

Intervocalic /d/ deletion in Málaga: Frequency effects and linguistic factors

Rob Bedinghaus & Beatriz Sedó
Indiana University

Abstract

The current study examines the deletion of intervocalic /d/ in the Spanish of Málaga, Spain, based on a corpus of transcribed interview data (Lasarte Cervantes, Sánchez Sáez, Ávila Muñoz, & Villena Ponsoda, 2009). The focus of the study is the analysis of the effect of type and token frequency on the deletion of intervocalic /d/ from the perspective of Bybee's Usage-Based Model (2003), an approach only recently applied to the study of intervocalic /d/ deletion (Díaz-Campos and Gradoville, 2011, for the Spanish of Caracas).

The results of this study indicate that the deletion rate was higher for Málaga than for other varieties of Spanish (Cedergren, 1973; Díaz-Campos & Gradoville, 2011; D'Introno & Sosa, 1986; López Morales, 1983; Samper Padilla, 1996) both in general and in the context [+ado, +past participle]. We suggest that the more frequent use of present perfect in Peninsular Spanish (Howe, 2006; Schwenter & Torres Cacoullós, 2008) may have an effect on the deletion rate in the past participle. We also propose that the absence of a statistical effect of lexical frequency is due to the fact that the phenomenon is so advanced in this variety of Spanish that it has generalized to less frequent items too.

Keywords: Spanish, frequency effects, usage-based model, intervocalic /d/, variable-rule analysis

1. Introduction

In Spanish, the phoneme /d/ has different phonetic realizations depending on its position in a word. In word-initial position and after homorganic nasals and laterals it is prescriptively produced as a stop [d] (*disco* ‘disk’ [ˈdisko], *caldo* ‘broth’ [ˈkaɫdo]). However, word-internally between vowels, following other consonants, and word-finally it is produced as an approximant [ð] (*cada* ‘each’ [ˈkaða], *ciudad* ‘city’ [θiˈðað]). However, previous research has showed that there is significant variation in the realization of /d/. It has been reported that /d/ can remain occlusive in intervocalic position (Michnowicz, 2012) or become voiceless word-finally ([θ] or [t]), as in [θiˈðaθ] or [θiˈðat] for *ciudad* ‘city’, in areas of Spain and in Peruvian Spanish as well (Hualde, 2005; Schwegler, Kempff, & Ameal-Guerra, 2010). Finally, Spanish /d/ can also be deleted in intervocalic position and in word final position (e.g., *lado* ‘side’ [ˈlao], *verdad* ‘truth’ [berˈða]). Our focus in this paper will be the deletion of /d/ in intervocalic position.

The deletion of intervocalic /d/ has been the object of several studies where data from different dialects of Spanish have been analyzed. However, most of these studies have focused on the sociolinguistic factors that influence the variation and have not revealed anything about the effect of type and/or token frequency on the phonetic reduction of intervocalic /d/. There are two studies to our knowledge, Díaz-Campos and Gradoville (2011) and Bybee (2002), which have examined the effects of frequency on the reduction of intervocalic /d/ in Spanish. Moya Corral and García Wiedeman (2009) and Blas Arroyo (2006) also mentioned frequency in their studies, but did not analyze it in depth.

The present investigation contributes to the discussion on the effects of token and type frequency on the deletion of intervocalic /d/ in Spanish, comparing data from Díaz-Campos and Gradoville (2011) from Caracas, Venezuela, with data from the Málaga corpus that is part of the multi-dialectal *Proyecto Para el Estudio Sociolingüístico del Español de España y de América* (PRESEEA) (Lasarte Cervantes et al., 2009). We will also discuss the potential influence of the higher rates of use of the present perfect in Peninsular Spanish than other varieties (Howe, 2006; Schwenter & Torres Cacoullos, 2008), since this results in a higher frequency of past participles ending in *-ado* and *-ido*, which has been repeatedly shown to be the grammatical category that most favors /d/ deletion (Alba, 1999; Cedergren, 1979; D’Introno & Sosa, 1986; Navarro, 1983; Samper Padilla, 1996; Uruburu Bidaurrázaga, 1994).

2. Intervocalic /d/ deletion in Spanish

Previous studies of intervocalic /d/ deletion in Spanish have shown variation in rates of overall deletion across dialects and geographic regions, variation according to social factors such as sex, age, socioeconomic class (Alba, 1999 for the Dominican Republic; Cedergren, 1973 for Panama; D’Introno & Sosa, 1986 for Venezuela; Samper Padilla, 1996 for the Canary Islands); variation according to linguistic factors such as preceding and following phonological context (D’Introno & Sosa, 1986), grammatical category of the word (Blas Arroyo, 2006;

Samper Padilla & Pérez Martín, 2008), and, most importantly to our investigation, variation according to token and type frequency of the word (Bybee, 2002; Díaz-Campos & Gradoville, 2011).

In the investigation of social factors that affect the realization of intervocalic /d/ in Spanish, studies have generally found that deletion is typically favored by men (D’Introno & Sosa, 1986; Navarro, 1983; Samper Padilla, 1996; Uruburu Bidaurrázaga, 1994). However, others have found a slight preference among women for deletion that is rarely significant (Alba, 1999; Cedergren, 1973, 1979; Turell, 1996). The deletion of intervocalic /d/ is also generally favored by the lower socioeconomic classes (Alba, 1999; Cedergren, 1973, 1979; D’Introno & Sosa, 1986; Navarro, 1983; Uruburu Bidaurrázaga, 1994). When the factor of age is significant, older speakers tend to delete /d/ more than younger speakers (Alba, 1999; Turell, 1996). Nevertheless, a few studies have shown a slight favoring of deletion among the younger speakers over older speakers (Navarro, 1983; Uruburu Bidaurrázaga, 1994).

Of the linguistic factors that affect the realization of /d/, the most prominent is the phonological context. Many studies have clearly shown that the deletion of /d/ is cross-dialectally most favored in the phonological context /ado/, with the low back vowel /a/ preceding /d/ and the mid back vowel /o/ following it (Alba, 1999; Cedergren, 1973, 1979; Díaz-Campos & Gradoville, 2011; D’Introno & Sosa, 1986; Samper Padilla, 1996; Samper Padilla & Perez Martin, 1998; Uruburu Bidaurrázaga, 1994). Other favoring phonological contexts are /odo/ (as in *todo* ‘all-masc’), /oda/ (as in *toda* ‘all-fem’), and /ada/ (as in *nada* ‘nothing’). Regarding stress, Navarro (1983) found that the presence of a tonic vowel before /d/ led to the greatest deletion, while a tonic vowel following /d/ led to the least deletion. Another significant factor is the grammatical category of the word. Specifically, regular past participles, which in Spanish end in *-ado* and *-ido*, show higher deletion of /d/ than other grammatical categories such as nouns and verbs (Díaz-Campos & Gradoville, 2011; Samper Padilla, 1996). Adjectives can also show high rates of /d/ deletion (Alba, 1999; D’Introno & Sosa, 1986; Uruburu Bidaurrázaga, 1994). This is particularly true for adjectives that have participial endings such as *-ado* or *-ido* (Uruburu Bidaurrázaga, 1994). Furthermore, the *-ado* past participle ending almost always shows higher rates of deletion than the *-ido* participle ending (true in all of the above studies except for the Las Palmas data in Samper Padilla & Perez Martin, 1998). This effect is likely due to type frequency, given that the number of *-ar* verbs (which take the *-ado* ending in the participle) is much larger than the number of *-er* and *-ir* verbs (which take the *-ido* ending) (Hualde, 2010).

One study that has investigated in depth the deletion of intervocalic /d/ in Málaga is that of Villena Ponsoda, Díaz Montesinos, Ávila Muñoz & Lasarte Cervantes (2011). Using the data from Lasarte Cervantes et al. (2009) they found a deletion rate of 25% and found that younger speakers elided /d/ the most while older speakers elided the least. As had been found in other studies, there was more deletion with past participles and in the context *-ado*. Additionally, more deletion was found in paroxytonic words and in words with more than two syllables. Then, they focused their attention only in the words ending in *-ado* and found more

deletion when there was more solidarity, in spontaneous speech and among young speakers and male speakers. With respect to their Variable Rule Analysis the authors found that /d/ deletion was only favored in the *-ado* context, the *-ada* contexts and *todo*, *toda* and *nada*. Moreover, /d/ deletion was also favored by the younger speakers, the male speakers and when the /d/ was in postonic position. Both proper nouns and technical terms retained the /d/ to a high degree.

However, studies apart from Bybee (2002) and Díaz-Campos and Gradoville (2011) have not discussed frequency as a contributing factor. These two recent analyses have begun to look at /d/ deletion in a different way under the assumptions of Bybee's usage-based model.

3. The Usage-based Model

Bybee (2002) argued that frequency strongly influences sound change. She mentioned two possible outcomes depending on the degree of frequency of the words. First, she posited that words with a higher frequency of use tend to present higher degrees of reduction and deletion, such as t/d deletion in English and [ð] deletion in many dialects of Spanish. Her reasoning was that the more frequent the item is, the more automated the articulation becomes, resulting in the overlapping of sounds and the reduction or deletion of unnecessary sounds (p. 288). On the other hand, she also discussed changes in low-frequency words, which are principally analogical in nature. She used irregular verbs as an example to explain changes in low-frequency words, namely that while high-frequency irregular verbs maintain their irregularity, low-frequency irregular verbs tend to be regularized.

Frequency of constituency also has an impact on the realization of the word. A sound in certain words will have a higher degree of reduction or deletion when that word is frequently combined with some other word. For example, it has been shown that the Spanish tap /ɾ/ in the preposition *para* deletes more often when *para* appears with a following word frequently, as in the expression *para que* ('in order that'), than when it appears with a following word less frequently, as in *para estudiar* ('in order to study', Bedinghaus, 2013). According to Bybee (2002), a high degree of frequency of context will lead to the formation of an exemplar cluster, a set of constituents that are mentally stored as a unit, as shown by Bybee and Sheibman (1999), Díaz-Campos, Fafulas, and Gradoville (2012), and Bedinghaus (2013), which will lead to the accumulation of reductions and coarticulations through usage.

3.1 The effect of frequency on /d/ deletion in Spanish

Bybee (2002) analyzed /d/ deletion in the Spanish of New Mexico. For the study, all instances of past participles present in the data were removed prior to the analysis given that the previous literature had shown a higher rate of /d/ deletion in this context. Bybee (2002) found that high frequency tokens showed higher deletion rates and that this process diffuses throughout the lexicon affecting high frequency words first and extending to lower frequency words over time.

Díaz-Campos and Gradoville (2011) used Bybee's usage-based model to examine the effects of type and token frequency (Bybee, 2003) on the deletion of intervocalic /d/ in sociolinguistic interviews taken from the *Estudio Sociolingüístico de Caracas* (Bentivoglio and Sedano, 1993). In their study, they extracted a total of 7,200 tokens of intervocalic /d/, which they coded for both corpus frequency and CREA¹ frequency. Additionally, their dependent variable was binary (deletion vs. retention).

The results of their study showed that the overall deletion rate in the corpus was 10.1%. The multivariate analysis revealed that high frequency items in both corpora favored deletion (Goldvarb factor weights: CREA .55, Caracas corpus .53) as predicted by the usage-based model. For grammatical category, the past participles favored deletion the most (.73) followed by prepositions (.57) and adverbs (.56). Pronouns, adjectives, verbs and nouns all disfavored deletion. Within the category of past participle, they compared /d/ deletion in participles ending in *-ado* and *-ido* and *elsewhere*, meaning all other words not ending in *-ado* or *-ido*. Participles ending in *-ado* favored deletion with a factor weight of .81, followed by *-ido* with a factor weight of .69. The *elsewhere* category disfavored deletion with a weight of .47. Thus, as other previous work has shown, the past participle form *-ado* favors deletion the most and they demonstrate that it is a result of high type frequency in that *-ado* occurred in 269 different verbs and *-ido* only in 74 different verbs. They conclude that while both token and type frequency were shown to be influential in favoring /d/ deletion, type frequency is more important than lexical token frequency in the past participles since *-ado* favors deletion more than *-ido* forms or any other context. Having shown more deletion in the past participles, Díaz-Campos and Gradoville (2011) then analyzed the effect of frequency on the data without the participles. When the participles were excluded, the frequency effect increased in magnitude for both corpus and CREA frequencies, indicating a pattern of lexical diffusion. There was "more erosion of intervocalic /d/ in high frequency tokens in a fashion that shows lexical diffusion" (Díaz-Campos & Gradoville, 2011:9).

4. Past participle and present perfect

As previously mentioned, the context that has been reported to most favor the deletion of intervocalic /d/ is the context [+ado +participle]. In Spanish, regular past participles are formed by adding the endings *-ado/-ido* to the stem of the verb. For example, *hablar-INF/hablado-PART* 'speak-spoken', *comer-INF/comido-PART* 'eat-eaten' and *vivir-INF/vivido-PART* 'live-lived'. This form is found in the present perfect, which shows variation cross-dialectally in Spanish, with different varieties presenting different rates of use of the present perfect with respect to the use of the preterit.

¹ CREA: (Corpus de Referencia del Español Actual), by the Real Academia Española (Spanish Royal Academy).

It has been reported that the use of the present perfect in Peninsular Spanish is more extended than in other varieties of Spanish (Howe, 2006), and that it is advancing as the perfective marker by default, whereas this is not occurring in Mexican Spanish (Schwenter & Torres Cacoullós, 2008). In fact, according to Schwenter and Torres Cacoullós (2008), the present perfect in Peninsular Spanish is almost categorical as the hodiernal perfective and it is also used with non-specific temporal contexts.

Howe (2006) compared data from different studies that focus on the use of the present perfect as opposed to the use of the preterit and showed that, while in Latin American dialects the preterit was the most frequent form for past events, the present perfect was more frequent in the case of Peninsular Spanish (Table 1).

Table 1. Cross-dialectal distribution of present perfect vs. preterit (Howe, 2006).

TENSE	Spain	Mexico	Argentina	El Salvador	Peru
Present Perfect	53.4% (N = 953)	14.9% (N = 335)	13% (N = 232)	22% (N = 838)	29.5% (N = 1082)
Preterit	46.6% (N = 830)	85.1% (N = 1899)	87% (N = 1602)	78% (N = 2932)	70.5% (N = 2585)
TOTAL	1783	2234	1834	3770	3667

5. Research Questions

Given the clear findings that Peninsular Spanish is more advanced in the process of extending the PP to other contexts than any other Spanish dialects such as Caracas Spanish, which has not been shown to exhibit this phenomenon, one of the main purposes of the current study is to determine if a higher frequency of use of PP (assumed for this variety of Spanish, given the previous literature on PP) may lead to a higher rate of intervocalic /d/ deletion in past participle forms in Málaga than Caracas.

With this goal in mind, the following research questions guided this study:

- a) Is there an effect for token and/or type frequency in the deletion of intervocalic /d/ in the Málaga corpus?

- b) Is there a higher rate of deletion of intervocalic /d/ in past participle forms than in Díaz-Campos and Gradoville (2011) for Caracas Spanish?
- c) Given that a higher rate of use of the present perfect has been reported in Peninsular Spanish than in American varieties, how does the rate of deletion in this context in Málaga compare to varieties reported in other studies?

6. Methodology

We collected our data from the *Corpus Oral de la Universidad de Málaga*, a corpus of 60-minute sociolinguistic interviews with 12 men and 12 women, all of high socioeconomic class (Lasarte Cervantes et al., 2009). This corpus is a part of the multi-dialectal *Proyecto Para el Estudio Sociolingüístico del Español de España y de América* (PRESEEA). One of the male speakers was not included in our study because the transcript of the conversation, for whatever reason, was not available through the source from which we accessed the transcriptions and the recording was difficult to comprehend. We wanted to avoid mistakes due to not being able to sufficiently understand the speaker. While socioeconomic class was labelled the same for all speakers by Lasarte Cervantes and colleagues, rendering impossible a sociolinguistic analysis of that variable, age was not the same for all speakers. There were three age groups defined in the corpus (20-34, 35-54, and over 55), each group having four men and four women. However, we did not analyze extralinguistic factors for this study. These factors had already been analyzed by Villena Ponsoda et al. (2011). Villena Ponsoda et al.'s study focuses on social, stylistic and contextual factors, whereas the present study focuses on both frequency of usage (i.e., token frequency) and type frequency, given the fact that very few studies have investigated the deletion of intervocalic /d/ under the scope of frequency of usage.

For each of the 23 interviews we began the data extraction at the 15:00 minute mark in order to decrease a possible observer's paradox effect on the data, with the exception of one interview, where, due to the speaker participating significantly less than other speakers in the interview, the extraction began at the 6:00 minute mark in order to obtain sufficient tokens. From this point, we extracted the first 100 tokens (or less if the interview did not contain 100 tokens) of intervocalic /d/, retained or deleted, found in each interview for a total of 2,208 tokens. The tokens included cases in which orthographically the <d> appeared word-initially but phonetically was an intervocalic [ð] (ej. *casa de mi madre* ['ka.sa.ðe.mi.'ma.ðre]). Lasarte Cervantes and colleagues transcribed the interviews, which they checked multiple times for accuracy and specifically marked all cases of intervocalic /d/ deletion. In the corpus, the retained tokens are not specially marked while the deleted tokens are specifically marked with the following transcription convention: habla<(d)>o. We utilized their categorization of retention and deletion for each token as a benchmark. However, we verified each token impressionistically, and in the cases where any of the authors disagreed with the transcription both researchers listened to the recordings and concluded together that each was retained or deleted.

After extracting each token and coding for the dependent variable of retention or deletion, we coded for four groups of factors (for a total of nine factors). First, followed Díaz-Campos and Gradoville (2011) and coded for both corpus frequency (i.e., lexical frequency in the 23 interviews of the Málaga corpus) and also the frequency of each lexical item in the CREA corpus from the Royal Spanish Academy (i.e., Corpus de Referencia del Español Actual - ‘Reference Corpus of Contemporary Spanish’). Given that Díaz-Campos and Gradoville (personal communication) examined CREA frequency only for Venezuela and limited the search to oral tokens, we examined CREA frequency only for Spain, also limited to oral tokens, since we are examining an oral corpus from Spain. To determine high and low frequency for each lexical item, Díaz-Campos and Gradoville (2011) examined the distribution of frequencies in both corpora and determined the best point at which to divide high and low frequency tokens. To do this, they searched for the clearest division in the frequencies between words to divide high and low frequency (Díaz-Campos, personal communication). Thus, we followed the same procedure in order to define high and low frequency in both the CREA and Málaga corpora. We determined that for the CREA corpus, any word with 2,276 appearances or higher in the corpus was considered high frequency, and any word with less was considered low frequency. The highest frequency word in the CREA corpus was the pronoun *de*, with 138,149 tokens. For the Málaga corpus, we determined that any word with 138 or more appearances in the corpus of 23 interviews was considered high frequency, and less than 138 was considered low frequency. The highest frequency word in the Málaga corpus was also *de*, with 7,638 tokens. Although it appears that *de* has a much higher frequency than the point at which we made the division between high and low in both corpora, it is important to recognize that *de* is the most frequent word in both corpora by a very large margin. If *de* is excluded, the remaining words are much more comparable in frequency. It is also important to note that after the division between high and low frequency, both corpora had a comparable distribution of lexical items that were high and low frequency. There was only a very small number of lexical items which were considered frequent in one corpus and not the other. The corpora contained 24 (CREA) and 23 (Málaga) high frequency words.

The next factor for which we coded was the grammatical category of each word, since this has been shown to have an effect on the deletion of intervocalic /d/ (Díaz-Campos & Gradoville, 2011; D’Introno & Sosa, 1986). It should be noted that the definition of ‘participle’ for this factor only includes verbal participles (e.g., he **hablado** ‘I have spoken’), categorizing participles used as adjectives (e.g., estoy **cansado** ‘I am tired’) as adjectives.

We then coded for the surrounding phonetic context of each intervocalic /d/, including information about height and frontness of both the preceding and the following vowels. Finally, we coded for the *-ado/-ido* participial ending, including only the verbal past participles (e.g., he **comido**), to the exclusion of adjectival participles.

7. Analysis

To analyze the data we ran a binary logistic regression using GoldVarb X (Sankoff, Tagliamonte & Smith, 2005) to determine the effects of frequency, grammatical category, phonetic context and type of ending (*-ado/-ido*) on the deletion of intervocalic /d/. This analysis in Goldvarb requires a binary dependent variable and produces a factor weight between 0 and 1 that describes the relationship between a given factor (i.e. independent variable) and the dependent variable. In other words, a factor weight of greater than .50 indicates a favoring relationship between a given factor and the dependent variable. For example, a hypothetical factor weight of .75 for the factor category of high lexical frequency in the Málaga corpus would indicate that a high frequency in the Málaga corpus strongly favors /d/-deletion. On the other hand, a factor weight less than .50 would indicate a disfavoring relationship, namely that high frequency in the Málaga corpus did not lead to more deletion. Finally, it should be noted that after a preliminary analysis the independent variable (IV) *grammatical category* had to be excluded from the regression due to unresolvable interactions with other factors.

8. Results

In this section we will present the results of the binary logistic regression analysis. The overall deletion rate in our corpus was 14.9% (N=329/2,208), which is higher than the rate obtained by Diaz-Campos & Gradoville (2011) (10.1%; N=727/7,200). Villena Ponsoda et al. (2011) found an overall deletion rate in the corpus of 25%, which is higher than what was found in the present study. However, this could be explained by several factors. First, Villena Ponsoda et al. removed from the analysis those cases where the stressed syllable was not either the syllable containing the /d/ or an adjoining one, because they did not find any case of deletion. In the present study, these contexts were not removed from the analysis. Additionally, one of the male speakers was eliminated from the current analysis due to some limitations with the audio and, as was discussed in Villena Ponsoda et al.'s study, male speakers favored deletion more than female speakers. These two differences in the study might explain the differences in the deletion rate found in the data.

Of the nine factors included in our analysis, Goldvarb only selected four as being significant: frontness of the following vowel, past participle ending (*-ado* vs. *-ido*), frontness of the preceding vowel, and the height of the following vowel.

As can be observed in Table 2, the deletion of /d/ is favored when it is followed by a mid back vowel and when it is preceded by a central vowel. This corresponds to the context [+ado], which has already been associated with a higher deletion rate in many previous studies. This can also be seen in the second most significant factor selected by GoldVarb, which shows that the past participle ending *-ado* favors the deletion of /d/ by a wide margin over *-ido*.

Table 2. Factors favoring /d/ deletion

		Weight	%	N
Frontness following vowel	Back	.83	39.5%	245/620
	Central	.61	11.1%	49/441
	Front	.26	3%	35/1150
	<i>Range</i>	<i>57</i>		
Past participle ending	-ado	.74	93.2%	82/88
	-ido	.26	20.9%	18/86
	<i>Range</i>	<i>48</i>		
Frontness preceding vowel	Central	.75	30.7%	240/783
	Back	.39	7.4%	46/576
	Front	.33	5.3%	43/806
	<i>Range</i>	<i>42</i>		
Height following vowel	Mid	.58	19.5%	275/1410
	Low	.48	11.1%	49/441
	High	.25	1.4%	5/355
	<i>Range</i>	<i>33</i>		

Log likelihood = -590.286

As mentioned in the previous section, given the interactions encountered in the first analysis, the factor *grammatical category* had to be eliminated from the regression, but a cross-tabulation was performed in order to see the distribution of deletion and retention based on the grammatical category. As can be seen in Table

3 there is more retention in all the grammatical categories, except for the past participle, which shows more deletion than retention.

Table 3. Distribution according to grammatical category in the current study.

Gram. Category	Deletion	Retention	Total
Adjectives	25.6% (N=106)	74.4% (N=308)	414
Adverbs	10.4% (N=14)	89.6% (N=121)	135
Prepositions	2.7% (N=13)	97.3% (N=461)	474
Pronouns	19.6% (N=31)	80.4% (N=127)	158
Nouns	13.3% (N=56)	86.7% (N=364)	420
Past Participles	57.5% (N=100)	42.5% (N=74)	174
Verbs	2.1% (N=9)	97.9% (N=424)	433
TOTAL	329	1879	2208

A comparison with data from previous studies shows that the deletion rates are higher for Málaga Spanish than for other dialects of Spanish (Table 4) in the context *-vSdv(c)*, except for Las Palmas (Samper Padilla, 1996). However, it should be noted that the deletion rate of 37.6% for Las Palmas includes speakers of all social classes, whereas the data from Málaga, which shows a deletion rate of 32.8% for that context, only comes from the upper educated class, which, according to previous literature, tends to exhibit lower deletion rates than lower social classes. This might indicate that, should we analyze data from all the social classes in Málaga, we might find even higher deletion rates, perhaps even surpassing the deletion rates found in Las Palmas.

Comparing the Málaga data to other Peninsular dialects, as reported by Samper Padilla (1996), we also see that Málaga shows the highest rate of deletion in the [+ado, +past participle] context (Table 5).

Table 4. Comparison of /d/ deletion rates among dialects in the context -vSdv(c) (Data taken from different studies and summarized by Samper Padilla, 1996).

Dialect	% of deletion
Las Palmas (upper class)	16.1%
Las Palmas (all classes)	37.6%
San Juan de Puerto Rico	21%
Panamá	20%
Caracas	11.5%
Málaga (upper class; current study)	32.8%

Table 5. Comparison of [+ado, +past participle] /d/-deletion rates in Peninsular dialects

Variety	% deletion
Las Palmas (upper class)	46.8%
Jaén	88.7%
Granada coast	88%
Valladolid	90%
Málaga (upper class; current study)	93.2%

Finally, if we compare the results to those from Díaz-Campos and Gradoville (2011), who also used frequency in their analysis, we observe several differences. First of all, the overall deletion rate was slightly higher for Málaga Spanish than for Caracas Spanish (14.9% vs. 10.1%). Secondly, the differences in the deletion rate increased when we only look at the past participles, and even more so in the case of past participles in *-ado*. Lastly, the other observed difference is that, whereas in Díaz-Campos and Gradoville (2011) tokens with high lexical frequency favored deletion, this factor was not selected by Goldvarb in the present study, not even when the participles were excluded from the analysis.

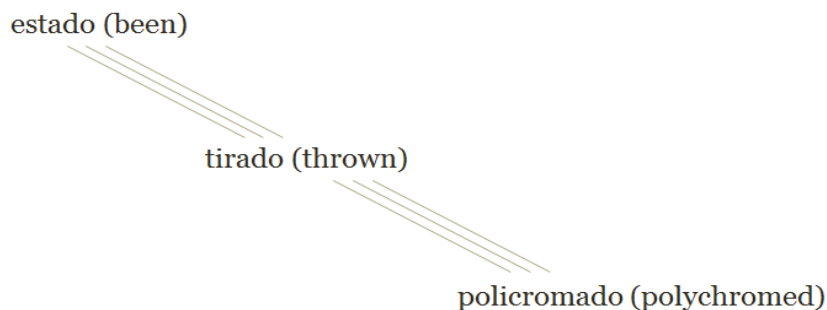
Table 6. Comparison with the results by Díaz-Campos and Gradoville (2011)

	Caracas	Málaga
Overall deletion rate	10.1%	14.9%
Deletion rate for past participles	21.7%	56.5%
Deletion rate for past participles in -ADO	30.4%	93.2%
Frequency effects	High frequency tokens favor deletion Increased effect when participles were excluded from the analysis	Not selected by GoldVarb When participles were excluded frequency still wasn't selected

9. Discussion

In order to answer the first research question regarding the effects of token and type frequency on the deletion of intervocalic /d/ in Málaga, we will look at the results of the binomial logistic regression. What is immediately clear is that token frequency was not selected as a significant factor. In other words, there was no significant effect of higher frequency of lexical items on the deletion of /d/. On the surface this is a surprising result, given the results of Díaz-Campos and Gradoville (2011) and other studies that have found a significant effect for token frequency (Bybee, 2002). However, a possible explanation along the lines of the usage-based model is that the process of the lexical diffusion of intervocalic /d/ deletion in Málaga is further along than in Caracas, resulting in the spread of deletion from more frequent words to more infrequent words over time (Figure 1). Lexical diffusion is the gradual spreading of a sound change through the lexicon from more frequent words, which are affected first, to more infrequent words as the change becomes more commonplace (Bybee, 2002). Therefore, it is plausible that a sufficient number of the infrequent words in the corpus used in this study have reached a level of /d/ deletion similar to the more frequent words to render the difference between deletion in frequent and infrequent words statistically indiscriminable. However, this is not the case for type frequency, which was selected as a significant factor by the regression.

Figure 1. Representation of the lexical diffusion with tokens from the present study ordered based on frequency of use.



Although it is not the first factor selected by the regression, there is reason to believe that type frequency is the primary force driving the variation. While grammatical category was not included in the statistical analysis due to interactions, it does paint an interesting picture. What is clear is that past participles show 57.5 percent deletion, whereas none of the other grammatical categories surpass 25.6 percent (adjectives), with most categories showing rates much lower than that. Furthermore, out of the 293 adjectives that show /d/ deletion, 159 are adjectives formed with *-ado/-ido* endings (*-ado*: 137; *-ido*: 22). When we look at the past participle category specifically, it is clear that there is an effect for type frequency in that *-ado* participles show much more frequent deletion than *-ido* participles. This is in line with previous research. However, previous research did not delve into why *-ado* participles would show more deletion. We are arguing that the reason behind this is the overall greater frequency of *-ado* participles in speech, given that *-ar* verbs represent the vast majority of verbs in Spanish (Hualde, 2010). This is not necessarily the case in the small *PRESEEA* corpus used for this study, as the frequency of *-ado* and *-ido* participles is almost identical. However, it has been shown that *-ado* participles tend to be more frequent than *-ido* participles generally. For example, Díaz-Campos and Gradoville (2011) found over twice as many *-ado* participles than *-ido* participles in their data.

As for the results found in this variety of Spanish, compared to other varieties, Málaga Spanish showed higher rates of /d/ deletion. When taking the context [+ado +participle] into account and comparing Málaga Spanish with other varieties of Peninsular Spanish, we can observe that, although all the varieties (except for Las Palmas) show very high deletion rates (88%-90%), Málaga Spanish shows more deletion (93.2%), even when it only includes data from the educated class. Additionally, when comparing the general deletion rates with those from non-Peninsular varieties of Spanish, we observe that Málaga Spanish also shows more deletion than other dialects. An explanation for this could be the more frequent use of the present perfect in Peninsular Spanish (Howe, 2006; Schwenter & Torres Cacoullós, 2008). More use of the present perfect will lead to

a higher frequency of use of the *-ado/-ido* participle forms, which, at the same time, could translate into higher degrees of deletion, even in the upper class. In fact, when comparing the results for /d/ deletion in past participles between Caracas Spanish (Díaz-Campos & Gradoville, 2011) and Málaga Spanish we observe a difference of over 30% (21.7% vs. 56.5% respectively), which becomes even larger when the past participle ends in *-ado* (30.4% vs. 93.2% respectively). Furthermore, the more extended use of the present perfect and, therefore, the higher frequency of the past participle in this dialect, would help explain why the deletion process is so advanced that it has extended to less frequent items.

10. Conclusion

The present study has investigated the effects of token and type frequency on the deletion of intervocalic /d/ in Málaga Spanish, as well as determined if the deletion rate was higher in this variety than in other varieties of Spanish, such as Caracas Spanish (Díaz-Campos & Gradoville, 2011).

In the case of the role of frequency, the absence of a statistical effect for lexical frequency on the data seems to suggest that the deletion process is so advanced in this variety that it has extended to less frequent items, following Bybee's (2002) theory of lexical diffusion, by which the change starts in more frequent items but then spreads to less frequent items. The fact that no statistical effect was found in Málaga Spanish but was found in Caracas Spanish, for example, might indicate that the process is more advanced in Málaga Spanish than in Caracas.

Frequency is also likely behind the fact that *-ado* past participles showed more deletion than *-ido* participles. Although both past participle endings showed more or less the same amount of frequency in the data used for the present study, Díaz-Campos and Gradoville (2011), for example, found that in their corpus the *-ado* participles were more frequent than the *-ido* participles (i.e., 385 tokens of *-ado* and 184 of *-ido*), which could explain the differences between the rates of /d/ deletion for both past participle endings.

As for the comparison of the deletion rates between Málaga Spanish and other varieties of Spanish, it has been shown that the deletion rate was higher for this variety than for any non-Peninsular variety of Spanish. We argue that this is like a result of a more frequent use of the present perfect in Peninsular Spanish, which would have an effect on the deletion rate of intervocalic /d/ in the past participle. Additionally, in the more specific [+ado +past participle] context, Málaga Spanish also shows a higher deletion rate compared to other Peninsular varieties, exhibiting a deletion rate of 93.2%.

Future studies on the deletion of intervocalic /d/ in Spanish could go in three different directions. First, more cross-dialectal studies are needed to verify the connection between the use of present perfect and the deletion of intervocalic /d/ in past participles and, by extension, in other contexts in general. Additionally, a study on this variety of Spanish that included social factors (socioeconomic status, gender, etc.) could be done to allow for a better comparison with other dialects of Spanish. Finally, an acoustic analysis rather than an impressionistic

analysis would be preferable, not only involving a dichotomous variable (retention vs. deletion), but also a continuum with different degrees of /d/ deletion, since it has been shown that intervocalic /d/ does not always present the same degree of weakening (D’Introno & Sosa, 1986). Finally, a statistical analysis that allows for a continuous frequency variable would be preferable to the binary frequency division that is required by Goldvarb. In order for this to be possible, other statistical programs and methods that allow for a continuous DV, such as Rbrul, should be considered for future studies on the effects of frequency on the deletion of intervocalic /d/ in Spanish, or any other variable that may be affected by frequency.

Acknowledgements

We would like to thank Dr. Manuel Díaz-Campos for his help, support, and feedback. We would also like to thank the anonymous reviewers for their comments. Any mistakes remain our own.

References

- Alba, O. (1999). Elisión de la/d/intervocálica postónica en el español dominicano. In *Estudios de lingüística hispánica: homenaje a María Vaquero* (pp. 3-21). Universidad de Puerto Rico.
- Bedinghaus, R. (2013). ¿Vas pa(ra) Málaga? The Reduction of *para* in Málaga, Spain: Effects of Frequency, Syntactic Category, and Social Factors. In *Selected Proceedings of the 16th Hispanic Linguistics Symposium*, ed. Jennifer Cabrelli Amaro et al., 238-252. Somerville, MA: Cascadilla Proceedings Project.
- Bentivoglio, P., & Sedano M. (1993). Investigación sociolingüística: sus métodos aplicados a una experiencia venezolana. *Boletín de Lingüística* 8. 3-35.
- Blas Arroyo, J. L. (2006). Hasta aquí hemos llega(d)o. ¿Un caso de variación morfológica? Factores estructurales y estilísticos en el español de una comunidad bilingüe. *Southwest Journal of Linguistics* 25. 39-72.
- Bybee, J. L. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. In *Language Variation and Change* 14. (pp. 261-290).
- Bybee, J. L. (2003). *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Bybee, J. L., & Scheibman, J. (1999). The effect of Usage on Degrees of Constituency: the reduction of don’t in English. *Linguistics* 37(4). 575-596.
- Cedergren, H. (1973). *Interplay of social and linguistic factors in Panama* (Doctoral Thesis). Ithaca: Cornell University.
- Cedergren, H. (1979). La elisión de la /d/: un ensayo de comparación dialectal. *Boletín de la Academia Puertorriqueña de la Lengua Española* VII. 19-29.
- Díaz-Campos, M., & Gradoville, M. (2011). An analysis of frequency as a factor contributing to the diffusion of variable phenomena. Evidence from

- Spanish data. In *Selected Proceedings of the 13th Hispanic Linguistics Symposium* (pp. 224-38). Somerville, MA: Cascadilla Proceedings Project.
- Díaz-Campos, M., Fafulas, S., & Gradoville, M. (2011). Going retro. An analysis of the interplay between socioeconomic class and age in Caracas Spanish. In *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics* (pp. 65-78). Somerville, MA: Cascadilla Proceedings Project.
- D’Introno, F., & Sosa, J. M. (1986). Elisión de la /d/ en el español de Caracas: aspectos sociolingüísticos e implicaciones teóricas. *Estudios sobre la fonología del español del Caribe*, (pp. 135-165). Caracas: Casa del Bello.
- Howe, C. (2006). *Cross-dialectal features of the Spanish present perfect: A typological analysis of form and function*. Ph.D. dissertation, Ohio State University.
- Hualde, J. I. (2005). *The Sounds of Spanish*. Cambridge, UK: Cambridge University Press.
- Hualde, J. I. (2010). *Introducción a la Lingüística Hispánica*. 2nd Ed. Cambridge, UK: Cambridge University Press.
- IBM Corp. Released 2011. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.
- Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. (2001). Probabilistic relations between words: evidence from reduction in lexical production. Frequency and the emergence of linguistic structure, ed. by Joan Bybee & Paul Hopper, (pp. 229-254). Amsterdam/Philadelphia: John Benjamins.
- Lasarte Cervantes, M. de la C., Sánchez Sáez, J. M., Ávila Muñoz, A. M., & Villena Ponsoda, J. A. (Ed.). (2009). *El español hablado en Málaga. Corpus oral para su estudio sociolingüístico. Nivel de estudios superior* (Vol. III). Málaga: Editorial Sarriá.
- López Morales, H. (1983). *Estratificación social del estado español de San Juan de Puerto Rico* (Vol. 18). Universidad Nacional Autónoma de México, Instituto de Investigaciones Filológicas, Centro de Lingüística Hispánica.
- Moya Corral, J. A. & García Wiedemann, E. J. (2009). *La elisión de /d/ intervocálica en el español culto de Granada: factores lingüísticos*. <http://rodin.uca.es:8081/xmlui/handle/10498/9895>. (22 January, 2011.)
- Navarro, M. (1983). La variación del segmento /d/ en Puerto Cabello. *Boletín de la Academia Puertorriqueña de la Lengua Española* XI(I). 65-72.
- Samper Padilla, J. A. (1996). El debilitamiento de la -/d/- en la norma culta de las Palmas de Gran Canaria. In *Actas del X congreso internacional de la Asociación de Lingüística y Filología de la América Latina*, 791-796. Veracruz: UNAM.
- Samper Padilla, J. A. & Pérez Martín, A. M. (1998). La pérdida de -/d/- en dos modalidades del español canario. *Philologia Canariensis*, 392-412. Las Palmas de Gran Canaria: Facultad de Filología de la Universidad de Las Palmas de Gran Canaria.
- Sankoff, D., Tagliamonte, S., & Smith, E. (2005). *Goldvarb X. A variable rule application for Macintosh and Windows*. Toronto: University of Toronto Department of Linguistics.

- Schwegler, A., Kempff, J. & Ameal-Guerra, A. (2010). *Fonética y fonología españolas* (4th ed.). Hoboken, NJ: Wiley.
- Schwenter, S.A., & Torres Cacoullos, R. (2008). Defaults and indeterminacy in temporal grammaticalization: The 'perfect' road to perfective. *Language Variation and Change* 20. 1-39.
- Turell, M. T. (1996). El contexto de la variación lingüística y su aplicación al estudio del morfema español -ADO. In *El español, lengua internacional (1492-1992)* (pp. 639-654). Asociación Española de Lingüística Aplicada, AESLA.
- Uruburu Bidaurrázaga, A. (1994). El tratamiento del fonema /-d-/ en posición intervocálica en la lengua española hablada en Córdoba (España). In *La Linguistique* 30 (1). 85-104.
- Ponsoda, J. A. V., Montesinos, F. D., Muñoz, A. Á., & Cervantes, M. C. L. (2011). Interacción de factores fonéticos y gramaticales en la variación fonológica: La elisión de/d/intervocálica en la variedad de los hablantes universitarios en la ciudad de Málaga. In *Variación lingüística y contacto de lenguas en el mundo hispánico: in memoriam Manuel Alvar* (pp. 311-359). Iberoamericana.