in the adaptation of English loanwords in Spanish. Even though some researchers claim that orthography only plays a limited role in loanword phonology (LaCharité & Paradis 2005: 241), others underline the importance of the graphic system (Pustka 2021: 505; Vendelin & Peperkamp 2006; Peperkamp & Dupoux 2003: 369): Since Spanish has a very transparent orthography, it can favor the realization of Spanish grapheme-phoneme correspondences. This analysis clearly shows many cases of grapheme-phoneme correspondences. However, the limitation of the chosen method, namely the reading of a word list, has to be considered. An influence of the visual code cannot be excluded. Therefore, future studies should also investigate the phonic realization of Spanish Anglicisms in more naturalistic data, such as interviews. However, comparability between the data is not as high as it is in the case of reading of a list, since Anglicisms occur only rarely in spontaneous speech (cf. Section 2.1.2). Moreover, picture naming tasks or discourse completion tasks, where at least the graphic code is not present, might be a fruitful possibility to analyze phonic adaptation.

Concerning the effects of the different variables on the realization of the phoneme, the generalized linear mixed model revealed a significant effect of *LEAS*. Further, the apparent effect of *area* and *frequency* might in fact only be the result of an interaction with *LEAS*. The effects of the variables *age*, *education*, *English level*, *foreign languages*, *area*, and *frequency* found when integrated separately into the model are partially supported by previous literature. Rodríguez González (2017:

127) and Gómez Capuz (2001: 42) suggest that speakers with a higher level of competence in English tend to imitate the English model. Rodríguez González (2017: 102, 123) concludes the same for speakers with a higher level of education, while Rodríguez González (2018: 354) states that more frequent loanwords are more likely to undergo phonic adaptation.

However, as this study examines this phenomenon by analyzing comparative data of various speakers in connection with sociogeographical variables, it allows for statistical modelling of multiple predictors. The statistical model suggests that the abovementioned predictors, also discussed in previous research, do not directly explain the change of preference regarding the realization of the English phoneme, but that these predictors are rather explained by the collinear predictor of English exposure and affinity (i.e., *LEAS*). In other words, the actual decisive predictor that drives the preference of imitation of the English model is the exposure and affinity to the English language and American culture. However, the limitation of the chosen method – the reading of a word list – has to be considered: speakers might adopt the pronunciation of the Anglicisms in different context situations.

Moreover, the statistical analysis did not reveal significant differences between speakers in Mexico and Spain. The assumption expressed in previous literature (Pustka 2021: 504; Rodríguez González 2017: 127; Oncins-Martínez 2009: 116) that English has a greater influence in Hispanoamerica than in Spain could not be confirmed by these data. Therefore, at least the analyzed data suggest that in a globalized world, geographic proximity to the US might not be as relevant, since speakers can expose themselves to the English language via easily accessible media. Besides written input, spoken English is available via streaming services as Netflix or video blogs shared via social media – just to mention some. However, more research is needed to compare the realization in different Spanish varieties. Moreover, the analysis did not reveal a significant influence of speakers' contact with people who have previously lived in the US and therefore may use a higher number of Anglicisms. In Mexico, in particular, one might assume that such contact situations could influence the pronunciation of speakers who have not lived in the US. Since the data analyzed in this study did not reveal such a correlation, it is reasonable to assume that contact with English in general, which does not necessarily have to happen through returning migrants, is a more crucial predictor. In addition, the analysis did not reveal any significant influence of the internal variable *years in dictionary* on the realization of the Anglicisms. The assumption made in previous literature (Gómez Capuz 2001: 21) that older Anglicisms are less likely to be pronounced source language-like, cannot be confirmed with these data.

The finding of the descriptive analysis described above – that the pronunciation of the phoneme in question varies significantly between words – can be confirmed by the statistical analysis. However, the analysis of the random effects shows that it is not only the history of the word itself that influences the realization of the phoneme. The realization also depends on the quality of the following vowel (low/medium/high), indicating that when followed by the high vowel /i/, the phoneme is more likely to be realized grapheme-phoneme correspondent. Future research could also incorporate Anglicisms where /dʒ/ is followed by /u/ and /o/ to test these occurrences. In this sample, the position of the phoneme (word initially versus internally) did not influence the realization when it comes to the difference between imitation and realization of grapheme-phoneme correspondences. However, within the group of imitation, it can be assumed that word initially, more speakers opt to import the phoneme than in intervocalic position (as described above). This might be due to the fact that the consonant is more salient in word-initial position. Nevertheless, more data in intervocalic position are needed.

The results of the statistical modelling show that when multiple potential predictors are included in complex statistical models, effects that would eventually remain hidden by other types of analysis can be presented. Therefore, and due to the fact that complex variables such as *LEAS* have not been used in previous research, to the best of my knowledge, future research on the realization of Spanish Anglicisms should be based on comparative corpora that include sufficient sociodemographic data to allow for statistical modelling.

5 Conclusion

Concerning the distribution of the different realizations, the analysis of the corpus data revealed that in 32.2 percent of the analyzed cases, speakers imported the English phoneme /dʒ/ when realizing the respective Anglicisms. The importation of the voiced alveopalatal affricate /dʒ/, which is not native to the Spanish phoneme system, is therefore not rare. Besides phoneme importation, in other cases of phoneme imitation, speakers rather realize a voiced palatal sound ([j] or [ɟ]) than the unvoiced alveopalatal affricate ([tʃ]). The voicing might therefore play a major role in the perception of the English phoneme. The realization of the grapheme-phoneme correspondence also plays an important role in the realization of the English phoneme, which may be due to the transparent nature of Spanish orthography.

Moreover, concerning the sociogeographic variables influencing the realization of the English phoneme, it has been shown that primarily speakers' expo-

sure and affinity to the English language and American culture was decisive; the higher one speakers' *LEAS*, the more likely imitation of the English model was observed. In contrast, at least the data studied here suggest that in a globalized world, geographic proximity may – at least regarding the regions compared in this study – no longer be a relevant factor.

Abbreviations

EIS	English Influence on Spanish
DAG	directed acyclic graph
LEAS	Language Exposure and Affinity Score
NTLLE	Nuevo tesoro lexicográfico de la lengua Española

Acknowledgements

I want to kindly thank Frederik Hartmann (University of North Texas) for providing advice and expertise for statistical analysis and comments on the manuscript, and Elissa Pustka (University of Vienna), the editors and two blind reviewers for a critical reading of this paper. All remaining errors are my own. Furthermore, I want to thank Luke Green and Martin Baumgartner for conducting the additional phonetic transcriptions and of course all the speakers who patiently participated in the study. The research stays in Mexico and Spain were funded by grants from the University of Vienna and the Rosita Schjerve-Rindler Fonds.

References

- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.
- Bäumler, Linda. 2023. *Anglicisms in Mexico and Spain in times of globalization: Use, pronunciation and perception*. University of Vienna. (Doctoral dissertation).
- Boersma, Paul & David Weenink. 2021. *Praat: Doing phonetics by computer [Computer program]*. www.praat.org (18 December, 2021).
- Briz, Antonio & Grupo Val.Es.Co. 1995. *La conversación coloquial. Materiales para su estudio*. Valencia: Universidad de Valencia.
- Caballero Fernández-Rufete, Salvador. 1994. Anglicismos: Un trabajo de campo. *Revista de filología inglesa* (18). 143–160.
- Cabanillas, Isabel De La Cruz & Cristina Tejedor Martinez. 2012. Email or correo electrónico? Anglicisms in Spanish. *Revista Española de Lingüística Aplicada* 25(1). 95–118.

- Cassano, Paul Vincent. 1973. The influence of American English on the phonology of American Spanish. *Orbis* 22(1). 201–214.
- Chambers, J. K. & Peter Trudgill. 1998. *Dialectology*. 2nd edn. Cambridge: Cambridge University Press.
- Chaudron, Craig. 2007. Data collection in SLA research. In Catherine J. Doughty & Michael H. Long (eds.), *The handbook of second language acquisition*, 762–828. Malden, MA: Blackwell.
- Eychenne, Julien. 2021. Corpus phonology. In Christoph Gabriel, Randall Gess & Trudel Meisenburg (eds.), *Manual of romance phonetics and phonology*, 319–341. Berlin/Boston: De Gruyter.
- Fleiss, Joseph L., Bruce Levin & Myunghee Cho Paik. 2003. *Statistical methods for rates and proportions*. 3rd edn. Hoboken, NJ: Wiley.
- Fussell, Elizabeth. 2004. Sources of Mexico's migration stream: Rural, urban, and border migrants to the United States. *Social Forces* 82(3). 937–967.
- Gómez Capuz, Juan. 2001. Estrategias de integración fónica de los anglicismos en un corpus de español hablado: Asimilación, compromiso y efectos estructurales. *Estudios de Lingüística* 15. 5–86.
- Gray, Jonathan. 2007. Imagining America: *The Simpsons* go global. *Popular Communication* 5(2). 129–148.
- Hamm, Bernd & Russell Smandych (eds.). 2005. *Cultural imperialism. Essays on the political economy of cultural domination*. Toronto: University of Toronto Press.
- Haspelmath, Martin. 2009. Lexical borrowing: Concepts and issues. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 35–54. Berlin/New York: De Gruyter Mouton.
- Haugen, Einar. 1950. The analysis of linguistic borrowing. *Language* 26(2). 210–231.
- Hualde, José Ignacio. 2014. *Los sonidos del español*. Cambridge: Cambridge University Press.
- INE, Instituto Nacional de Estadística. 2019. *Cifras oficiales de población resultantes de la revisión del Padrón municipal a 1 de enero*. <https://www.ine.es/jaxiT3/Datos.htm?t=2881#!tabs-tabla> (17 November, 2021).
- INEGI, Instituto Nacional de Estadística Geografía e Informática. 2009. *Prontuario de información geográfica municipal de los Estados Unidos Mexicanos*. https://www.inegi.org.mx/contenidos/app/mexicocifras/datos_geograficos/15/15030.pdf (17 November, 2021).
- Kang, Yoonjung. 2011. Loanword phonology. In Marc van Oostendorp, Colin Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 2258–2282. Malden, MA: Wiley-Blackwell.

- Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý & Vít Suchomel. 2014. The Sketch Engine: Ten years on. *Lexicography* 1. 7–36.
- LaCharité, Darlene & Carole Paradis. 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36(2). 223–258.
- Landis, J. Richard & Gary G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33(1). 159–174.
- Lorenzo, Emilio. 1997. Tratamiento del vocalismo inglés en español. *Estudios ingleses de la Universidad Complutense* 5. 81–90.
- Martínez Celdrán, Eugenio. 2015. Naturaleza fonética de la consonante ‘ye’ en español. *Normas* 5. 117–131.
- Meisenburg, Trudel. 1992. Graphische und phonische Integration von Fremdwörtern am Beispiel des Spanischen. *Zeitschrift für Sprachwissenschaft* 11(1). 47–67.
- Milroy, Lesley & Matthew Gordon. 2003. *Sociolinguistics: Method and interpretation*. Chichester: Wiley-Blackwell.
- Oncins-Martínez, José. 2009. Towards a corpus-based analysis of Anglicism in Spanish: A case study. *International Journal of English Studies* 1. 115–132.
- Peperkamp, Sharon & Emmanuel Dupoux. 2003. Reinterpreting loanword adaptations: The role of perception. In *Proceedings of the 15th international congress of phonetic sciences*, 367–370.
- Pratt, Chris. 1980. *El anglicismo en el español peninsular contemporáneo*. Madrid: Gredos.
- Pustka, Elissa. 2007. *Phonologie et variétés en contact. Aveyronnais et Guadeloupéens à Paris*. Tübingen: Narr.
- Pustka, Elissa. 2021. Loanword phonology in Romance. In Christoph Gabriel, Randall Gess & Trudel Meisenburg (eds.), *Manual of romance phonetics and phonology*, 503–528. Berlin/Boston: De Gruyter.
- Pustka, Elissa, Christoph Gabriel, Trudel Meisenburg, Monja Burkard & Kristina Dziallas. 2018. (Inter-)Fonología del Español Contemporáneo (I)FEC: Metodología de un programa de investigación para la fonología de corpus. *Loquens* 5(1). 1–16.
- R Core Team. 2021. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Real Academia Española. 2019. *Nuevo tesoro lexicográfico (NTLLE)*. <https://apps.rae.es/ntlle/SrvltGUISalirNtlle>.

- Real Academia Española & Asociación de Academias de la Lengua Española. 2011. *Nueva gramática de la lengua española: Fonética y fonología*. Madrid: Espasa.
- Rodríguez González, Félix. 2017. Variaciones fonológicas en el uso del anglicismo: Panorama y revisión crítica. *Revista Española de Lingüística* 47(2). 99–134.
- Rodríguez González, Félix. 2018. Aspectos ortográficos del anglicismo. *Lebende Sprachen* 63(2). 350–373.
- Sandøy, Helge. 2014. Linguistic globalization: Experiences from the Nordic laboratory. In Eline Zenner & Gitte Kristiansen (eds.), *New perspectives on lexical borrowing: Onomasiological, methodological and phraseological innovations*, 225–249. Berlin/Boston: De Gruyter Mouton.
- Schweickard, Wolfgang. 1991. Anglizismen im Spanischen. *Terminologie et Traduction* 1. 75–86.
- Signorell, Andri et mult. al. 2021. *DescTools: Tools for descriptive statistics*. R package version 0.99.44. <https://cran.r-project.org/package=DescTools>.
- Sim, Julius & Chris C. Wright. 2005. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy* 85(3). 257–269.
- Textor, Johannes, Benito van der Zander, Mark S Gilthorpe, Maciej Liśkiewicz & George TH Ellison. 2016. Robust causal inference using directed acyclic graphs: The R package ‘dagitty’. *International Journal of Epidemiology* 45(6). 1887–1894.
- Vendelin, Inga & Sharon Peperkamp. 2006. The influence of orthography on loan-word adaptations. *Lingua* 116(7). 996–1007.
- Wickham, Hadley. 2016. *Ggplot2: Elegant graphics for data analysis*. New York: Springer. <https://ggplot2.tidyverse.org>.