The Relative Importance of Lexical Frequency in Syllable- and Word-Final /s/ Reduction in Cali, Colombia

Earl K. Brown
California State University, Monterey Bay

1. Introduction

The literature on phonological variation and change abounds with studies about syllable- and word-final /s/ reduction in Spanish. In fact, "the aspiration and deletion of /s/ in dialects of Spanish may be the most extensively treated of all sound changes being investigated from an empirical, variationist perspective" (Ferguson, 1990, p. 64). Many factors have been shown to significantly affect this linguistic phenomenon. Terrell (1979) finds word length to be a significant factor in his Cuban data, with more deletion in polysyllabic words than in monosyllabic ones. Additionally, Terrell shows that redundant plural markers in noun phrases (that is, all but the first plural marker) are prone to deletion, such as the /s/ in *muchas*, *luchas*, and *internas* in the phrase "*...por las muchas luchas internas*" '*...because of the many internal struggles*' (p. 605, my translation). Besides word length, prosodic stress has been shown to be a significant factor. Alba (1982) demonstrates that syllable- and word-final /s/ is reduced more often in unstressed syllables than in stressed ones. This finding concurs with what is known about reductive processes in general, that is, that phonetic reduction is more likely in unstressed syllables than in stressed one. In fact, longer and larger articulatory gestures can be considered part of the definition of a stressed syllable (Hualde, 2005). Finally, one factor has been shown almost invariably to have the strongest conditioning effect on syllable- and word-final /s/ reduction: the following phonological context. Lipski (1984) shows that across fifteen dialects of Latin American Spanish, in general, word-final /s/ is reduced most before a consonant, less so before a vowel, and least before a pause.

In addition to these linguistic factors, several extralinguistic factors have also been shown to significantly constrain syllable- and word-final /s/ reduction. Alba (2003) shows that the level of formality of the communicative situation in which the speech is produced significantly conditions reduction, with more reduction in less formal situations. Additionally, Alba shows that the socioeconomic class of the speaker is influential, with more reduction among lower-class speakers. Fonetanella de Weinberg (1973) demonstrates that the sex of the speaker also contributes to reduction, with, in general, more reduction in men's speech. Finally, the geographic origin of the speaker greatly influences the level of syllable- and word-final /s/ reduction. Traditionally, dialectologists (see Zamora Vicente, 1970) have classified Spanish dialects according to the overall rate of syllable- and word-final /s/ reduction, making a distinction between *Tierras Altas* '*High Lands*' and *Tierras Bajas* 'Low Lands,' which *Tierras Altas* displaying relatively low levels of reduction and *Tierras Bajas* displaying high levels of reduction. This terminology is based on general trends observable in Latin American Spanish where syllable- and word-final /s/ reduction is the common in the Caribbean Basin and coastal areas, but uncommon in highland areas, such as central Mexico and the Andean regions of South America.

Despite the many studies that document syllable- and word-final /s/ reduction in Spanish, the lexical frequency of the word in which /s/ occurs has only recently been addressed in the literature. Brown & Torres Cacoullos (2002, 2003) show that final /s/ is reduced more often in high-frequency words than in low-frequency ones in the speech of four men in Chihuahua, Mexico, but only in word-final position. Brown (2004) finds the opposite effect in traditional New Mexican Spanish: lexical frequency is more influential in word-internal position than in word-final position. Minnick Fox (2006) also finds a significant effect from lexical frequency in word-internal position but not in word-final position in fifteen dialects of Latin American Spanish. Further, File-Muriel (2007) reports that lexical

frequency is a significant conditioning factor in the reduction of word-internal, syllable-final /s/ in Barranquilla, Colombia.

This paper contributes to this small, but growing, body of literature documenting the influence of lexical frequency on syllable- and word-final /s/ reduction by analyzing another dialect of Colombian Spanish: the spoken Spanish of Cali, Colombia. It is hypothesized that syllable- and word-final /s/ will be reduced more often in high-frequency words than in low-frequency ones in this dialect of Spanish, as has been found in other dialects of Spanish (Brown, 2002, 2003; Brown, 2004; Minnick Fox, 2006; File-Muriel, 2007). The results show that, as expected, syllable- and word-final /s/ is reduced more often in high-frequency words than in low-frequency ones. Further, the results demonstrate that lexical frequency has a stronger effect in word-internal position than in word-final position. This paper also shows that, even though lexical frequency is a significant factor, the following phonological segment is *the* most influential factor to condition this linguistic phenomenon in these data.

2. Lexical Diffusion

Sound changes have been studied for many years. In the nineteenth century, linguists took note of the apparent systematicity with which sound changes have affected the lexicon of various languages. This regularity led to the formulation of the Neogrammarian principle, which holds that phonetically motivated sound changes affect every word in a language that has the appropriate phonological context (Schuchardt, 1885). Many examples from the evolution of Latin into the Romance languages support this idea (Trask, 1996).

Despite the many examples of regularity, not all sound changes follow such a consistent pathway nor do all words necessarily follow the sound change to completion. Penny (2000, p. 72) points out that competing sound changes in Spanish created a residue of words unaffected by one of the sound changes. For example, initial /f/ in Latin changed to /h/ and later [ø] in Spanish, such as in *FAMINE > hambre* 'hunger,' and *FEMINA > hembra* 'female.' At the same time, there was a sound change that diphthongized stressed short /o/ of Latin into /we/ in Spanish, as in *PONTE > puente* 'bridge.' These two sound changes interacted with each other and left some words unaffected by the /f/ > /h/ > [ø] change, such as *fuera* 'outside,' *fuerte* 'strong,' and *fuego* 'fire.'

Instead of following a regular path, there is ample evidence that sound changes diffuse across the lexicon in an irregular manner, or in other words, sound changes spread by the process of "lexical diffusion" (Bybee, 2000; Phillips, 2006). In phonetically-motivated sound changes, the more frequent the word, the more likely the sound change will occur. Evidence is seen in the deletion of post-stress schwa in English. Hooper (1976) shows that high-frequency *every* is usually pronounced with two syllables, because the post-stress schwa is deleted, while low-frequency *mammary* most commonly has three syllables. As mentioned, recent studies have shown a similar effect of lexical frequency on syllable- and word-final /s/ reduction in Spanish: more reduction in high-frequency words (Brown & Torres Cacoullos, 2002, 2003; Brown, 2004; Minnick Fox, 2006; File-Muriel, 2007).

3. Data

The data (Travis 2005) used in this study come from spontaneous conversations between a female university student and her friends, family members, and colleagues. The corpus was recorded in Cali, Colombia, and all of the participants, including the university student, are native speakers of this dialect, having been born and raised there, or having moved there as a child. The university student was given a tape recorder for two months and instructed to "make ... recordings in conversational settings which naturally arose. ... Thus all conversations were entirely spontaneous, arising from the participants' own desire to talk with each other" (Travis, 2005, p. 13). As such, the recordings represent the most relaxed speech available to a conspicuous[1] recording device, and therefore is much

---

[1] Only a hidden device could record a more relaxed speech style. However, the practice of recording speakers without their knowledge has been deemed unethical by the linguistic community and is no longer acceptable (see Labov, 1984, pp. 51-2).

more likely to contain the speakers' vernacular. (More about the vernacular will be discussed in Section 4.) Therefore, this corpus lends itself very well to studying syllable- and word-final /s/ reduction.

The recordings consist of the speech of 31 speakers; 17 women and 14 men, recorded in groups of between two and four persons. They range in length from 8 to 40 minutes. The recordings were made in the places where the research assistant happened to be: in the kitchen of her home, in a restaurant, in her university's cafeteria, traveling in a car, etc. The topics are many and varied, and include the normal topics of conversation between friends and family members: food, family news, plans for the holidays, jobs, car trouble, etc.

The number of tokens of syllable- and word-final /s/ in this study is 5,661. Three-fourths of the tokens occur in word-final /s/ position while only one-fourth of the tokens occur in word-internal position, similar to the distribution of tokens between these two word positions in other studies of syllable- and word-final /s/ reduction (Brown, 2004; Minnick Fox, 2006).

4. Methodology

This study employs variationist methodology (Sankoff, 1988; Tagliamonte, 2006) because this method is able to detect patterns in apparently unsystematic variation. Through multivariate analysis (Sankoff, Tagliamonte, & Smith, 2005), this methodology can disentangle the influence of competing factors on a phenomenon of linguistic variation and change. For example, upon hearing the sentence "*E[s] que, una[ø] vece[h] me pongo a pensar*" 'It's that, sometimes I begin to think,' a listener may wonder why the speaker maintains the first word-final /s/ as a sibilant, but then reduces the next two. A variationist study may find that factors X, Y, and Z contribute to the maintenance of /s/ in *es*, but that factors A, B, and C contribute to the reduction of /s/ in *unas* and *veces*. In other words, variationist studies seek to explain the "why" of variation; or, what is the most likely reason a certain variant is used in a given context. In short, through the use of multivariate analysis variationist methodology attempts to create order in an apparently chaotic linguistic world.

The goal of most variationist researchers is to study linguistic variation in the vernacular, or the variety of speech that speakers pay the least amount of attention to (see Tagliamonte, 2006, p. 8). The importance of trying to obtain and analyze speakers' vernacular is that this register of speech is the "most systematic data for linguistic analysis" (Labov, 1984, p. 29), and therefore, more likely to be free of hypercorrection and style-shifting. Concerning our study, patterns of syllable- and word-final /s/ reduction and maintenance are the most consistent and systematic in the vernacular as compared to all other registers of linguistic production. A very effective manner to access speakers' vernacular is through spontaneous conversations, that is, conversations that arise out of the desire of the speakers to talk to each other. This technique keeps the level of formality, and therefore attention to speech, to a minimum.

5. Coding

The most cited variants of syllable- and word-final /s/ reduction in the literature are: maintenance [s], aspiration [h], and deletion [ø]. In this paper, aspiration is grouped with deletion, instead of maintenance, as the purpose of this study is to determine the conditioning effect of lexical frequency on syllable- and word-final /s/ reduction of any kind, whether [h] or [ø]. Performing a detailed analysis of the different realizations of reduction is outside the scope of this study. Thus, both aspiration and deletion are grouped together as reduction in this paper.

Determining whether each token of syllable- and word-final /s/ is maintained [s] or reduced [h, ø] was accomplished by impressionistic, auditory coding, as has been done in previous studies (Brown & Torres Cacoullos, 2002, 2003; Brown, 2004). This researcher listened to the audio recordings and each token of syllable- or word-final /s/ was coded as maintained if any sibilance, whether voiceless [s], voiced [z], or an assibilated aspiration [h^s], was detected. If no sibilance could be detected, the token was coded as reduced. Often, each token of syllable- and word-final /s/ was listened to several times, occasionally as many as ten times, to ensure accuracy.

To ensure the accuracy of the codes of the dependent variable, inter-rater reliability checks were performed. A sequential set of 1% of the number of tokens was coded by another listener who did not see the original codes assigned by the researcher. The percentage of identical codes is 88%. In addition to an inter-rater reliability check, an intra-rater reliability check was also performed. That is, this researcher listened to a sample of 1% of the data and recoded syllable- and word-final /s/ again without seeing the original codes. The amount of time that had lapsed between the original coding and recoding was approximately six months. The percentage of identical codes after the six-month time period had lapsed is 92%.

The choice to use impressionistic coding rather than coding instrumentally with a spectrographic analysis is based on the binary distinction between the two variants in this study: the presence of sibilance (maintained /s/) versus the lack of sibilance (reduced /s/). Impressionistic coding is optimal for this study as "discrete variants, as the label implies, tend to be easily distinguished and their coding is typically a relatively straightforward process" (Milroy & Gordon, 2003, p. 144). The wave format of each interview was viewed in order to highlight and play a few-second window in which a given token of syllable- and word-final /s/ occurred.

Word-final /s/ could not be coded when the following word began with an /s/. Distinguishing between the word-final /s/ in the one word and the word-initial /s/ in a following word is virtually impossible in conversational speech, as the two segments of sibilance merge together into one. Only when there was a pause between the two words could the word-final /s/ be accurately distinguished.

After the dependent variable was coded, the various independent factor groups were also analyzed. They are: following phonological segment, previous phonological segment, syllabic stress of the word in which /s/ occurs, realization of previous syllable- and word-final /s/ (as either maintained or reduced), and lexical frequency. In addition to the syllabic stress of the word in which syllable- and word-final /s/ occurs, and as part of the following phonological segment category, the stress of vowels following word-final /s/ was also taken into account. Terrell (1981) notes that word-final /s/ in Dominican Spanish is more likely to be realized as a sibilant when it is followed by a stressed vowel in article + noun phrases, such as *mis hijos* 'my children/sons.'

In order to calculate the lexical frequency of each word and arrange them into frequency categories, a large corpus of spoken Spanish with 3.5 million words was employed. This corpus consists of six small corpora and two large corpora: the *Corpus Oral de Referencia de la Lengua Española Contemporánea* (COREC) 'Oral Reference Corpus of Contemporary Spanish' (Marcos Marín, 1992) and eight cities that took part in the *Proyecto de Estudio Coordinado de la Norma Lingüística Culta de las Principales Ciudades de Iberoamérica y de la Península Ibérica* (*Norma Culta*[2]) 'Coordinated Project of the Study of the Educated Linguistic Norm of the Principal Cities of Latin America and the Iberian Peninsula' (Amparo Morales, 1990; Barrenechea, 1987; Caravedo, 1989; Esgueva & Cantarero, 1981; Instituto Caro y Cuervo, 1986; Lope Blanch, 1971; Rabanales & Contreras, 1979; Rosenblat, 1979).

A search of the frequency corpus was performed for the words in which syllable- and word-final /s/ occurs and the number of hits was recorded. Frequency groups (high versus low, etc.) were created based on the number of hits of exact forms of words. For example, the word *altos* 'tall PLU MASC' occurs 72 times in the frequency corpus, while the word *altas* 'tall PLU FEM' occurs 47 times.

In order to perform the multivariate analyses, discrete, rather than gradient, variants are required in each factor group. Consequently, the results from the various multivariate analyses reported below are based on an arbitrary dividing line that places 60% of the tokens in the high-frequency category and 40% in the low-frequency group.

6. Results

Syllable- and word-final /s/ reduction in these data patterns like it does in other studies (Brown & Torres Cacoullos, 2002, 2003; Brown, 2004; Minnick Fox, 2006; File-Muriel, 2007). First, these Cali

---

[2] The electronic files of the *Norma Culta* interviews were pain-stakingly prepared by J. Halvor Clegg at Brigham Young University. I express thanks to him for his work.

data have higher rates of reduction of /s/ in word-final position than in word-internal, syllable-final position, as seen in Table 1.

Table 1: Distribution of Final /s/ According to Word Position

Word position	Total	% data	Reduced	% reduced
Internal	1,436	25%	188	13%
Final	4,225	75%	1,860	44%
Total	5,661	100%	2,048	36%
Reduction: $p \leq 0.001$; chi-square = 444.11; df = 1				

Second, the following phonological segment is the most influential factor to constrain reduction in both word positions. Several other phonetic factors cited in other studies are also significant in these data, such as the previous phonological segment (Brown, 2004). More importantly for this paper, the lexical frequency was also selected as significant by two separate multivariate analyses for both word positions.

6.1. Word-Internal Position

The multivariate analysis selected three of the five factor groups included in the analysis as significant. In order of magnitude of effect, they are: following segment, lexical frequency, and the realization of the previous syllable- or word-final /s/. The two factor groups not selected as significant are the previous segment and stress. See Table 2.

Table 2: Multivariate Analysis of Variables Contributing to the Reduction of Word-Internal, Syllable-Final /s/

Independent Factor	Total	% data	Factor weight	% reduced	Reduced	Order
Following segment						First
/p, b, f, d, l, m, n, k, g/	344	24%	0.86	30%	102	
/t/	1,092	76%	0.36	8%	86	
			range = 50			
Lexical frequency						Second
High frequency	863	60%	0.67	16%	142	
Low frequency	573	40%	0.26	8%	46	
			range = 41			
Previous /s/						Third
Reduced	300	34%	0.63	19%	56	
Maintained	575	66%	0.43	9%	51	
	total = 875		range = 20			
$p \leq 0.001$; Log likelihood = -462.29; Input = 0.078; Total chi-square = 76.29; Chi-square/cell = 1.34; N = 1,436						
Factors not selected as significant: Previous segment, Syllabic stress						

Word-internal, syllable-final /s/ reduction is conditioned the most by the following segment, as in other studies. In this word position, final /s/ is reduced more before sounds other than /t/. In other words, a following /t/ has a maintaining effect on word-internal, syllable-final /s/, as seen in Table 2 above, and in more detail in Table 3 below (syllable-final /s/ before /f/ and /n/ has lower rates of reduction, but there are too few tokens to have any serious implications).

Table 3: Distribution of Word-Internal, Syllable-Final /s/ According to the Following Phoneme

	Total	% data	Reduced	% reduced
/f/	2	0.1%	0	0%
/n/	3	0.2%	0	0%
/t/	**1,092**	**76%**	**86**	**8%**
/k/	136	9%	14	10%
/g/	14	1%	3	21%
/p/	115	8%	29	25%
/m/	48	3%	34	71%
/l/	6	0.4%	5	83%
/d/	19	1%	16	84%
/b/	1	0.1%	1	100%
Total	1,436	100%	188	13%

As seen in the above tables, word-internal, syllable-final /s/ is followed the vast majority of the time by /t/: 76% of the time to be exact. The reason for such a high rate of a following /t/ is the frequent occurrence of a few words, as noted by Brown & Torres Cacoullos (2002, 2003). The copula verb *estar* and the proximal demonstrative *este* and their inflected forms make up 37% (527 tokens) of the total 1,436 tokens of word-internal, syllable-final /s/. A few other frequent words add another 13% (185 tokens): the verb *gustar* 'to please' and its inflected and derived forms, the preposition *hasta* 'until,' the second person singular and plural subject pronouns *usted* and *ustedes*, and the adjective *listo* 'ready.' Combined, these few words and their inflected and derived forms constitute 50% (712 tokens) of the total 1,436 tokens of word-internal, syllable-final /s/. Table 4 presents these few frequent words.

Table 4: Distribution of Word-Internal, Syllable-Final /s/ According to the Word or Lexeme in which it Occurs

LEXEME/ *word*	Total	% of data	Reduced	% reduced
listo	21	1%	0	0%
GUSTAR	47	3%	1	2%
ESTE	159	11%	6	4%
usted/es	88	6%	10	11%
hasta	29	2%	4	14%
other /t/	380	26%	6	15%
ESTAR	368	26%	59	16%
All others	344	24%	102	30%
Total	1,436	100%	188	13%

The reason syllable-final /s/ is maintained so much more often before /t/ than before other consonants lies in the place of articulation of /s/ and /t/. Penny (2002, pp. 101-3) demonstrates that the Golden Age (16th and 17th centuries) Spanish alveolar sibilants [s, z] evolved into the modern day dento-alveolar ones [ṣ, ẓ] in Andalusia and America, but maintained their alveolar place of articulation in northern and central Spain, in the Castilian dialect. Méndez Dosuna (1985) proposes that because /s/ and /t/ have an identical place of articulation (in most dialects of American Spanish), the amount of articulatory overlap between these two sounds is minimal. In other words, the sibilance of /s/ is less likely to be overlapped with the occlusion of /t/ because the front part of the tongue (whether the apex or the predorsum) is used to produce both sounds. Mendez Dosuna (1985, pp. 647-8) states:

Los movimientos articulatorios de /s/ y de /p/ son, pues, en buena medida simultáneos. Por el contrario, en el grupo /st/, dado que la articulación de ambos sonidos se realiza con la parte delantera de la lengua, los respectivos movimientos articulatorios son completamente incompatibles. La articulación de /s/ impide que el ápice de la lengua pueda iniciar el gesto de oclusión de /t/ y, así, las dos consonantes se suceden sin que exista prácticamente ninguna superposición entre los correspondientes períodos articulatorios. La duración de

la silbante resulta por ello considerablemente más prolongada que en el grupo /sp/. Por su parte, el grupo /sk/ se sitúa en una posición intermedia....

Therefore, the articulatory movements of /s/ and /p/ are, to a large extent, simultaneous. On the other hand, with the /st/ cluster, the respective articulatory movements are completely incompatible because the articulation of both sounds is realized with the front part of the tongue. The articulation of /s/ impedes the tip of the tongue from beginning the occlusive gesture of /t/ and, thus, both consonants occur with practically no superposition of the corresponding articulatory periods. Because of this, the duration of the sibilant is considerably longer than the /sp/ cluster. As far as the /sk/ cluster is concerned, it is in an intermediary position... (my translation).

Further, and as alluded to by Méndez Dosuna, the difference in manner of articulation (/s/ is a fricative while /t/ is a stop) also inhibits articulatory overlap. Evidence in support of this assertion comes from the high levels of /s/ reduction when followed by a /d/, as this consonant is most often articulated as an approximant instead of a stop when following an /s/. Consequently, /s/ is more likely to be maintained when followed by the stop /t/ than by other consonants. The data in this study support Méndez Dosuna's proposal.

In addition to the following segment, the lexical frequency of the word in which word-internal, syllable-final /s/ occurs significantly contributes to the reduction of this sound in these data. As expected, syllable-final /s/ is reduced more often in high-frequency words than in low-frequency ones, 16% and 8% reduction, respectively, as seen in Table 2.

As discussed in Section 5, the results reported in Table 2 are based on an arbitrary dividing line that places 60% of the tokens in the high-frequency category and 40% in the low-frequency group. However, even when the tokens are distributed more evenly across five frequency groups the effect from the lexical frequency is still present, but not in a uniform manner. As seen in Table 5, word-internal, syllable-final /s/ is reduced incrementally more often as lexical frequency increases, until the fifth and final group. Instead of having the highest amount of reduction of syllable-final /s/, as would be expected because the lexical frequency of these words is the highest, reduction falls sharply: from 25% reduction in Group 4 to 13% reduction in Group 5.

Table 5: Distribution of Word-Internal, Syllable-Final /s/ among Five Frequency Groups

Freq. group	Total	% data	Reduced	% reduced	Types
1 (least freq.)	278	19%	21	7.6%	195
2	275	19%	23	8.4%	117
3	295	21%	34	12%	51
4	295	21%	73	25%	10
5 (most freq.)	293	20%	37	13%	4
Total	1,436	100%	188	13%	377

Linear regression of Groups 1 through 5: $p \leq 0.3$; R-squared = 0.37; not significant
Chi-square test of Groups 4 and 5: $p \leq 0.001$; chi-square = 14.19; df = 1

Upon closer examination of these five frequency groups, it becomes evident that the uneven distribution of the following segment among the five groups creates the decrease in /s/ reduction in Group 5. As seen in Table 5 above, Group 4 has ten word types, that is, distinct words, while Group 5 only has four. Of the ten word types in Group 4, seven have a following /t/, while in Group 5, all four word types have a following /t/. The three tokens in which syllable-final /s/ is not followed by /t/ in Group 4 are the words *después* 'after,' *desde* 'since,' *mismo* 'same.' The first word has 35% reduction of /s/ while the second and third words have a substantially higher rate of reduction: 84%. The categorical presence of a following /t/ in Group 5 causes the average reduction in Group 5 to be smaller than would otherwise be the case. In fact, when the three words *después, mismo, desde* are removed from Group 4, the average rate of /s/ reduction in that group drops from 25% to 11%, placing it below Group 5 (13%). Further, when the tokens of /s/ followed by a sound other than /t/ are removed from all five frequency groups, the influence from lexical frequency is present in the predicted fashion: more reduction as frequency increases, as seen in Table 6.

Table 6: Distribution of Word-Internal, Syllable-Final /s/ Followed by /t/ among Five Frequency Groups

Frequency group	Total	% data	Reduced	% reduced
1 (least freq.)	151	14%	3	2%
2	175	16%	3	2%
3	256	23%	19	7%
4	217	20%	24	11%
5 (most freq.)	293	27%	37	13%
Total	1,092	100%	86	8%

Linear regression: $p \leq 0.009$; R-squared = 0.93

These results corroborate the findings of the multivariate analysis, namely, that the lexical frequency of the word in which syllable-final /s/ occurs significantly influences the rate of reduction of this sound. Further, these results also confirm the fact that the following phonological segment has a stronger effect than lexical frequency in word-internal position in these data. The rate of reduction of the ten words in frequency Group 4 varies more between words with a following /t/, /p/, /d/, and /m/ than does the rate of reduction between the five frequency groups.

Further evidence of the stronger influence of the following segment as compared to the lexical frequency is seen when the two factors are compared with each other, as in Table 7. Two-thirds of the tokens of word-internal, syllable-final /s/ that are followed by /t/ occur in high-frequency words, while two-thirds of the tokens that are followed by some other consonant occur in low-frequency words. If the lexical frequency were more influential than the following segment, we would expect the tokens of /s/ followed by /t/ to be reduced more than the tokens followed by other consonants. However, the opposite is true, thus supporting the claim that the following segment has a stronger effect than lexical frequency.

Table 7: Cross Tabulation of the Following Segment and Lexical Frequency

Tokens	Total	/t/	Other segments
High freq.	863 (60%)	748 (68%)	115 (33%)
Low freq.	573 (40%)	344 (32%)	229 (67%)
Total	1,436 (100%)	1,092 (100%)	344 (100%)
$p \leq 0.001$; chi-square = 134.15; df = 1			
Reduction			
High freq.	16%	11%	54%
Low freq.	8%	2%	17%
Total	13%	8%	30%

The same pattern is seen in the few frequent words (ESTAR, ESTE, *hasta*, etc.) presented above in Table 4. Almost all of the tokens (95%) of word-internal, syllable-final /s/ in those words are of high lexical frequency. Yet, the highest rate of reduction is only 16%, as seen above in Table 4. This is only half of the amount of /s/ reduction when followed by some other consonant besides /t/ (30%).

6.2. Word-Final Position

The multivariate analysis selected all five factor groups included in the analysis as significant. See Table 8.

Table 8: Multivariate Analysis of Variables Contributing to the Reduction Word-Final /s/

Independent Factor	Total	% data	Factor weight	% reduced	Reduced
Following segment					
/b, d, f, g, ɟ, l, m, n, r, tʃ, x/	1,273	30%	0.82	76%	971
/a, e, o/	559	13%	0.63	56%	311
/á, é, ó/	208	5%	0.56	48%	100
/i, u/	194	5%	0.47	38%	74
/í, ú/	86	2%	0.42	34%	29
/p, k/	904	21%	0.28	24%	213
Pause	824	20%	0.21	17%	142
/t/	177	4%	0.14	11%	20
			range = 68		
Previous /s/					
Reduced	1,117	40%	0.65	54%	599
Maintained	1,684	60%	0.40	37%	622
	total = 2,801		range = 25		
Previous segment					
Non-high vowel	4,132	98%	0.51	44%	1,837
High-vowel	93	2%	0.28	25%	23
			range = 23		
Syllabic stress					
Unstressed syllable	2,635	62%	0.53	46%	1,199
Stressed syllable	1,590	38%	0.45	42%	661
			range = 8		
Lexical frequency					
High	2,530	60%	0.52	47%	1,189
Low	1,695	40%	0.47	40%	671
			range = 5		

$p \leq 0.04$; Log likelihood = -2265.31; Input = 0.42; Total chi-square = 215.12; Chi-square/cell = 1.63; N = 4,225

The following segment has the largest magnitude of effect on the reduction of final /s/ in word-final position, as it does in word-internal position. Word-final /s/ is reduced most when followed by a consonant other than a voiceless stop, and least when followed by /t/, similar to the patterns in word-internal position. The influence of a following vowel falls between these two extremes, with non-high vowels favoring the reduction of word-final /s/ more than high vowels, and unstressed vowels favoring reduction more than stressed ones. The height of the following vowel is more influential than its stress, evidenced by the fact that the stressed non-high vowels (/á, é, ó/) fall immediately below their unstressed counterparts (/a, e, o/) and immediately above high vowels (/i, u/ or /í, ú/) in the hierarchy of constraints, as seen in Table 8. As expected from the results in previous studies (such as Minnick Fox, 2006), a following pause falls near the bottom of the hierarchy; that is, a following pause favors maintenance of word-final /s/. Only a following /t/ favors maintenance more.

The height of the following vowel contributes to the reduction or maintenance of word-final /s/. The mouth is a more closed position during the articulation of the high vowels /i, u/ than it is during the articulation of the low and mid vowels /a, e, o/. As such, less articulatory effort is needed to create the sibilance of /s/ when it is followed by a high vowel, as a speaker (unconsciously) anticipates creating a more closed position of the mouth in order to articulate the high vowels /i, u/. On the other hand, as a speaker anticipates the more open mouth positions of /a, e, o/, the tongue moves away from the roof of the mouth, thus requiring more effort to create sibilance and therefore increasing the likelihood of /s/ reduction. Although these trends are not categorical, the results in Table 8 support this argument.

In addition to the height of the following vowel, the stress of the following vowel significantly contributes to the reduction of word-final /s/. Word-final /s/ is more likely to be reduced when followed by an unstressed vowel. The reason is that reductive processes in general are more likely to occur in unstressed syllables than in stressed ones. In fact, larger and longer articulations are part of the definition of stressed syllables. The strong tendency in Spanish towards open CV syllable structure causes word-final /s/ to be converted into a syllable-initial /s/ by the process of resyllabification. This occurs when one word ends in a consonant and the following one begins with a vowel, such that the final consonant and the initial vowel combine to form an open CV syllable, for example: *dos años* > *do-sa-ños.* Consequently, when word-final /s/ is followed by a stressed vowel, it not only occurs in a newly formed stressed syllable, it also occurs in syllable-initial position, both of which favor maintenance of /s/. The result is a lower rate of word-final /s/ reduction when it is followed by a stressed vowel.

As seen in the data for word-internal position, lexical frequency was selected as significant in the multivariate analysis in word-final position, with more reduction in high-frequency words than low-frequency ones. This is the case when the tokens are divided up into two frequency groups, as in Table 8. However, when the tokens are divided among five frequency groups, the influence from lexical frequency is not consistent. The first three frequency groups show the expected trend: more reduction with each successive increase in lexical frequency. However, the last two groups, Groups 4 and 5, which contain the most frequent words, break with the trend in that the rate of reduction actually decreases, as seen in Table 9.

Table 9: Distribution of Word-Final /s/ among Five Frequency Groups

Frequency group	Total	% data	Reduced	% reduced	Types	Fav. Cont.
1 (least freq.)	846	20%	309	37%	542	39%
2	849	20%	362	43%	211	47%
3	756	18%	395	52%	28	56%
4	789	19%	375	48%	4	56%
5 (most freq.)	985	23%	419	43%	2	46%
Total	4,225	100%	1,860	44%	787	48%
Linear regression: $p \leq 0.5$; R-squared = 0.2; not significant						

Interestingly, the explanation for this unexpected behavior, like in word-internal position, is based on the interaction between the following segment and lexical frequency. It appears that the following segment overpowers the influence from the lexical frequency. As seen in Table 9, the amount of word-final /s/ reduction in each frequency group roughly mirrors the frequency with which this sound occurs in a favorable phonological context for reduction, seen in far right column. In other words, the likely reason that the high-frequency groups (Groups 4 and 5) show a decrease in the amount of reduction of word-final /s/ is because there is an accompanying decrease in the occurrence of this sound in a phonological context favorable to reduction. Phonological contexts favorable to reduction are those in which /s/ is reduced more than the average rate of reduction. Examples of phonological contexts favorable to reduction include, among other contexts, a following /b, d, g/.

These results support the claim being made here that the following segment has the largest conditioning effect on the reduction of syllable- and word-final /s/ in the Colombian data, even though lexical frequency exerts a significant influence when not overpowered by the following segment.

To further investigate the influence of lexical frequency in word-final position in these data, three morphological groups based on the morphemic status of word-final /s/ were created: monomorphemic /s/ (no morphological function), plural /s/, and verbal /s/ (*tú* '2s SUBJ PRON' and *nosotros* '1p SUBJ PRON' verbal desinences). Lexical frequency has a statistically significant effect in the expected direction (more reduction with high-frequency words) with plural and verbal /s/, but not with monomorphemic /s/.

In summary, the lexical frequency of the word in which word-final /s/ occurs has a significant influence on the rate of reduction of this sound when the tokens are divided among two groups (high versus low). However, when the tokens are divided among five frequency groups, the effect from the

lexical frequency is lost because of the more frequent occurrence of Groups 4 and 5 in phonological contexts unfavorable to reduction. Additionally, the various morphological functions that word-final /s/ can have influence the magnitude of effect that lexical frequency has; plural and verbal /s/ are significantly affected by lexical frequency while monomorphemic /s/ is not.

7. Discussion and Conclusions

This study elucidates the classification of the Cali dialect among the various dialectal regions of Colombia. Lipski (1994, pp. 209-13) divides Colombia into four dialectal regions based on phonological characteristics: (1) central highlands, (2) Caribbean coast, (3) Pacific coast, and (4) Amazonian region. He states that in the central highlands, syllable- and word-final /s/ is retained as a sibilant, that is, [s]. Further, Lipski (1994, p. 351) proposes that the "consonant-retaining Bogotá dialect exercises an influence even in coastal areas where consonantal reduction prevails." With the relatively low levels of overall reduction of syllable- and word-final /s/ in these Cali data, it is safe to conclude that Cali falls within the influence of the consonant-retaining Bogotá dialect despite being located closer to the coast than to Bogotá. A comparison of levels of reduction of word-internal, syllable-final /s/ reduction in these Cali data and in data from the coastal dialect of Barranquilla (File-Muriel, 2007) support this classification; while these Cali data display a 14% reduction of word-internal, syllable-final /s/, File-Muriel (2007) reports a 51% reduction in his Barranquilla data.

The results of this study show that, in these data, the following phonological context has the largest conditioning effect on syllable- and word-final /s/ reduction. Two separate multivariate analysis runs assign a substantially larger range to the following segment than to the other factor groups. In both cases, a following consonant other than /t/ favors reduction more than other sounds, while a following /t/ favors maintenance the most. This result supports Méndez Dosuna's (1985) proposal that there is little articulatory overlap between /s/ and /t/ because both sounds are pronounced with the front part of the tongue and because the articulation of /s/ inhibits the tip of the tongue from beginning the occlusive gesture of /t/.

In additional to the following phonological context, other phonetic factors also significantly contribute to the reduction of syllable- and word-final /s/ in these data: the previous phonological segment, the realization of the previous syllable- and word-final /s/, and the stress of the syllable in which syllable- and word-final /s/ occurs. These results concur with previous reports about this linguistic phenomenon (Brown & Torres Cacoullos, 2002, 2003; Brown, 2004; Minnick Fox, 2006; File-Muriel, 2007).

In addition to these phonetic factors, the usage-based factor of lexical frequency also significantly conditions syllable- and word-final /s/ reduction. In an exemplar model of mental representation (Pierrehumbert, 2001; Bybee, 2001, 2006), the exemplars of high-frequency words are more likely to have a reduced variant of /s/ than are low-frequency words, as it is proposed that sub-phonemic detail is recorded in memory. Thus, high-frequency words are more prone than low-frequency words to display a reduced variant of /s/ in production even when all other factors are controlled for, an effect of the mental representation of the high- and low-frequency words in memory. The results of this study support and are best explained within this exemplar model of mental representation, as syllable- and word-final /s/ is reduced more often in high-frequency words than in low-frequency ones, even when other factors are held constant, as measured in two multivariate analyses. With a detailed sub-phonemic mental representation of words in memory, there is little need in the exemplar model adopted in this study for complicated rules and transformations in order for the appropriate sound to be produced in speech. Thus, this model is in stark contrast to traditional models of phonology that rely on complicated rules, such as generative phonology and Optimality Theory.

While lexical frequency is a significant conditioning factor on both syllable- and word-final /s/ reduction, the effect from this factor is stronger in word-internal position. As seen in Table 2 and Table 8, lexical frequency receives a range of 41 points in word-internal position, but a range of only 5 points in word-final position, indicating a much stronger influence in word-internal position. This result coincides with results from Brown (2004) and Minnick Fox (2006).

The difference in conditioning strength of lexical frequency in the two word positions appears to be the result of the alternating versus uniform phonological environment in word-final position and word-internal position, respectively. For example, the word-internal, syllable-final /s/ in e<u>s</u>tás is always followed by /t/ and is therefore in a uniform environment, while the word-final /s/ in está<u>s</u> is followed by many different sounds, depending on the word (or pause) that follows. It appears that the uniformity of the following phonological environment in word-internal position allows the lexical frequency of the word to exercise a stronger conditioning effect on the reduction of final /s/ in that word position. Conversely, in word-final position, it appears that the following phonological segment overpowers the influence of lexical frequency since the following segment is always changing.

Although it exerts a strong influence in word-internal position, lexical frequency places second in the hierarchy of significant factor groups in that word position, falling below the most influential factor, the following phonological segment. This finding contradicts File-Muriel's (2007) recent study of another dialect of Colombia: Barranquilla. That author employs an elicitation task with 33 university students between the ages of 20 and 26 who read sentences extracted from editorials printed in the *El Tiempo* newspaper. He employed this experimental design in order to solicit tokens of word-internal, syllable-final /s/ in all possible phonological contexts, something difficult to do with spontaneous speech, as most tokens of word-internal, syllable-final /s/ are followed by /t/. One of File-Muriel's conclusions is that "lexical frequency is the single most significant factor in the prediction of *s*-lenition in this variety and style of Spanish" (p. 84). He bases this conclusion on the fact that a multivariate analysis selected lexical frequency first among the significant factor groups. However, when the factor weight ranges (File-Muriel, 2007, p. 83) are used to create a hierarchy of significant factor groups instead of using the order in which the significant groups were selected, as File-Muriel does, the lexical frequency falls to second place in the hierarchy of significant factor groups. In this respect, File-Muriel's Barranquilla results parallel the Cali results presented in this paper in word-internal position, namely, the following segment (manner of articulation in File-Muriel's study) is more influential than lexical frequency.

Despite this, both File-Muriel (2007) and this paper provide strong evidence that the lexical frequency of the word in which /s/ occurs significantly conditions the reduction of word-internal, syllable-final /s/ in Colombian Spanish. Additionally, this paper shows that lexical frequency also conditions word-final /s/ reduction, but at substantially lower rates than in word-internal position. In the end, this paper contributes to the growing body of research that shows that the usage-based factor of lexical frequency must be accounted for in any study of syllable- and word-final /s/ reduction in Spanish.

References

Alba, O. (1982). Función del acento en el proceso de elisión de la /s/ en la República Dominicana. In O. Alba (Ed.), *El español del Caribe* (pp. 15-26). Santiago de los Caballeros: Universidad Católica Madre y Maestra.

Alba, O. (2003). *¿Cómo hablamos los dominicanos?* Santo Domingo: Grupo León Jimenes. Retrieved May 2007, from {http://www.glj.com.do/?s=bv-cc-hablamos}.

Amparo Morales, M. V. (1990). *El habla culta de San Juan.* Río Piedras, Puerto Rico: Editorial de la Universidad de Puerto Rico.

Barrenechea, A. M. (1987). *El habla culta de la ciudad de Buenos Aires: materiales para su estudio.* Buenos Aires: Universidad Nacional de Buenos Aires, Facultad de Filosofía y Letras, Instituto de Filología y Literaturas Hispánicas "Dr. Amado Alonso."

Brown, E. L. (2004). *The reduction of syllable initial /s/ in the Spanish of New Mexico and Southern Colorado: A usage-based approach.* Unpublished doctoral dissertation, University of New Mexico, Albuquerque.

Brown, E. L., & Torres Cacoullos, R. (2002). ¿Qué le vamoh aher?: Taking the Syllable out of Spanish /s/ Reduction. In D. Johnson & T. Sanchez (Eds.), *University of Pennsylvania Working Papers in Linguistics: Papers from NWAV 30* (pp. 17-32). Philadelphia: University of Pennsylvania Press.

Brown, E. L. & Torres Cacoullos, R. (2003). Spanish /s/: A different story from beginning (initial) to end (final). In R. Núñez-Cedeño, L. López, & R. Cameron (Eds.), *A Romance perspective in language knowledge and use. Selected papers from the 31st linguistic symposium of Romance languages (LSRL)* (pp. 22-38). Amsterdam: John Benjamins.

Bybee, J. (2000). The phonology of the lexicon: evidence from lexical diffusion. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 65-85). Stanford, CA: Center for the Study of Language and Information.

Bybee, J. (2001). *Phonology and Language Use.* Cambridge: Cambridge University Press.

Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language variation and change, 14,* 261-90.

Bybee, J. (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language*, 82, 711-33.

Caravedo, R. (1989). *El español de Lima: materiales para el estudio del habla culta.* Lima: Pontificia Universidad Católica del Perú, Fondo Editorial.

Esgueva, M., & Cantarero, M. (1981). *El habla de la ciudad de Madrid: materiales para su estudio.* Madrid: Consejo Superior de Investagiciones Científicas, Instituto "Miguel de Cervantes."

Ferguson, C. A. (1990). From esses to atiches: Identifying pathways of diachronic change. In W. Croft, K. Denning, & S. Kemmer (Eds.), *Studies in Typology and Diachrony for Joseph H. Greenberg* (pp. 59-78). Amsterdam: John Benjamins.

File-Muriel, R. (2007). *A Study Of Lenition: The Role Of Lexical Frequency And Phonetic Context In The Weakening Of Lexical /s/ In The Spanish Of Barranquilla.* Unpublished doctoral dissertation, Indiana University, Bloomington.

Fontanella de Weinberg, M. B. (1973). Comportamiento ante –s de hablantes femeninos y masculinos del español bonaerense. *Romance Philology*, 27, 50-58.

Hooper, J. B. (1976). Word frequency in lexical diffusion and the source of morphophonological change. In W. Christie (Ed.), *Current Progress in Historical Linguistics* (pp. 96-105). Amsterdam: North Holland.

Hualde, J. I. (2005). *The Sounds of Spanish.* Cambridge & New York: Cambridge University Press.

Instituto Caro y Cuervo. (1986). *El habla de la ciudad de Bogotá: materiales para su estudio.* Bogotá: Publicaciones del Instituto Caro Y Cuervo LXXV. Imprenta Patriótica del Instituto Caro Y Cuervo.

Labov, W. (1984). Field Methods of the Project on Linguistic Change and Variation. In J. Baugh & J. Sherzer (Eds.), *Language in Use: Readings In Sociolinguistics* (pp. 28-53). Englewood Cliffs, NJ: Prentice-Hall.

Lipski, J. (1994). *Latin American Spanish.* New York: Longman.

Lope Blanch, J. M. (1971). *El habla de la Ciudad de México: materiales para su estudio.* México, D. F.: Instituto de Investigaciones Filológicas, Universidad Nacional Autónoma de México.

Marcos Marín, F. (1992). Corpus oral de referencia de la lengua española contemporánea. Madrid: Universidad Autónoma de Madrid. Accessed at http://elvira.lllf.uam.es/does_es/corpus/corpus.html.

Méndez Dosuna, J. (1985). La duración de S en los grupos SP, ST, SK: A propósito del orden regular de difusión en algunos cambios fonéticos. In J. L. Melena (Ed.), *Symbolae Ludovico Mitxelena Septuagenario Oblatae* (pp. 647-655). Vitoria, Spain: Universidad del País Vasco.

Milroy, L., & Gordon, M. (2003). *Sociolinguistics: Method and Interpretation.* Malden, MA: Blackwell.

Minnick Fox, M. A. (2006). *Usage-based Effects in Latin American Spanish Syllable-final /s/ Lenition.* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Penny, R. (2000). *Variation and change in Spanish.* Cambridge: Cambridge University Press.

Penny, R. (2002). *A History of the Spanish Language.* 2nd edition. Cambridge: Cambridge University Press.

Phillips, B. S. (2006). *Word Frequency and Lexical Diffusion.* NY: Palgrave Macmillan.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency and the Emergence of Linguistic Structure* (pp. 137-157). Amsterdam: John Benjamins.

Piñeros, C. (2006). The Phonology of Nasal Consonants in Five Spanish Dialects. In F. Martínez-Gil & S. Colina (Eds.), *Optimality-Theoretic Studies in Spanish Phonology* (pp. 146-171). Amsterdam & Philadelphia: John Benjamins.

Rabanales, A. & Contreras, L. (1979). *El habla culta de la ciudad de Santiago: materiales para su estudio.* Santiago de Chile: Editorial Universitaria, Anejo Nº 2 del Boletín de Filología, Facultad de Filosofía y Letras, Departamento de Lingüística y Filología.

Rosenblat, A. (1979). *El habla culta de Caracas: materiales para su estudio.* Caracas: Ediciones de la Facultad de Humanidades y Educación, Instituto de Filología Andrés Bello, Universidad Central de Venezuela.

Sankoff, D. (1988). Variable rules. In U. Ammon, N. Dittmar, & K. J. Mattheier (Eds.), *Sociolinguistics: an international handbook of the science of language and society* (pp. 984–997). Berlin: Walter de Gruyter.

Sankoff, D., Tagliamonte, S., & Smith, E. (2005). Goldvarb X: A variable rule application for Macintosh and Windows [Computer software]. Department of Linguistics, University of Toronto. Retrieved from http://individual.utoronto.ca/tagliamonte/goldvarb.htm

Schuchardt, H. (1885). *Über die Lautgesetze: gegen die Junggrammatiker.* Oppenheim, Berlin (Reprinted and translated in Vennemann, T. & Wilbur, T. 1972).

Tagliamonte, S. (2006). *Analysing Sociolinguistic Variation.* New York: Cambridge University Press.

Terrell, T. D. (1981). Diachronic reconstruction by dialect comparison of variable constraints: s-aspiration and deletion in Spanish. In D. Sankoff and H. Cedergren (Eds.), *Variation Omnibus*, pp. 115-124. Edmonton, Alberta: Linguistic Research, Inc.

Trask, R. L. (1996). *Historical Linguistics.* New York: St. Martin's Press.

Travis, Catherine E. (2005). *Discourse markers in Colombian Spanish: A study in polysemy.* Berlin / New York: Mouton de Gruyter.

Vennemann, T. & Wilbur, T. (1972). *Schuchardt, the Neogrammarians, and the Transformational Theory of Phonological Change*. Frankfurt: Athenäum Yerlag.

Zamora Vicente, A. (1970). *Dialectología española*, 2nd ed. Madrid: Editorial Gredos, S. A.

Selected Proceedings of the
11th Hispanic Linguistics Symposium

edited by Joseph Collentine,
Maryellen García, Barbara Lafford,
and Francisco Marcos Marín

Cascadilla Proceedings Project   Somerville, MA   2009

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Brown, Earl K. 2009. The Relative Importance of Lexical Frequency in Syllable- and Word-Final /s/ Reduction in Cali, Colombia. In *Selected Proceedings of the 11th Hispanic Linguistics Symposium*, ed. Joseph Collentine et al., 165-178. Somerville, MA: Cascadilla Proceedings Project.

or:

Brown, Earl K. 2009. The Relative Importance of Lexical Frequency in Syllable- and Word-Final /s/ Reduction in Cali, Colombia. In *Selected Proceedings of the 11th Hispanic Linguistics Symposium*, ed. Joseph Collentine et al., 165-178. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2211.