USAGE-BASED EFFECTS IN LATIN AMERICAN SPANISH SYLLABLE-FINAL /s/ LENITION

Michelle Annette Minnick Fox

A DISSERTATION

in

Linguistics

Presented to the Faculties of the University of Pennsylvania in Partial

Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2006

____
Mark Liberman
Supervisor of Dissertation

____
Eugene Buckley
Graduate Group Chairperson

Acknowledgments

I am very grateful to my advisor, Mark Liberman, for all of his support and advice in this project, and to committee members Gene Buckley and Gillian Sankoff for their insightful comments and suggestions.

Many others have been kind enough to give me help and comments on this work over the past few years. In particular, I thank Rolf Noyer, Bill Labov, Pam Beddor, Susan Garrett, Terri Lander, and Corey Miller, for their insights. I thank Susan Garrett and Doug Chavez for performing the coding of the CallHome corpus. I thank Kazuaki Maeda for all of his help with technical issues.

I am fortunate to have started at Penn with a wonderful group of classmates. I thank Jason Baldridge, Atissa Banuazizi, Patricia Chow, Ron Kim, Eon-Suk Ko, Masato Kobayashi, Eleni Miltsakaki, Kieran Snyder, and Alexander Williams for their camaraderie. For countless enjoyable hours in the phonetics lab, I thank John Bell, Susan Garrett, Eon-Suk Ko, Terri Lander, Kazuaki Maeda, and Amanda Seidl. Several students who started ahead of me have been an inspiration, particularly Corey Miller, Sean Crist, and Beth-Ann Hockey.

It has often been challenging to work on the dissertation from Michigan. I thank the Phondi group at the University of Michigan linguistics department, and especially Pam Beddor, José Benkí, and San Duanmu for their hospitality.

I thank my parents who have been constantly supportive of my pursuit of the PhD. This work would not have been possible without continuous motivation and help from my mother.

My deepest gratitude is to my husband, Jason, who has shown unsurpassed patience throughout all phases of graduate school, and to my three children born while I have been working on this dissertation, Andrea, Jeremy, and Brandon. Their love and enthusiasm have been an inspiration to me.

ABSTRACT

USAGE-BASED EFFECTS IN LATIN AMERICAN SPANISH

SYLLABLE-FINAL /s/ LENITION

Michelle Annette Minnick Fox

Supervisor: Mark Liberman

Previous studies have identified factors that contribute to the weakening and deletion of syllable-final /s/ in Latin American Spanish, including the dialect and sex of the speaker, the phonetic environment and grammatical status of the /s/, and functional considerations. Proponents of a usage-based model of language claim that the structure of language is shaped by how it is used, and therefore factors such as word frequency, word predictability, and the context that words appear in have an effect on the form of language. This dissertation investigates the extent to which these usage-based factors contribute to syllable-final /s/ lenition in addition to the previously identified factors. Automated speech recognition methods were used to code three dependent variables for a corpus of over 50,000 tokens of syllable-final /s/: deletion or retention of /s/, duration of retained /s/, and the spectral center of gravity of retained /s/. Multiple regression was performed for each of the dependent variables, on all of the data combined and on several subsets of the data. For each multiple regression, usage-based factors were added to a base model to determine which of them improve the model. Word frequency and word predictability based on the following word both have an effect in the expected direction, with more frequent and more predictable words having higher levels of lenition. Word predictability based on the preceding word has the opposite effect, with more predictable words having lower levels of lenition. The phonetic context that words appear in most frequently also has an effect, with words that are more often followed by a consonant having more advanced lenition, even after taking into consideration the actual phonetic context. These usage-based factors contribute both to the categorical process of deletion and the gradient processes of shortening and weakening of articulation. This data supports the claim that these usage-based variables form part of the speaker's knowledge and that speakers have knowledge of low-level phonetic detail. This study suggests that the extent of lenition may be determined both within the lexical entry as in an exemplar model and by processing during production.

Table of Contents

Acknowledgments ii
ABSTRACT iv
Table of Contents vi
1 Introduction 1
2 Description of Syllable-final /s/ lenition 4
2.1 Extralinguistic factors 4
2.1.1 Dialect 4
2.1.2 Age, Sex, Social Class 8
2.1.3 Speech Style 10
2.2 Linguistic factors 12
2.2.1 Phonetic environment 13
2.2.2 Grammatical category 16
2.2.3 Functional considerations 18
2.2.4 Number of syllables in the word 21
2.2.5 Lexical exceptions 21
2.3 Summary 22
3 Usage-based accounts of variation 24
3.1 Frequency 25
3.2 Word probability 28
3.3 Exemplar theory and lexical representations 32
3.4 Usage-based factors in relation to syllable-final /s/ lenition 36
4 Data Collection 38
4.1 Spanish pronunciation lexicon 38
4.2 CallHome Spanish 40
4.2.1 The LDC Hub5-LVCSR CALLHOME Spanish Speech Corpus 40
4.2.2 CallHome Spanish Data Coding 43
4.2.3 Categorical Data 46
4.2.4 Data Accuracy 47
4.3 Broadcast News 52
4.3.1 The LDC 1997 HUB-4 Spanish Speech Corpus 52
4.3.2 Broadcast News Data Labeling 54
4.3.2.1 Data Preparation 55
4.3.2.2 Creation of Monophone HMM acoustic models 57
4.3.2.3 Issues encountered during the HTK coding procedure 59
4.3.3 Categorical Data 61
4.3.4 Acoustic Measurements 63
4.3.4.1 Duration 66
4.3.4.2 Spectral Moments of /s/ 71
4.3.4.3 Intensity 75
4.3.4.4 Preceding vowel formants 78
4.3.4.5 Summary 83
4.3.5 Data Accuracy 85
4.3.5.1 Categorical Data 85

4.3.5.2 Measurement Data....88
4.4 Summary....94
5 Independent Variables....96
5.1 Linguistic and extra-linguistic factors....96
5.1.1 Speaker....96
5.1.2 Following segment....100
5.1.3 Preceding vowel....105
5.1.4 Stress....106
5.1.5 Grammatical Status....111
5.1.6 Number of syllables....115
5.1.7 Speech rate....117
5.1.8 Lexical identity....119
5.2 Usage-based Independent Variables....121
5.2.1 Calculation of usage-based variables from the corpora....122
5.2.2 Word frequency....123
5.2.3 Relative frequencies....128
5.2.4 Conditional probability....130
5.2.5 Repetition....134
5.2.6 Local deletion rate....136
5.2.7 Degree of ambiguity of the word....139
5.2.8 Phonetic context proportion for individual words....141
6 Data Analysis....144
6.1 Model for all tokens of syllable-final /s/....144
6.1.1 Categorical deletion data....144
6.1.2 Continuous duration data....151
6.1.3 Continuous Spectral COG data....154
6.2 Model for Content Words and Function Words....156
6.2.1 Deletion in Content Words....157
6.2.2 Deletion in Function Words....159
6.2.3 /s/ duration in content words....161
6.2.4 /s/ duration in function words....164
6.2.5 Spectral COG in content words....165
6.2.6 Spectral COG in function words....167
6.3 Word-final /s/ followed by a consonant or vowel....168
6.3.1 Deletion for word-final /s/ followed by a consonant....169
6.3.2 Deletion for word-final /s/ followed by a vowel....171
6.3.3 /s/ duration for word-final /s/ followed by a consonant....172
6.3.4 /s/ duration for word-final /s/ followed by a vowel....174
6.3.5 Spectral COG for word-final /s/ followed by a consonant....176
6.3.6 Spectral COG for word-final /s/ followed by a vowel....177
6.4 Plural -s....179
6.4.1 Deletion of plural -s....179
6.4.2 /s/ duration of plural -s....182
6.4.3 Spectral COG of plural -s....183
6.5 Verbal ending for first person plural (-mos)....184

6.5.1 Deletion in -mos 185
6.5.2 /s/ duration in -mos 186
6.5.3 Spectral COG 187
6.6 Word-internal syllable-final /s/ 188
6.6.1 Deletion in word-internal /s/ 189
6.6.2 /s/ duration in word-internal -/s/ 190
6.6.3 Spectral COG in word-internal -/s/ 191
7 Discussion 193
7.1 Word frequency 193
7.2 Other measures of word probability 196
7.3 Repetition 200
7.4 Phonetic context proportion 201
7.5 Local deletion rate 202
7.6 Conclusion 203
References 207

Tables

Table 1. Realization of syllable-final /s/ in Latin American Spanish dialects (data from Lipski 1984) 6

Table 2. Percentage of syllable-final /s/ aspiration or deletion for news broadcasts and speech. Data taken from Lipski (1983a, 1985).......... 12

Table 3. Word-final /s/ deletion data for Cuban and Puerto Rican speakers, separated by grammatical category and lexical exceptions (taken from Terrell 1977). 18

Table 4. Number of speakers and number of syllable-final /s/ tokens in the CallHome Spanish corpus.......... 42

Table 5. Number of speakers having the number of syllable-final /s/ tokens in each range in the CallHome Spanish corpus. 42

Table 6. Breakdown of tokens that could not be coded for the CallHome corpus. 45

Table 7. Percent of tokens with each code for the CallHome Corpus. 47

Table 8. Percent of tokens with each code for the CallHome Corpus, using the dialect categories "coastal" and "interior". 47

Table 9. Cross tabulation of codes within and across coders for tokens coded more than once.......... 48

Table 10. Cross tabulation of retained vs. deleted codes within and across coders for tokens coded more than once. 49

Table 11. Reliability measurement κ for the CallHome corpus.......... 50

Table 12. Number of speakers and number of syllable-final /s/ tokens in the Spanish Broadcast News corpus.......... 53

Table 13. Number of speakers having the number of syllable-final /s/ tokens in each range in the Spanish Broadcast News corpus.......... 53

Table 14. Percent of tokens with each coding by dialect for the Broadcast News Corpus. 61

Table 15. Average duration of the immediately preceding vowel, the syllable-final /s/, and the following segment, for all automatically generated data.......... 71

Table 16. Summary of measurement data. 85

Table 17. Comparison of HTK-generated codes and hand-labeled codes. 86

Table 18.Reliability measurement κ for the Broadcast News corpus. 87

Table 19. Comparison of average values of acoustic measurements using HTK and Hand-labeling. 94

Table 20. Percent deletion for the CallHome corpus by dialect type and sex. 99

Table 21. Percent deletion and measurement data for the Broadcast News corpus by dialect type and sex. 99

Table 22. Percent deletion and measurements for speakers in the Broadcast News corpus having more than 1,000 tokens of syllable-final /s/. 100

Table 23. Percentage deletion in the CallHome corpus separated by position in the word and following segment type. 102

Table 24. Percentage deletion and measurement data in the Broadcast News corpus separated by position in the word and following segment type. 103

Table 25. Percent deletion by the place of articulation of the following consonant for the CallHome corpus. 104

Table 26. Percent deletion and measurement data by the place of articulation of the following consonant for the Broadcast News corpus. 104

Table 27. Percent deletion in the CallHome corpus, by preceding vowel. 105

Table 28. Percent deletion, average /s/ duration and average Spectral COG in the Broadcast News corpus, by preceding vowel. 106

Table 29. Percent deletion for the CallHome corpus by following stress. 108

Table 30. Percent deletion and measurement data for the Broadcast News corpus by following stress. 109

Table 31. Percent deletion by syllable stress type in the CallHome corpus. 110

Table 32. Percent deletion and measurement data by syllable stress type in the Broadcast News corpus. 110

Table 33. Percent Deletion in the CallHome corpus by grammatical category. 113

Table 34. Percent deletion and measurement data in the Broadcast News corpus by grammatical category. 113

Table 35. Percent deletion of word-final /s/ in the CallHome corpus, separated by content words and function words. 114

Table 36. Percent deletion and measurement data of word-final /s/ in the Broadcast News corpus, separated by content words and function words. 115

Table 37. Percent deletion by the number of syllables in the word and word class, for word-final tokens in the CallHome corpus. 116

Table 38. Percent deletion and measurement data by the number of syllables in the word, for word-final tokens in the Broadcast News corpus....................................... 116

Table 39. Percent deletion by speech rate (syllables/second) for tokens of -/s/ that are not phrase-final the CallHome corpus. 119

Table 40. Percent deletion and measurement data by speech rate (syllables/second) for tokens of -/s/ that are not phrase-final in the Broadcast News corpus................ 119

Table 41. Distribution of unique words and total number of tokens by word frequency for the Broadcast News corpus. 125

Table 42. Deletion rate for the CallHome corpus by word frequency. 127

Table 43. Deletion rate and measurement data for the Broadcast News corpus by word frequency. 127

Table 44. Deletion rate by bigram relative frequency with the following word ($w_i w_{i+1}$) for the CallHome corpus. 130

Table 45. Data by bigram relative frequency with the following word ($w_i w_{i+1}$) for the Broadcast News corpus... 130

Table 46. Deletion rate by word frequency and forward conditional bigram probability for the CallHome corpus. 133

Table 47. Deletion rate and measurement data by word frequency and forward conditional bigram probability for the Broadcast News corpus......................... 134

Table 48. Percent deletion by repetition number for the CallHome corpus. 135

Table 49. Percent deletion and measurement data by repetition number for the Broadcast News corpus. 136

Table 50. Deletion rate as a function of the local deletion rate for the CallHome corpus. 138

Table 51. Deletion rate as a function of the local deletion rate for the Broadcast News corpus... 138

Table 52. Deletion rate as a function of the local deletion rate for the speaker Jorge Ramos in the Broadcast News corpus.. 138

Table 53. Deletion data for plural /s/, based on whether deletion of /s/ would lead to ambiguity for the CallHome corpus..139

Table 54. Deletion and measurement data for plural /s/, based on whether deletion of /s/ would lead to ambiguity for the Broadcast News corpus. 140

Table 55. Deletion data for plural /s/ by the word singular-to-plural ratio for the CallHome corpus. 140

Table 56. Deletion and measurement data for plural /s/ by the word singular-to-plural ratio for the Broadcast News corpus...141

Table 57. Percent deletion by following segment type and how often the word appears in pre-consonant position, for the CallHome corpus. 142

Table 58. Percent deletion by following segment type and how often the word appears in pre-consonant position, for the Broadcast News corpus..................................143

Table 59. Binary logistic regression for deletion data...150

Table 60. Multiple regression for /s/ duration data. 154

Table 61. Multiple regression for spectral COG data. 156

Table 62. Binary logistic regression for deletion data for word-final /s/ in content words. 159

Table 63. Binary logistic regression for deletion data for word-final /s/ in function words. 161

Table 64. Multiple regression for /s/ duration data for word-final /s/ in content words. 163

Table 65. Multiple regression for /s/ duration data for word-final /s/ in function words. 165

Table 66. Multiple regression for spectral COG in word-final /s/ in content words. 166

Table 67. Multiple regression for spectral COG in word-final /s/ in function words.....168

Table 68. Binary logistic regression for deletion data for word-final /s/ followed by a consonant...170

Table 69. Binary logistic regression for deletion data for word-final /s/ preceding a vowel...172

Table 70. Multiple regression for /s/ duration for word-final /s/ followed by a consonant. 174

Table 71. Multiple regression for /s/ duration data for word-final /s/ preceding a vowel. 175

Table 72. Multiple regression for spectral COG data in word-final /s/ preceding a consonant........ 177

Table 73. Multiple regression for spectral COG in word-final /s/ preceding a vowel. ... 178

Table 74. Binary logistic regression for deletion data for plural *–s*. 181

Table 75. Multiple regression for /s/ duration for plural *–s*. 183

Table 76. Multiple regression for spectral COG in plural *–s*. 184

Table 77. Binary logistic regression for deletion data for the 1st person plural morpheme -*mos*. 186

Table 78. Multiple regression for /s/ duration data for the 1st person plural morpheme -*mos*. 187

Table 79. Multiple regression for spectral COG for the 1st person plural morpheme -*mos*. 188

Table 80. Binary logistic regression for deletion data for word-internal /s/. 190

Table 81. Multiple regression for /s/ duration data for word-internal /s/. 191

Table 82. Multiple regression for spectral COG for the word-internal -/s/. 192

Table 83. Direction of effect of higher word frequency on /s/ deletion, duration, and spectral COG in each of the subsets tested. 194

Table 84. Direction of effect of higher probability based on following word (bigram frequency, bigram mutual information, conditional bigram probability) in each of the subsets tested. 198

Table 85. Direction of effect of higher probability based on preceding word(s) (bigram/trigram frequency, conditional bigram/trigram probability) in each of the subsets tested. 199

Table 86. Direction of effect of higher word repetition number in each of the subsets tested. 200

Table 87. Direction of effect of consonant proportion in each of the subsets tested. 202

Table 88. Direction of effect of higher local deletion rate on /s/ deletion, duration, and Spectral COG in each of the subsets tested. '+' indicates more advanced lenition. 203

Figures

Figure 1. Boxplot of /s/ duration for hand-labeled tokens with retained /s/. 67

Figure 2. Duration of the immediately following segment, by segment type and hand-labeled code. 69

Figure 3. Duration of the immediately preceding vowel by hand-labeled code. 70

Figure 4. Spectral COG (in Hz) by hand-labeled code, for tokens with retained /s/. 74

Figure 5. Spectral Second Moment (in Hz) by hand-labeled code, for tokens with retained /s/. 75

Figure 6. /s/ Intensity (measured in dB) with respect to the immediately preceding vowel. 77

Figure 7. Boxplot of F1 of preceding vowel [e] by speaker for each of the hand labeled codes. 79

Figure 8. Boxplot of F2 of preceding vowel [e] by speaker for each of the hand labeled codes. 80

Figure 9. Boxplot of F1 of preceding vowel [o] by speaker for each of the hand labeled codes. 81

Figure 10. Boxplot of F1 of preceding vowel [o] by speaker for each of the hand labeled codes. 82

Figure 11. Duration and Spectral COG for hand-labeled tokens from the Broadcast News Corpus. 84

Figure 12. Comparison of label and duration (automatically generated vs. hand-labeled), for automatically generated label [s]. 89

Figure 13. Comparison of label and duration (automatically generated vs. hand-labeled), for automatically generated label [h]. 90

Figure 14. Comparison of label and duration (automatically generated vs. hand-labeled), for automatically generated label [z]. 91

Figure 15. Comparison of spectral COG calculated from automatically generated time alignments and hand labeled time alignments. 92

Figure 16. Comparison of /s/ intensity calculated from automatically generated time alignments and hand labeled time alignments. 93

Figure 17. Deletion proportion for 195 speakers in the CallHome corpus....................... 97

Figure 18. Deletion proportion for 201 speakers in the Broadcast News corpus. 98

Figure 19. Deletion proportion by word in the CallHome corpus. 120

Figure 20. Deletion proportion by word in the Broadcast News corpus. 121

1 Introduction

Weakening of syllable-final /s/ in Spanish is a wide-spread phenomenon. It occurs to some extent in most dialects of Spanish spoken in Latin America. Syllable-final /s/ lenition in Spanish has been of great interest to sociolinguists due to the great variability with which this process occurs over a large number of dialects.

The result of lenition is variable and may be the loss of the oral constriction – *aspiration*, where the underlying /s/ becomes the placeless fricative [h] – or it may be total deletion of the segment. Initially, aspiration occurred only before consonants, and later spread to other word-final contexts in order to reduce allomorphy (Terrell 1981:121). The results of aspiration and deletion are shown in (1).

(1) a. __C: word-internally before a consonant

ministro à [ministro], [minihtro], [minitro] 'minister'

b. __##C: word-finally before a consonant

los libros à [los], [loh], [lo] 'the books'

c. __##V: word-finally, before a vowel

los estados à [los], [loh], [lo] 'the states'

d. __||: before a pause

los estados *unidos* || à [uniðos], [uniðoh], [uniðo] 'the United States'

Although aspiration and deletion appears from this description to be a categorical process that has three surface forms /s/ à [h] à Ø, the process of /s/ weakening can also result in voicing assimilation when followed by a voiced consonant, so that /s/ may be pronounced [z] or as a voiced variant of [h]. A weakened underlying syllable-final /s/ may be also realized with little or no aspiration, but with the following consonant lengthened to the point of being a geminate consonant (Terrell 1977).

Numerous social and linguistic factors that favor or resist the weakening of syllable-final /s/ have been identified by other researchers. Social factors include dialect, sex, social class, and speech style. Linguistic factors that have previously been explored include the phonetic environment, grammatical category of the word and functional factors.

In a usage-based model of language, the structure of language is claimed to be primarily shaped by how the language is used. In this type of framework, factors such as the frequency of particular words or word sequences, the probability of a word given surrounding words, and the context in which words are used all are considered to be important in forming the speaker's linguistic knowledge.

This study tests whether these usage-based factors contribute to a model of syllable-final /s/ lenition that takes into consideration social and linguistic factors that have been previously identified. Two large speech corpora from the Linguistic Data Consortium (LDC) have been labeled for syllable-final /s/: the CallHome Spanish corpus, which consists of informal telephone conversations, and the Hub4 Spanish Broadcast News corpus, which consists of more formal speech from news broadcasts. The labeling for the CallHome Spanish corpus was done by human listeners who coded each token of syllable-final /s/ as an [s], [h] or deletion. For the Spanish Broadcast News corpus, automated methods were used to train acoustic models for a speech recognizer which then generated labels for the speech files and calculated several different measurements.

Using speech corpora such as the ones used in this study has several advantages. These corpora are very large and because they have been orthographically transcribed, automated methods can be used to identify all instances of the phenomenon under study. For this study, over 70,000 tokens of syllable-final /s/ have been identified and analyzed from the two corpora. This large number of tokens is necessary to evaluate interactions between the different factors that promote syllable-final /s/ lenition. In addition, the corpora are available to the research community at large, which means that other researchers will be able to verify the findings of this study and perform other studies, using the same speech data used in this study. Finally, using an existing corpus means that the recording conditions are more transparent.

This dissertation is organized as follows. Chapter 2 consists of a literature review of syllable-final /s/ lenition. In Chapter 3, I discuss different usage based theories put forward by various researchers and report on some of their findings. Chapter 4 is a description of the data collection process and methodology used. In Chapter 5, I report my results by independent variable. Chapter 6 consists of multivariate statistical analyses of the data using Minitab 14, a software package that has the capability of sophisticated data analysis. Finally, in Chapter 7, I discuss the significance of my results and the statistical analyses in relation to other researchers' findings.

2 Description of Syllable-final /s/ lenition

Weakening and deletion of syllable-final /s/ in Spanish is attested as early as the 16th century (Lapesa 1980, as cited by Lipski 1984). The exact origins of syllable-final /s/ lenition are unclear, but it most likely began in the southern region of Andalucia, and then spread to other regions of Spain as well as the port areas of Latin America, first in the port areas of the Caribbean (Terrell 1981). Today lenition occurs to some extent in nearly all regions of Latin America.

The process of lenition has been widely studied, both across dialects and within dialects. Studies on dialectal differences, such as work done by Canfield and Lipski, generally explore whether or not lenition occurs in a given dialect, but do not extensively explore the individual constraints on lenition using quantitative methods. Many of the more extensive quantitative studies which look at the contributions of individual constraints on lenition, such as those performed by Poplack, Terrell, and Lipski, tend to focus on the Spanish of one dialect.

2.1 Extralinguistic factors

2.1.1 Dialect

Lenition of syllable-final /s/ is one of the main phonological variables that have been used to classify dialects of Latin American Spanish. For example, Canfield (1981) presents dialect classification maps for each country of Latin America based on the pronunciation of syllable-final /s/ as well as several other linguistic variables. Countries of the Caribbean (notably Cuba, the Dominican Republic, Puerto Rico, Venezuela, Honduras, and Nicaragua) are particularly well known for s-aspiration, and many studies have been conducted of these dialects. According to Canfield's description, lenition also occurs in many other regions of Latin America as well: El Salvador, Panama, Chile, Paraguay, Uruguay, much of Argentina, part of Bolivia, and small parts of Peru, Ecuador, and Colombia. Syllable-final /s/ is mostly retained only in the inland highland areas of Mexico, and Central and South America, and even there, pre-consonantal /s/ is subject to some lenition (Lipski 1985).

Table 1 shows the percentages of syllable-final /s/ pronounced as [s], [h] and deleted for a number of different dialects of Latin American Spanish, reported in Lipski (1984, 1986). Lipski's data is taken from many studies performed by a variety of researchers.[1] As described later, syllable-final /s/ lenition is highly influenced by the phonetic context in which it occurs, and so researchers have reported separate aspiration and deletion rates for each of the main contexts: (1) word-internal, pre-consonantal; (2) word-final, with a following consonant; (3) word-final, with a following pause; (4) word-final, with a following stressed vowel; and (5) word-final, with a following unstressed vowel. The data presented in Table 1 is largely in keeping with Canfield's description. The Caribbean countries generally have the most advanced lenition rates; Cuba, Puerto Rico, and Nicaragua have high percentages of aspiration, Venezuela has approximately equal (high) rates of aspiration and deletion, and the Dominican Republic has mostly deletion. Among the countries of the Caribbean, Honduras has the highest rate of /s/

---

[1] Data for Cuba is taken from Terrell (1979); for Puerto Rico, Terrell (1978b); for Argentina, Terrell (1978a); for Panama, Cedergren (1978); for Honduras, Lipski (1983b) and Lipski (1986a); for Nicaragua, Lipski (1984b); for the Dominican Republic, Alba (1982) and Núnez Cedeño (1980); for El Salvador, Lipski (1982).

pronounced as [s] (63% in word-internal position and 19% in word-final, pre-consonantal position). The other countries listed by Canfield as having advanced lenition also show relatively low levels of /s/ retained as [s] in Lipski's table: El Salvador (10% [s] in word-final, pre-consonantal position), Panama (4%), Chile (4%), Paraguay (2%), Uruguay (4%), and Argentina (11%). Even the countries that have been reported to be more conservative such as Costa Rica and Guatemala show a fair amount of aspiration (29% and 30% respectively) in the word-final, pre-consonantal position.

	sC			s#C			s##			s#V (str.)			s#V (unstr.)		
	s	h	Ø	s	h	Ø	s	h	Ø	s	h	Ø	s	h	Ø
Argentina	12	80	8	11	69	20	78	11	11	93	7	0	94	6	0
Chile	7	93	0.5	4	88	8	63	33	4	90	10	0	76	22	2
Costa Rica	92	8	0	69	29	2	96	4	0	98	2	0	98	2	0
Cuba	3	97	0	2	75	23	61	13	26	48	28	25	10	53	37
Dominican Republic	8	17	75	5	25	70	36	10	54	50	5	45	17	22	61
El Salvador	55	44	1	10	71	19	86	12	2	44	47	9	28	69	3
Guatemala	93	7	0	69	30	1	93	3	0	100	0	0	100	0	0
Honduras	63	34	3	19	58	23	83	15	2	90	10	0	61	38	1
Nicaragua	13	83	4	2	86	12	35	59	6	28	70	2	7	90	3
Panama	13	52	35	4	48	48	25	21	54	62	13	25	9	67	27
Paraguay	14	86	0	2	92	6	83	15	2	47	53	0	15	84	1
Peru	53	47	0	21	71	8	91	8	1	94	6	0	91	9	0
Puerto Rico	3	92	5	4	69	27	46	22	32	45	32	23	16	53	31
Uruguay	20	79	1	4	88	8	85	13	2	98	2	0	93	7	0
Venezuela	7	40	53	3	47	50	38	16	46	57	26	17	15	52	33

Table 1. Realization of syllable-final /s/ in Latin American Spanish dialects (data from Lipski 1984)

No data is presented for Mexico in Lipski (1984a), presumably because Mexican Spanish had not been well studied due to the fact that it is commonly thought to retain syllable-final /s/ as [s] nearly all the time. However, in a later study, Lipski (1994) states that in fact /s/ is not retained to the same extent in all of Mexico; the highest rates of /s/

retention occur in educated speech of Mexico City. /s/ lenition was originally more common in the coastal areas, but /s/ retention has been spreading from Mexico City to educated speech in other areas (Lipski 1994:223). Aspiration and deletion occur at the highest rates in Mexico along the coast (both Caribbean and Pacific), with composite rates of 20-30%.

As shown by Lipski's (1994) description of Mexican Spanish and Canfield's (1981) description of Bolivia, Peru, Ecuador, and Columbia, looking at the data by only one dialect per country over-simplifies the situation (see also Hundley 1987 for Peru). In many cases, different regions in the same country display very different characteristics with respect to syllable-final /s/. This occurs in particular when different regions are geographically separated, usually with the coastal areas having significantly higher rates of aspiration and deletion than the interior or mountainous regions.

It also appears that there are dialectal differences within Chile. Cepeda's (1995) description of Chilean Spanish agrees with Lipski's data for two of the environments (an underlying /s/ is generally pronounced [s] in syllable-final position before a stressed vowel and [h] before a non-continuant consonant). However, in contrast with Lipski's data showing that syllable-final /s/ is mostly retained as [s] before an unstressed vowel (76% of the time) and before a pause (63% or the time), Cepeda claims that for the dialect she is studying, syllable-final /s/ is mostly deleted in these contexts.

There is also conflicting data for Argentina. In her study of the speech of Bahía Blanca, a city in the Province of Buenos Aires, Fontanella de Weinberg (1974) reports that the two principal variants of word-final /s/ are [s] and deletion, with [h] occurring less than 1% of the time. In contrast, the data reported by Lipski indicates that [h] is the most common variant in both word-internal and word-final, pre-consonantal contexts (80% and 69%, respectively).

It is also important to note that there are other significant extralinguistic factors affecting syllable-final /s/ lenition that are not controlled for in the data provided by Lipski (1984, 1986), most importantly the speaker's social class and the speech style. Lipski's data compiles the results of different researchers using different methodologies, which raises the possibility that the data may not be exactly comparable.

2.1.2 Age, Sex, Social Class

A number of studies have looked at the effect of age, sex, or social class on syllable-final /s/ lenition. However, in most cases, the researchers have not attempted to study the effect of all three.

As for sex, a number of studies report that the process of lenition is more advanced in males than in females. In his description of Cuban and Puerto Rican Spanish, Terrell (1977) states that there is "some evidence" that lenition generally occurs more frequently among males, although he does not give quantitative numbers, and adds that there may not be a large effect of sex in these dialects. Valdivieso and Magaña. (1991) find that in the speech of Concepción, Chile, both males and females tend to weaken syllable final /s/, but males have more advanced lenition: women have 80% [h] and 11% deletion, while men have 64% [h] and 24% deletion. Cepeda (1995) also reports more deletion among males (44%) than among females (36%) in Chile. Similar results have been reported by Poplack (1980a) for Puerto Rican Spanish and by Cedergren (1978) for Panamanian Spanish.

There are somewhat conflicting results with regard to the effect of age on syllable-final /s/ lenition. Cepeda (1995) reports slightly more deletion in the youngest age group (14-19 years old) - 43% - than in the middle age group (24-48 years old) - 37% - or in the oldest age group (54-78 years old) - 38% -. In contrast, Cedergren (1978) reports that younger speakers in Panama exhibit less deletion than older speakers. Other studies, such as Poplack (1980b) for Puerto Rico and Valdivieso and Magaña (1991) for Concepción, have found little or no effect of age on syllable-final /s/ lenition.

Studies have shown that social class has an effect on syllable-final /s/ lenition, with speakers of lower social classes generally demonstrating more advanced levels of lenition. For Chilean Spanish, Cepeda (1995) finds lower rates of deletion in the highest social class (32%) and middle social class (35%) when compared to the lowest social class (50%). Similar effects have been reported by Cedergren (1978) for Panama, Alba (1982) for the Dominican Republic, Lopez Morales (1983) for Puerto Rico, and Lipski (1994) for Mexico. Terrell (1977) also notes more lenition in the lower classes in the Spanish of Cuba and Puerto Rico, but he states that social class does not have a large effect in either of these dialects.

Fontanella de Weinberg (1974) considers the effect of educational level as well as social class for Argentinean Spanish. She reports a large effect of the speakers' social class, with the highest social class having a much lower rate of deletion (13%) than the lowest social class (68%). She finds a further contribution of the speakers' educational level in conjunction with the social class, with only 6% deletion for speakers in both the highest social class and highest educational class, compared with 73% for speakers in both the lowest social class and lowest educational class.

While not exactly a study of educational level, Valdivieso and Magaña (1991) test whether there is an effect of profession type in speakers of Concepción. In this study, which is limited to 16 university professionals (thus all well educated), they find little effect of whether the speakers are involved in the humanities or a scientific discipline.

In the analysis presented in the current study, the speaker's sex will be considered. Because information on the social class and age of the speakers is not available for the two corpora I have used in the present study, the effect of these two factors cannot be evaluated.

2.1.3 Speech Style

Numerous studies of speech style for different dialects show that more formal speech has less lenition than more casual speech. Fontanella de Weinberg (1974) studies deletion of word-final /s/ for 4 different speaking styles (spontaneous, formal, reading, and word lists) by Argentinean speakers of 4 different social classes and 4 educational levels. She finds that all social classes and educational levels delete word-final /s/ most often in informal speech, followed by formal speech and then slightly less in the reading style. Word-final /s/ is deleted the least in word lists. Among speakers who had completed university, Fontanella de Weinberg finds very little deletion (2%) in read speech.

For Chile, both Cepeda (1995) and Valdivieso and Magaña (1991) find a higher rate of [s] in more careful speech, and more deletion in more informal speech. In read speech, Valdivieso and Magaña find little lenition before a vowel or pause (between 80 and 90% [s]) and 70% aspiration before a consonant. In spontaneous speech, there is about 60% aspiration before a vowel and 40% aspiration before a pause, and nearly 90% aspiration before a consonant.

For the dialects with advanced levels of lenition, the speech style appears to have less of an effect. Lipski (1985) states that for Cuban and Puerto Rican Spanish, syllable- and word-final /s/ undergo a high rate of aspiration or deletion in almost all speech styles. Terrell (1977:37) notes that in these dialects, /s/ is retained as [s] "only in the most formal styles," such as when reading or giving speeches.

According to Lipski (1983a, 1985) broadcast speech is at the extreme of speech styles. In general, Hispanic broadcasters use an affected style, including exaggerated intonation and word choice that avoids regional terms. Radio broadcasters tend to "strive for a maximally clear and precise diction, as free as possible from regionalisms of pronunciation" and to pronounce each letter in the written word (Lipski 1985:221). Lipski identifies 3 levels of broadcast speech style: news, music, and sports. The news style is the most formal and the sports style is the closest to conversational speech. Table 2 shows /s/ lenition data taken from Lipski (1983a, 1985) comparing the news broadcast style with conversational speech style across dialects. In this table, Lipski has combined the categories [h] and deletion, and the numbers indicate the percentage of /s/ that is weakened to either of these. All of the dialects show considerably more retention in the broadcast news style than in conversational style for word-internal and word-final pre-consonant position. In pre-pause and pre-vowel position, Lipski reports no aspiration or deletion in the broadcast news style in any of the dialects. He also notes that Cuban Spanish is highly variable in the broadcast style; some broadcasters have a low rate of /s/ aspiration and deletion, while others have a much higher rate. Lipski hypothesizes that the reason for the higher rate of /s/ aspiration and deletion in some Cuban Spanish broadcasts is to avoid a perceived social class distinction and to make the radio broadcasters sound more like their audience.

		sC	s#C	s##	s#V (str.)	s#V (unstr.)
Argentina	News	10	22	0	0	0
	Speech	88	89	22	7	6
Chile	News	1	7	0	0	0
	Speech	93	96	36	10	22
Cuba	News	25	20	0	0	0
	Speech	97	98	39	53	90
Dominican Republic	News	11	12	0	0	0
	Speech	91	98	65	49	85
El Salvador	News	9	26	0	0	0
	Speech	45	90	14	37	72
Honduras	News	20	45	0	0	3
	Speech	41	89	28	15	39
Nicaragua	News	3	17	0	0	0
	Speech	87	98	64	72	93
Panama	News	10	11	0	0	0
	Speech	95	95	66	53	80
Paraguay	News	3	17	0	0	0
	Speech	86	98	17	53	85
Puerto Rico	News	4	5	0	0	0
	Speech	94	96	54	55	84
Venezuela	News	2	12	0	0	0
	Speech	95	49	98	41	89

Table 2. Percentage of syllable-final /s/ aspiration or deletion for news broadcasts and speech. Data taken from Lipski (1983a, 1985).

.

2.2 Linguistic factors

Studies have been conducted of speakers of different Latin American dialects in order to identify linguistic constraints on deletion. These factors include: the phonetic environment of the syllable-final /s/, grammatical status of the /s/, functional considerations, and the number of syllables in the word.

2.2.1 Phonetic environment

The most important factor of phonetic environment on syllable final /s/ lenition is the type of following segment. As described above, while a word-internal syllable-final /s/ must necessarily be followed by a consonant, an underlying syllable-final /s/ that is word-final may be followed by a consonant, pause, or a vowel. Other aspects which have been shown to be important are position in the word and following stress. Because phonetic context has such a great impact, the data reported in Lipski (1984a) and shown in Table 1 above is separated into the five phonetic contexts that have been used by many researchers: (1) word-internal, pre-consonantal; (2) word-final, with a following consonant; (3) word-final, with a following pause; (4) word-final, with a following stressed vowel; and (5) word-final, with a following unstressed vowel. According to Lipski (1986b) and Terrell (1981), /s/ lenition started in the context before a consonant in word-final position, and has subsequently spread to the other positions in order to reduce allomorphy.

In Lipski's table, in the word-final position, in all dialects, there is the most deletion before a consonant. Even in the dialects with the most advanced lenition before a word-final consonant (Cuba, Puerto Rico, Venezuela, Chile, Panama and Uruguay), syllable-final /s/ is retained as [s] in some contexts. For example, in these dialects, there is generally more /s/ pronounced [s] before a stressed vowel than consonant (ranging from 28% for Nicaragua to 90% or more for Chile and Uruguay). The lenition rate for a following pause generally falls between the rate for a following consonant and a following stressed vowel. Among the countries with advanced lenition, only El Salvador and Paraguay have significantly less aspiration and deletion before a pause than before a vowel. Countries with the most [s] in a preconsonantal position (Costa Rica and Guatemala) have very little lenition before a pause.

Not many researchers divided this category up into subcategories but Valdivieso and Magaña (1991), in their study of the speech of Concepción, Chile, break down 'following consonant' into 3 categories: occlusives/affricates, resonants, and fricatives. There is the least deletion (8%) before occlusive/affricates, the most deletion before fricatives (23%), with resonants in the middle (16%).

Brown and Torres Cacoullos (2003) have also split up the following segment into several categories in their varbrul analysis of the Spanish of Chihuahua, Mexico. This dialect is rather unusual in that it actually has more lenition word-internally in syllable initial position (before a vowel: *ese* – 34%) than in syllable final position (*este* – 22%), and more lenition word-finally before a pause or vowel than before a consonant. For word-final /s/, Brown and Torres Cacoullos find that the following phonological environment is the most significant factor, with the most favoring environment being a liquid (63% aspirated or deleted), followed by /f,x,y/ (52%), nasals (54%), /a/ (53%), pause (54%), voiced consonants (50%), other vowels (44%), other consonants (37%), and the least favoring environments of voiceless stops (28%) and finally /s/ (22%).

The position in the word (word-internal or word-final) also has an impact on the lenition of a syllable-final /s/ followed by a consonant. Lipski's data shows that in general, dialects that have the highest rates of aspiration and deletion in word-final pre-consonant position tend also have high rates of lenition rates before a consonant in word-internal position. The dialects that have lower overall aspiration/deletion tend to have more lenition in word-final pre-consonant position than in word-internal pre-consonant position. For example, for Honduras, the word-final retention of /s/ pronounced as [s] before a consonant is 19% and the deletion rate is 23%, while in word-internal position /s/ is retained as [s] 63% and is only deleted 3%.

Whether the syllable following the /s/ is stressed or not has also been reported to be an important factor, at least when the following segment is a vowel. For all but one of the dialects (El Salvador) in Lipski's table, there is either the same or more deletion before an unstressed vowel than a stressed one. Valdivieso and Magaña (1991) also note more deletion before unstressed vowels (24%) as compared with stressed vowels (8%). Poplack (1980a, 1981) and Reynolds (1994) similarly report that a following unstressed syllable favors lenition over a following stressed syllable.

The role of following stress is perhaps a bit more complicated than this because Terrell (1977) states that the stress of a following vowel only has an effect for word-final /s/ that is part of a determiner or an object pronoun.

Brown and Torres Cacoullos (2003) also look at two other aspects of the phonetic environment: the preceding vowel, and the stress of the syllable that the /s/ occurs in. For preceding vowel, they find that /a/ is the most favorable environment (57% lenition), while /u,i,ei/ are the least favorable environments (46% lenition). For the stress, they find that a word-final /s/ in a stressed syllable is more likely to be weakened (53%) than a word-final /s/ in an unstressed syllable (38%).

2.2.2 Grammatical category

In addition to phonetic influences, the grammatical status of a word-final /s/ also has been reported to affect the probability of lenition. In Spanish, a word-final /s/ may be simply part of the lexical item and therefore not contribute to its meaning (*entonces,* 'then') or it can be inflectional, as a plural marker or a verbal ending. Plurality is indicated by adding an /s/ to nouns (*libro, libros,* 'book, books') adjectives (*bueno, buenos*, 'good'), determiners (*la, las*, 'the'), pronouns (*lo, los*, 'him, them'). It can also be part of an intrinsically plural, such as quantifiers and numbers. In verbs, a final /s/ disambiguates the second person singular from the third person singular of verbs (*tú estás*, 'you are', vs. *él está*, 'he is'), appears in the first person plural ending *–mos* (*estamos*, 'we are'), and in the third person singular of *ser* (to be): *es*.

In her study of Puerto Rican Spanish, Poplack (1980a, 1980b, 1981) finds that word-final /s/ in a monomorphemic lexical item is less likely to be aspirated or deleted than inflectional /s/, and that adjectives and nouns favor deletion while determiners inhibit deletion. Hundley (1987) also finds that word-final /s/ is deleted more frequently when it forms the plural morpheme (26%) than when it is in a monomorphemic lexical item (only 12%) in Peru. Other studies of grammatical status have also found similar results. For Chile, Cepeda (1995) finds the most deletion in nouns and "noun post modifiers" (i.e. adjectives) and the least deletion in lexical /s/ and "noun pre-modifiers"(i.e. determiners). In Chihuahua, Brown and Torres Cacoullos (2003) find the most lenition in the second person singular verbal form (58%), the least in plural markers (44%), with lexical /s/ in the middle (45%).

Terrell (1977) has the most in-depth look at different grammatical categories, as shown in Table 2. Terrell has split his data for word-final /s/ into many different categories based on the grammatical status or part of speech as well as some lexical exceptions. Included in the broad category "plural marker" are items that have *low deletion* (first modifier, numerical modifier, and the pronoun *nos*), moderate deletion (*ellos/ellas,* proper names, unmodified nouns, object pronouns), and *high deletion* (modified nouns, adjectives with redundant plural marker and the pronoun *nosotros*). The pronouns that are listed by their lexical identity are actually included in the broad category *plurals.* According to Terrell, lexical entries have "moderate" deletion—less than modified nouns and adjectives, but about the same as unmodified pronouns.

	Cuban	Puerto Rican	
más	1	3	
First Modifier	3	11	
Numerical Modifier	9	13	LOW DELETION
es	4	19	
nos	18	6	
ellos/ellas	8	23	
Lexical	21	19	
Proper Names	11	32	
Unmodified Nouns	19	35	MODERATE DELETION
Object Pronoun *los/las/les*	27	30	
pues	24	37	
Verb form *tu*	43	30	
Modified Nouns	39	44	
Adjective with redundant plural marker	37	46	
Verb form *nosotros*	66	31	HIGH DELETION
entonces	80	55	
Pronoun *nosotros*	85	56	

Table 3. Word-final /s/ deletion data for Cuban and Puerto Rican speakers, separated by grammatical category and lexical exceptions (taken from Terrell 1977).

2.2.3 Functional considerations

Unlike word-internal syllable-final /s/ or word-final /s/ that is part of the lexical representation of the whole word, in the case of the plural, the word-final /s/ may be the only indication of plurality. The plural marker is redundant across the entire noun phrase, in each determiner, noun, and adjective (*las casas blancas*, ‘the white houses’). If all of the plural markers are deleted, the result may be ambiguous. (*la casa blanca* could be 'the white house' or 'the white houses'). However if at least one of the plural markers is retained, then the meaning of plurality is maintained for the whole noun phrase.

Terrell's (1977) chart in Table 3 above for Cuban and Puerto Rican Spanish hints that the deletion of plural /s/ may be related to whether it is required to maintain the meaning. For example, "first modifiers" (i.e. the first item that contains the plural morpheme) has very low deletion (3% for Cuban and 11% in Puerto Rican) as compared to nouns or adjectives. He also splits up nouns into two separate categories, modified and unmodified. On the one hand, 'modified nouns' in which the plural marker would be redundant, have high deletion, and on the other hand, 'unmodified nouns,' in which the plural marker would be information-bearing, have only moderate deletion.

Poplack (1980a) also looks at how the grammatical category intersects with functional considerations in the expression of the plural. She finds that within an NP, the initial position has the strongest weight to inhibit deletion of the plural marker. For the following positions, there is a 'perseverance' effect; the deletion status of the preceding word(s) has a great impact on whether the plural marker –/s/ will delete; if the plural marker -/s/ is deleted in the first element in the noun phrase, the plural marker in the following elements are more likely to delete as well. This result is "anti-functional" in that the tendency is for either all elements to be redundantly marked or for none of them to be marked (see also Scherre and Naro 1992 for a similar effect in the Portuguese of Rio de Janeiro). Poplack considers whether a functional hypothesis can still be maintained by examining whether there is other information in the discourse that disambiguates the plural from the singular, such as inflectional (other /s/ in the NP), other morphological support (ex. verb form), non-morphological (syntactic, semantic), or a combination of these factors. She finds a weak functional effect of whether there is disambiguating information outside of the NP; when there is disambiguating information, the plural marker is more likely to be deleted. However, the lack of disambiguating information does not prohibit deletion; Poplack notes that out of over 1,000 tokens where plurality would only be expressed by the word-final /s/ in the NP, 39% of them were deleted. In these cases, there was no indication of plurality remaining.

Hundley (1987) uses a similar approach to Poplack to test for an effect of functional factors in the deletion of plural /s/ in Peruvian Spanish. Just as Poplack reports, Hundley also finds that there is a tendency to retain plural /s/ when there is no other disambiguating information and a tendency to delete plural /s/ when there is some sort of information available for disambiguation (morphological information: *los*, or verbal agreement, semantic information and/or syntactic constraints that help with disambiguation). In addition, he finds when there are multiple sources of disambiguation, the tendency to delete plural /s/ is stronger. Hundley also reports that just as in Puerto Rican Spanish, functional considerations are only tendencies in Peruvian Spanish; even when there is no disambiguating information outside of the NP, he reports that plural /s/ still was deleted 27% of the time.

For Chilean Spanish, Cepeda (1995) also notes an anti-functional effect of the plural marker: when the -/s/ is the only indication of plurality in a word, she finds more deletion than in words where the /s/ is part of the morpheme –/es/, and deletion of the /s/ would still leave a plural indicator remaining (*pared, paredes* ‘wall, walls’).

In addition to word-final /s/ lenition, Puerto Rican is undergoing another lenition process: velarization and deletion of word-final /n/. The two processes affect each other in a functional way because the verbal ending –/n/ disambiguates the third person plural from the third person singular. If both of these markers are missing, then plural meaning can be lost. Poplack (1981) finds a functional effect in that speakers who tend to delete plural /s/ tend to retain verbal /n/, and vice versa.

Other studies have been conducted of functional considerations for verbal /s/ for the second person singular. Normally pronoun subjects are not needed in Spanish because the verbal form indicates person and number. In dialects such as Puerto Rican where -/s/ is deleted from the second person plural, there is a functional effect to use the overt subject to reduce ambiguity (Hochberg 1986 and Cameron 1993). This is beyond the scope of the current study

2.2.4 Number of syllables in the word

A number of researchers have indicated that the number of syllables in the word affects the rate of lenition of word-final -/s/. Lipski (1984a), Barrutia & Terrell (1982), Cepeda (1995 for Chile), Lafford (1989 for Colombia), Terrell (1978a for Argentina) all find higher retention in monosyllabic words than in polysyllabic words. From the literature, it is not clear the important factor is monosyllabic words in comparison to all words with more than one syllable, or if there is an additional effect of the number of syllables for polysyllabic words as well.

2.2.5 Lexical exceptions

There are several references in the literature to particular words that appear to be lexical exceptions; they may have particularly advanced or conservative /s/ lenition even after taking into account the previous linguistic factors that have been identified. For example, Poplack (1981) notes low lenition rates for the determiners *los* and *las*. Lipski (1984:36) mentions that lenition in the word-final /s/ in the first person plural verbal ending *–mos* and in the word sequences *todos los* and *todas las* is more advanced than in other words. Terrell (1977) also points out several lexical items: deletion is rare in the words *es, más,* and *nos*, while deletion is common in *entonces, nosotros*, and in the verbal ending *–mos*.

2.3 Summary

Thus, as we have seen, numerous factors (both linguistic and extralinguistic) affecting syllable-final /s/ lenition have been studied by various researchers and reported on in the literature. One of the major extralinguistic factors is dialect and, indeed, syllable-final /s/ lenition is one of the variables used to classify the dialects of Latin American Spanish. There is significant variation in syllable-final /s/ lenition across dialects reported, with a general trend of more advanced lenition in the countries of the Caribbean and other coastal areas and more /s/ retention in interior or mountainous areas. In general, there has been more advanced lenition reported among males than females. Results on the effect of age have been conflicting. Studies of social class have found a fairly strong correlation, with the most advanced lenition among speakers in the lowest social class. Speech style has also been found to be an important factor, with formal speech (and particularly broadcast news speech) having far less lenition than informal or conversational speech.

As for linguistic factors, we have seen that the phonetic environment of syllable-final /s/ has a great impact on the rate of lenition. Results reported in the literature show that there is an effect of the grammatical status of word-final /s/ on lenition, with the morpheme -/s/ having higher deletion rates than -/s/ that is part of a lexical item.

Functional considerations also come into play for word-final /s/ as a plural marker. Other linguistic factors that have been reported in the literature to affect syllable-final /s/ lenition are the number of syllables in the word and lexical exceptions.

3 *Usage-based accounts of variation*

The term ***usage-based model*** of language was first coined by Langacker (1987). Its main premise is that the structure of language is primarily shaped by how the language is used. In generative models of language, a speaker's performance differs from their actual competence, and linguistic knowledge is primarily determined by innate linguistic capacities. In contrast, a usage-based model of language states that linguistic performance is a direct reflection of the speaker's competence. Linguistic knowledge is primarily based on experience with the language, and linguistic structure emerges as a result of language acquisition. While generative phonology claims that phonetic variation is extraneous, in a usage-based theory, linguistic knowledge ***is*** performance, and "statistical variation is intrinsic to the nature of language" (Pierrehumbert 1994:253).

According to Langacker (2000) (as well as Bybee 2001a, 2002, Pierrehumbert 2003, Fenck-Oczlon 2001, Hulst 2000, and others), language abilities are largely based on more general cognitive capabilities. Language-specific processes are posited only when structures cannot be derived from these more general cognitive functions. Langacker identifies the general psychological abilities that are critical to language as perception, memory, and categorization. Phonological units such as phonemes are created through the general cognitive ability of *abstraction*, which then causes an association or *schema* between segments in different words that have similar phonetic parameters. Through the process of *entrenchment*, any expression that is repeated can result in a unit that is stored in memory. There is no clear distinction between lexicon and grammar; they "grade into one another." (Langacker 2000:33)

According to Bybee (2001a), the most important usage-based characteristics that shape lexical representations are are *token frequency* and *type frequency*:

> "TOKEN FREQUENCY is the frequency of occurrence of a unit, usually a word, in running text – how often a particular word comes up. TYPE FREQUENCY refers to the dictionary frequency of a particular pattern (e.g., a stress pattern, an affix, or a consonant cluster)." (Bybee 2001a:10)

Token frequency shapes the structure of language and mental representations because higher-frequency words have a higher level of automation and therefore would have more advanced lenition (Bybee 2001a, 2002). High token frequency also creates high "familiarity," resulting in lower "cognitive cost" in production and perception (Fenk-Oczlon 2001:432). It has also been claimed that other usage-based factors such as word predictability and the contexts that the word appears in have an impact on reduction processes.

3.1 Frequency

Many studies have shown greater levels of reduction in high frequency words, and changes involving lexical diffusion often progress first in higher frequency words. For example, Hooper (1976) finds greater levels of schwa-deletion in English for high frequency words such as *memory* (with a word frequency of 91 occurrences per million words) than in low frequency words such as *mammary* (with a word frequency of 0 occurrences per million words). Phillips (1980, 1984) reports that the historical change of /a/ raising to /o/ before nasals in Old English is affected by word frequency, with the progressive spelling *o* occurring nearly categorically (98%) in frequent words (those with more than 400 tokens in the corpus she evaluates), but only 39% in infrequent words (those with 1 to 10 tokens). Similarly, Labov (1994) reports that short *a* in Philadelphia is undergoing a process of tensing, with the change progressing first in higher frequency words (although some of the highest frequency words have not been affected). Fosler-Lussier and Morgan (1999) report that higher frequency words in English tend to have less canonical pronunciations than those of lower frequency. Fenk-Oczlon (1990, 2001) looks at the length of verb forms with two different aspects (perfective/imperfective) in Russian, and finds that in 50 out of 67 pairs, the more frequent aspect for that particular verb is shorter (so word frequency is more important than semantic markedness in this case).

Bybee (2000, 2002) claims that frequency effects in word-final /t,d/ deletion in American English give evidence that at least some multi-morphemic units are stored in the lexicon. When divided into the categories *high frequency* (35 or more occurrences per million words) and *low frequency* (less than 35 occurrences per million words), she finds significantly more deletion among the high frequency words (54.4%) as compared to the low frequency words (34.3%). Further, this frequency effect is found even when considering only regular past tense (*-ed*) verbs (39.6% deletion for high frequency verbs compared to 18.9% deletion for low frequency verbs). Although others have claimed that a morphologically complex word would be stored as individual morphemes, Bybee states otherwise: “Since multi-morphemic words evince frequency effects, they must be stored in the lexicon” (Bybee 2002:272-273).

Frequency effects have also been reported for word sequences. Very high frequency phrases such as *I don’t know* and *I’m going to* undergo lenition (*I dunno, I’m gonna*) that does not occur within the same words when they occur in other contexts. Similarly, contractions occur more often among more frequent word sequences, as described by Krug (1998). He looks at how often contractions (such as *’ve*, *’s*, *’m*, *’re*, *’ll*) actually occur in constructions where they are allowed in English, and finds that there are frequency effects. Krug claims that the frequency of the sequences determines native speakers’ intuitions about how closely connected two adjacent words are. Sequences that are very frequent are more closely connected (and therefore more likely to be contracted). For example, the potentially contractible sequence *everything* + *is* or *has* is frequent (it occurs 220 times in his corpus), and the contracted form *everything’s* is used in 62% of the constructions, while the similar sequence *nothing* + *is* or *has* is less frequent (it occurs only 83 times), and is only contracted to *nothing’s* in 48% of the constructions.

Bush (2001) investigates palatalization of /t/ and /d/ across word boundaries when followed by high front segments (/i/ and /j/) in the CHILDES corpus (MacWhinney 1995). Palatalization of /t/ and /d/ when followed by /i/ and /j/ is a phonological rule that applies categorically word-internally and optionally across a word boundary. Bush finds that palatalization across a word boundary is more likely when the two-word sequence (*bigram*) has high relative frequency (such as *did you*) than when it has a low relative frequency (such as *backyard you*); palatalization occurs in 69% of high frequency bigrams as compared with 31% of low frequency bigrams). Bush claims that this is evidence that there is a single lexical entry consisting of the two-word chunk for the frequent bigrams. Since palatalization of /t/ and /d/ normally occurs word-internally, if the two-word unit is stored as one unit, it would explain why it applies across the word boundary in these instances.

In her study of French consonant liaison, Bybee (2001b) also concludes that multi-word units must be stored in the lexicon. Historically, in the 16th and 17th centuries, word-final consonants that had previously been pronounced underwent deletion, a change that progressed first in pre-consonant position. Today, the word-final consonants are pronounced only when followed by a vowel, and only in high frequency forms: suffixes, function words, and small closed classes such as prenominal adjectives, all of which are high frequency. There are no lexical or open class items that exhibit liaison. Because the constructions are high frequency, they can be stored separately. Because they are stored separately in memory, they have not been subjected to the same consonant lenition as low frequency constructions.

Phillips (2001) looks at several historical changes and changes in progress and claims that both word frequency and word class play a key role in sound change. She presents a *Frequency-Implementation Hypothesis*, which states that a sound change will affect more frequent words first if it does not require lexical analysis (for example, vowel reduction, deletion, and assimilation), and it will affect less frequent words first if it does require lexical analysis. Phillips (2001) claims that word frequency and *word class* can have separate effects. Philips particularly points out that function words (words that receive low sentence stress) behave differently because they are less analyzed, but she also shows that nouns and verbs can behave differently from each other as well.

3.2 Word probability

Gregory *et al*. (1999) and Jurafsky *et al.* (2001) claim that any measure of word probability affects the extent of phonetic reduction. They put forth a *Probabilistic Reduction Hypothesis*, which makes the claim that when a word is more probable, it is subject to greater reduction, and that the effects of probability should be unidirectional, with an increase in any measure of probability causing a greater extent of reduction. The overall probability of a word in context is affected by a number of different factors in addition to the frequency of the word or sequence of words. These include the conditional probability of the word (the probability of the word given one or more neighboring words), as well as the repetition number of the word in the discourse (with a word being more probable in its first or subsequent repetition than in its first use), and the semantic relatedness of the word to the discourse (with words more closely related semantically to other words in the discourse being more probable).

Quantifying the effect of these additional measures of word probability can be problematic. In his study of /t/ and /d/ palatalization across word boundaries, Bush (2001) attempts to measure the effect of what he terms the ***transitional probability***, which is the conditional probability of the second word, given the first word of the bigram (i.e.for *did you,* the probability of *you* given the word *did*). He finds that conditional probability is significant in predicting when palatalization occurs; palatalization applies in 60% of bigrams with high conditional probability and in 32% of bigrams with low conditional probability. In theory, relative frequency and conditional probability should be independent (the relative frequency and conditional probability could both be high or both be low, or one could be high and the other low). However, due to the small size of the corpus he evaluates, the relative frequencies and the conditional probabilities are correlated. There are only a total of 200 bigram sequences that have the appropriate conditioning environment, and the eight bigrams with high relative frequency (those occurring at least 5 times: *did you*, *didn't you*, *don't you*, *last year*, *that you*, *told you*, *what you*, and *would you*) also have high conditional probabilities (using a cut-off of

.0444 between high and low conditional probabilities). Because the relative frequency and conditional probabilities are related, it is not entirely possible to separate the effect of the relative frequencies and conditional probabilities.

Gregory *et al.* (1999) look at the weakening of word-final /t,d/ in English from a subset of the Switchboard corpus of recorded telephone conversations (Godfrey, Holliman, & McDaniel 1992). They evaluate two discrete variables – /t,d/ flapping or tapping and /t,d/ deletion – and one gradient variable, the duration of the /t/ or /d/. In each case, they start with a base model that includes many of the independent variables that have been previously identified in the literature as being important, including the speech rate, the following segment type (consonant/vowel), the preceding segment type (i.e. whether the /t,d/ is part of a consonant cluster), whether the next vowel is reduced, the grammatical status of the /t,d/ (past tense/monomorphemic), the identity of the underlying segment (/t/ or /d/), the number of syllables in the word, and word class (content word/function word). They evaluate several measures of the word: the relative frequency of the word; the relative frequency of bigrams (the word taken with either the preceding or following word); the conditional probability of the word based on one or more of the surrounding words; the mutual information (a measure of the cohesion between the words and the following word); and the semantic relatedness of the word to other words in the discourse. In the case of categorical deletion, word frequency, mutual information, and the conditional probabilities of the word given the preceding two and the following two words are all significant. In the case of /t,d/ duration, they find that word frequency, mutual information, semantic relatedness, conditional probability of the word given the preceding word, and the repetition number are all significant. Tapping of the /t,d/ is significantly affected only by the mutual information. For each of the measures of the extent of /t,d/ weakening, each of the significant probabilistic factors have the effect predicted by the Probabilistic Reduction Hypothesis: the more probable the word, the greater the weakening.

Jurafsky *et al.* (2001) focus on the effect of relative frequencies and conditional probabilities, and attempt to determine if the same constraints apply to content words and function words. For the content words, they look at /t,d/ deletion and duration in the content words from the Gregory *et al.* (1999) study. For the function words, they look at vowel reduction in the ten most frequent function words in English. They analyze both the discrete variable of phonemic pronunciation (full or reduced vowel) as well as the continuous variable of segment duration. The methodology is similar to that in Gregory *et al*.; for both phenomena, they start with a base model including previously identified factors and then evaluate which of the relative frequency and conditional probability variables are additionally significant. For vowel reduction, the base model includes the rate of speech, the following segment type (consonant/vowel), syllable type (open/closed), and whether the following vowel is reduced.

Jurafsky *et al.* find that vowel reduction is greatly affected by the predictability of the word in the context. For the discrete variable of whether the vowel is reduced or not, the conditional probability of the word given the preceding word is significant, with 48% reduced vowels for high conditional probability compared to 24% for low conditional probability. They find similar effects of conditional probability on duration, with shorter duration being predicted when the word is more probable given either the preceding word or the following word. The effect of conditional probability in the case of /t,d/ is "much weaker" (p. 245). For /t,d/ deletion, while word frequency is significant (with high frequency words twice as probable to delete /t,d/ compared to the lowest frequency words), conditional probability is not significant. For duration, both word frequency and the conditional probability given the following word are significant, although the magnitude is not large; they report that words with high conditional probability are only 12% shorter than words with low conditional probability.

From this data for /t, d/ deletion in content words and vowel reduction in function words, Jurafsky *et a*l. conclude that there is evidence that predictability does have an effect on reduction, and that it operates both on the categorical processes of deletion and vowel reduction, but also on the continuous processes of shortening. Furthermore, they claim that their results provide evidence that "probabilistic relations between words are represented in the mind of the speaker" (p. 230).

3.3 Exemplar theory and lexical representations

In order to account for the word frequency effects described above, and particularly the phonetically gradient effects, Bybee claims that the lexicon must contain fine-grained phonetic detail and information regarding the extent of allowable variation. Rather than having a lexicon with entries consisting of phonemes, each lexical entry is based on *exemplars*, tokens of that word that the speaker has previously encountered (see Nosofsky 1988 for general cognitive processes; Johnson 1997 and Pierrehumbert 2001 for phonology).

Pierrehumbert (2001) demonstrates how an exemplar-based model of phonology works. Each entry in the lexicon is stored in memory as a cloud of previously encountered exemplars stored in a sort of "map" of phonetic parameters. The more phonetically similar two exemplars are, the closer they will be to each other in this map. Pierrehumbert assumes that there is a minimum granularity to the phonetic parameters – related to the minimal difference that is perceivable, or Just Noticeable Difference – and below this granularity, all tokens are grouped together as one individual exemplar with a higher strength, proportional to the number of tokens it represents. When a new token is heard, its phonetic parameters are computed, and a similarity metric is calculated to each of the previously stored exemplars in the vicinity. In order to categorize the new token as one of the stored lexical items, for each of the stored lexical entries, the similarity metrics to all exemplars in that category are added together. The new token is given the label having the highest score. This token is then used to update the representation of that lexical entry, either by storing it as a new exemplar or by increasing the strength of a previously stored exemplar.

A model based on exemplars can account for phonological variation, and particularly for words that are lexical exceptions. Because words can be stored as exemplars rather than exclusively as sequences of phonemes, different words can be associated with slightly different phonetic parameters. Higher level categories such as phonemes and morphemes are abstracted when the type frequency (the number of different words containing the same pattern) is high, and a schema is created among words with the same category. Phonetic changes that occur in one word will cause changes in other words by analogy, with words that are more similar having greater analogical influence on each other than words that are less similar. Frequency effects in morphologically complex words such as in the case of /t,d/ deletion in regular *–ed* words reported by Bybee (2002) and in multi-word sequences are derived in the same manner as for monomorphemic words if they are stored as a whole in memory. However, low frequency forms that can easily be derived by analogy from a schema (such as the past tense of low frequency regular verbs) are not stored separately. Extremely high frequency tokens may operate independently from the schema because they are less analyzed into their constituents (Bybee 1999, Phillips 2001).

With an exemplar model, there is no need for a "rule" to account for lenition or other phonological changes; instead, there is a shift in the phonetic parameters contained within the stored exemplars. For example, assume a new token of the word *just* is pronounced with the final /t/ slightly reduced (presumably due to phonetic undershoot of the actual target). After this token is encountered and classified, the lexical representation of the word *just* is updated by adding a new exemplar with the /t/ slightly weakened. The next time the speaker wants to say the word *just*, when a phonetic target is chosen randomly from the stored exemplars, the newly added exemplar makes it more likely that the phonetic target will be slightly weakened. In this way the category will gradually change: phonetic undershoot of this slightly weaker phonetic target occurs, which causes the representation to be updated and shifted further, etc. In this model, lenition is more advanced for more frequent words because they have more opportunities to be affected. The extent to which the /t/ in *just* is analogically related to the /t/ in other words determines how much the change in one word affects the whole category.

Because each token encountered contributes to the lexical representation in an exemplar model, other factors are predicted to have an impact on the extent of lenition. For example, Bybee (1999:221) suggests that the "social contexts" that words are used in can have an effect on the representation. Each time a word is used, the lexical representation is updated with a new exemplar. When a word is primarily used in an informal social context that favors lenition, the lexical representation then contains more exemplars that contain the lenition. When the speaker uses the word in a more formal social context, because the phonetic target is based on the stored exemplars, the word will have more advanced lenition than other words that don't occur as often in informal social contexts. Likewise, the phonetic context a word most often occurs in can cause that particular lexical item to change at a different rate than other words:

> "Even holding frequency constant, a word that occurs more often in the right context for a change will undergo the change more rapidly than a word that occurs less often in the conditioning context." (Bybee 2002:274)

In her discussion of /t,d/ deletion, Bybee claims this explains why past tense regular verbs have lower overall deletion rates than monomorphemic words in all phonetic environments; for example, in pre-consonant position, regular *–ed* verbs delete 47% of the time, while all words delete 59%. Past tense regular verbs occur more often in an environment that disfavors deletion (i.e. before a vowel) than other words ending in /t/ or /d/ (40% for regular verbs compared to 21% overall).

An exemplar model can also explain the effect of ***neighborhood density***, or how many other words are similar to a given word. Because words in a dense neighborhood are pronounced more carefully than ones that are not (Wright 1998), they may resist a change that affects other words. In the exemplar model, these words are pronounced more carefully, and therefore there is less variation in the exemplars stored in the lexical representation.

3.4 Usage-based factors in relation to syllable-final /s/ lenition

As we have seen in this chapter, a number of usage-based factors have been described in the literature as affecting the extent of lenition in a number of different processes. In the following chapters, I will be evaluating the effect of some of these factors on syllable-final /s/ lenition: word frequency, measures of word predictability (relative frequency of word strings, conditional probabilities, repetition), and other factors related to the lexical representation: impact of whether there is a nearest neighbor and the effect of phonetic context. Note that some of the measures of word predictability identified by Gregory *et al*. (1999) are out of the scope of this investigation – namely, syntactic and semantic considerations.

The specific questions that this study addresses are the following:

- Which of the usage-based factors contribute to syllable-final /s/ lenition? In their study of word-final /s/ lenition in the Spanish of Chihuahua, Mexico, Brown and Torres Cacoullos (2003) find a significant effect of word frequency, with high frequency words favoring lenition. Other factors, such as conditional probabilities, have not been previously explored.
- Do these usage-based factors operate in the same way on categorical deletion data and on continuous data that capture the gradient nature of /s/ lenition?

Jurafsky *et al*. found that for /t,d/ weakening, there are some differences in the effect of word probability for categorical data and continuous data.

- Do all of the factors of word probability operate in the same direction, with higher word probability causing more weakening?
- Are there differences in the effect of usage-based variables, particularly word frequency, for function words vs. content words, as the data from Phillips (2001) and Jurafsky *et al*. (2001) suggest might be the case?
- How does the phonetic context of use affect lenition; do words that are more often followed by a consonant show greater levels of lenition, even when they are followed by a vowel?
- Do the same constraints apply to word-final /s/ in the plural morpheme and in the verbal morpheme *–mos* as to other words?

4 Data Collection

For the present study, two Spanish speech corpora, the LDC CallHome Spanish Speech Corpus and the LDC Broadcast News Spanish Speech Corpus, have been labeled for syllable-final /s/. The CallHome corpus consists of recordings of informal telephone speech, and the coding was done exclusively by listening. The Broadcast News corpus consists of speech recorded from radio news broadcasts, and the coding was generated automatically using a speech recognition tool. In addition, for the Broadcast News corpus, a number of acoustic measurements were also made: duration, spectral information and intensity measurements.

4.1 Spanish pronunciation lexicon

To facilitate the data collection, a Spanish pronunciation lexicon was used for pronunciation and other information. The main pronunciation lexicon used to determine the pronunciation for each word was the LDC Spanish Lexicon (Garrett *et al*., 1997), which was developed to cover the vocabulary in the LDC CallHome Spanish corpus, and to be used to support speech recognition applications. For each word entry, the following information is included:

- Morphological information: part of speech, tense, number, person, gender, etc.
- Pronunciation: the canonical pronunciation of the word spoken in isolation.
- Stress information: each word has 1 syllable marked as stressed
- A count of the number of times the word appears in 5 different corpora:
    - LDC Spanish CallHome corpus (143,394 words)

- Madrid Radio transcripts (941,199 words)
- Spanish Language Associated Press Newswire text (8,429.549 words)
- Reuters Latin American Business Report and Reuters Spanish Language Business Report (18,742,153 words)
- El Norte newswire (29,745,911 words)

When this lexicon was created, the pronunciations were automatically generated from the orthographic representation using pronunciation rules, and then hand-checked and corrected when necessary. Only a small number of words needed to be hand-corrected. Except for a handful of words that have 2 significantly different pronunciations, the lexicon includes only one canonical representation for each word; it does not attempt to account for any pronunciation variation due to dialect differences or fast speech rules.

In addition to the LDC Spanish Lexicon, a supplemental lexicon (Kingsbury, 1997) was also used for words that were not contained in the original LDC Spanish Lexicon. This supplemental lexicon contains only the pronunciation and stress information, generated in the same manner and using the same phoneme set as the LDC Spanish Lexicon. Between the two lexicons, nearly all of the words in the CallHome Spanish and Spanish Broadcast News corpus are covered.

With respect to /s/, since the lexicon does not attempt to account for any pronunciation variation, no syllable-final /s/ aspiration or deletion is incorporated, and all instances of an underlying /s/ are represented in the canonical pronunciation. However, the pronunciation rules generated one of two surface pronunciations, [s] or [z], depending on the context: when a word-internal /s/ was followed by a voiced consonant, then the /s/

was listed in the pronunciation as [z]. Because the surface [z] represents an underlying /s/, I have treated all syllable-final instances of [z] in the lexicon in the same way as syllable final /s/.

For the present study, I identified all tokens of underlying syllable-final or word-final /s/ using a perl script to check the pronunciation of each word in the transcript files for each of the corpora. Although syllabification is not specified in the pronunciation lexicon, syllable boundaries are easy to identify in Spanish because all instances of word-internal /s/ followed by a consonant are considered to be syllable-final (Barrutia and Terrell, 1982). This means that the perl script located each [s] or [z] and determined if it was syllable-final by checking to see if it was word-final or followed by a consonant. The perl script then generated a list of each occurrence of syllable-final /s/ to be coded for each corpus. Words with more than one syllable-final /s/ were included more than once on the list, with each /s/ identified separately.

4.2 CallHome Spanish

4.2.1 The LDC Hub5-LVCSR CALLHOME Spanish Speech Corpus

The first data set used for this syllable-final /s/ investigation is the LDC CallHome Spanish corpus, which was developed to support a Large Vocabulary Conversational Speech Recognition project and consists of telephone conversations between native speakers of Spanish.

The total corpus contains over 100 phone calls, each lasting up to 30 minutes, of which only a 5- or 10-minute section of each telephone call was transcribed. For the present study, 95 of the phone calls were used, all of which had a 10-minute section transcribed. The transcription generally starts about 2 minutes into the call. Word-level time-alignments are not included with the transcriptions, but the start and end time for each speaker turn is indicated to allow correlation of the text and the speech files. The transcriptions are purely orthographic; no attempt was made to include any pronunciation information or correct any grammatical mistakes. This made it straightforward to identify all occurrences of syllable-final /s/ using the orthographic transcription and the pronunciation lexicon.

Each caller was allowed to call whomever he or she wanted, and most called family members or close friends. The phone calls were unscripted and the speakers were allowed to talk to each other about any subject, so the speech style in this corpus is generally informal.

The speech in this corpus is well suited to the task of studying variation in the lenition of syllable-final /s/ across dialects and speakers due to the informal speech style and the large number of speakers from different Spanish dialects. Three attributes of each speaker are identified: gender, dialect (country), and approximate age (child/adult/elderly). The breakdown of numbers of speakers and tokens of syllable-final /s/ by dialect for each of the Latin American Spanish dialects included in the corpus are shown in Table 4. The fact that only a short portion of speech was transcribed for each speaker allows for larger speaker and dialect coverage. However, it also means that each speaker has a limited number of data points, as shown in the breakdown of speakers by the number of tokens of syllable-final /s/ in Table 5.

Dialect	Speakers	Syllable-final /s/		
		Total	Word-final	Word-internal
Argentinean	20	2,306	1,523	783
Bolivian	4	426	316	110
Chilean	43	4,275	3,024	1,251
Colombian	17	1,410	1,005	405
Costa Rican	2	151	87	64
Dominican	6	643	452	191
Guatemalan	5	398	300	98
Honduran	2	176	118	58
Mexican	31	3,377	2,401	976
Paraguayan	5	404	287	117
Peruvian	54	5,104	3,426	1,678
Puerto Rican	3	200	138	62
Salvadoran	5	596	439	157
Uruguayan	2	302	158	144
Venezuelan	12	1,497	1,052	445
Total	211	21,265	14,726	6,539

Table 4. Number of speakers and number of syllable-final /s/ tokens in the CallHome Spanish corpus.

#Tokens	# Speakers
< 10	7
10-49	40
50-99	54
100-199	103
200-299	7

Table 5. Number of speakers having the number of syllable-final /s/ tokens in each range in the CallHome Spanish corpus.

Since the CallHome Spanish corpus contains telephone speech, all of the recordings were made with a sampling rate of 8000 Hz, and the quality of the recordings is variable. The corpus includes some basic information regarding the overall quality of the files. Because the people who were doing the coding of syllable- final /s/ for this corpus were able to indicate when the recording quality was poor, the information in the corpus regarding recording quality was not used for the present study.

4.2.2 CallHome Spanish Data Coding

Two students at the University of Pennsylvania performed the coding of syllable-final /s/ for the CallHome Spanish corpus. The first was a female native speaker of English who is proficient in Spanish and a linguistics student, and the second was a male bilingual speaker of English and Puerto Rican Spanish who had not previously studied linguistics. Both were familiar with syllable-final /s/ lenition before beginning the project.

After I generated a list of all the occurrences of syllable-final /s/ in the corpus from the transcripts and the pronunciation lexicon, I added a large amount of redundancy to this list in order to measure both the repeatability of coding for each coder as well as the repeatability across coders. Since the task is a difficult one, it is important to validate the data by measuring the extent to which each coder consistently coded tokens and the extent to which the two coders used the same criteria.

The list of all tokens of syllable-final /s/ was then randomized so that tokens by the same speaker would not be repeatedly presented to the coders. This was an attempt to prevent any biases that could be created by knowledge of the speaker. For example, if the coder had an expectation that the speaker preferred to retain or delete syllable-final /s/, he or she might be more likely to hear what is expected. Randomizing the tokens also had the goal of keeping the coding criteria as consistent as possible over time; if a particular speaker pronounced /s/ very strongly in most cases, the coder's criteria might change so that a weaker /s/ would be mis-coded as a deletion.

In a further attempt to implement constant criteria across coders and across time, I developed a set of "golden" examples of syllable-final /s/ with two examples of each of the pronunciation of syllable-final /s/ as [s], [h], [z] and deletion. These examples were presented to the coders at regular intervals during the coding process: at the beginning of each coding session, and after about each 150 tokens during a coding session. During the coding process, I also did some spot-checking during the coding process to verify the accuracy of the coding.

I developed a Graphical User Interface using perl scripts and Bongo tools to facilitate the coding process. For each token of syllable-final /s/, the orthographic transcription of the entire sentence was presented on the computer screen, along with an indication of which syllable-final /s/ was to be coded. A word-level automatic alignment of the speech files was used to determine the approximate time location of the given word; from this alignment a window of speech starting 200 ms before the hypothesized beginning of the word and ending 200 ms after the end of the word was played. On average, this window corresponds to about the interval from the beginning of the preceding word to the end of the following word. The coder was able to replay the speech and to enlarge the window of speech to be played as needed.

The fact that the audio recordings were recorded with telephone quality made the coding significantly challenging. With a sampling rate of 8000 Hz, spectral content at frequencies above about 3000 Hz is greatly attenuated, and all spectral content above

4000 Hz is lost. The phoneme /s/ is characterized by high frequency sound energy, mostly above 3500 Hz. Spectrograms were therefore determined to not be helpful in identifying whether a syllable-final /s/ was retained or not, and the spectrograms were not used during the coding procedure.

The main coding categories available were: (1) [s]: the /s/ was retained; (2) [z]: the /s/ was retained and voiced; (3) [h]: the /s/ was retained, but only as aspiration; (4) deletion: the /s/ was deleted entirely. A number of alternate categories were also available in the event that a given syllable-final /s/ could not be coded. These categories were: (1) the recording was too distorted to determine the pronunciation of the /s/; (2) the following segment was also /s/, so the word-final /s/ in question could not be evaluated; (3) the entire syllable was truncated, or (4) the original transcript was incorrect. A breakdown of the number of tokens in each of these categories is shown in Table 6. All of the tokens with one of these codes was discarded and not used in the subsequent analysis. They are also not included in the token counts given in Table 4.

Category	# Tokens
Recording distorted	594
Following /s/	637
Truncated syllable	344
Incorrect transcription	469

Table 6. Breakdown of tokens that could not be coded for the CallHome corpus.

During the coding process, all tokens with word-final /s/ or /s/ followed by any consonant were coded. However, prior to compiling the data, two types of word-internal

/s/ tokens were removed from the data because the /s/ was determined to not be syllable-final. First, word-internal /s/ followed by the glides [w] and [y] were excluded. Although Spanish phonetics uses the term "semi-consonant" for glides when they occur before a vowel (Amastae 1989), an /s/ preceding a glide was determined to behave more like the onset of the following syllable, and during the data collection, it was never coded as being deleted. Secondly, tokens of /s/ that were preceded as well as followed by a (non-glide) consonant were also removed from the data because they were also never coded as being deleted. In Spanish, there are two consonants that precede an /s/ that is followed by a another consonant: [k] (for example, *experiencia* [e***ksp***eryensya]) and [n] (*inscrita* [i***nsk***rita]).

4.2.3 Categorical Data

Table 7 shows the percentage of tokens receiving each of the classifications [s], [h], and deleted for the CallHome corpus. Across all the Latin American dialects, 49.8% of the tokens were classified as [s], 44.4% were classified as deleted, 4.9% were classified as [z], and only 0.9% were classified as [h]. It is somewhat surprising that the category [h] occurred so infrequently given other studies have found much higher levels of aspiration. This may be due to the fact that the telephone-quality recording makes differentiation of [s] and [h] more difficult. Table 8 combines the data from the previous table into the two broad classifications "coastal" (high -/s/ deletion dialects) and "interior" (low -/s/ deletion dialects). The "coastal" dialects consist of Argentinean, Chilean, Dominican, Paraguayan, Puerto Rican, Salvadoran, and Venezuelan. All of the other dialects are included in the "interior" category.

Dialect	N	[s]	[h]	[z]	deleted
Argentinean	2,306	51.3	1.0	3.6	44.1
Bolivian	426	70.7	0.2	8.2	20.9
Chilean	4,275	23.2	1.9	3.2	71.7
Colombian	1,410	62.9	0.6	8.2	28.2
Costa Rican	151	72.2	0.7	2.6	24.5
Dominican	643	28.1	0.9	3.9	67.0
Guatemalan	398	63.8	0.0	8.3	27.9
Honduran	176	67.6	0.0	9.7	22.7
Mexican	3,377	68.3	0.3	9.0	22.4
Paraguayan	404	20.8	1.2	2.2	75.7
Peruvian	5,104	63.4	0.6	4.5	31.5
Puerto Rican	200	22.5	3.0	0.5	74.0
Salvadoran	596	50.2	0.7	2.0	47.2
Uruguayan	302	69.2	1.3	2.0	27.5
Venezuelan	1,497	24.9	1.2	2.1	71.7
Total	21,265	49.7	0.9	4.9	44.4

Table 7. Percent of tokens with each code for the CallHome Corpus.

Dialect	N	[s]	[h]	[z]	deleted
Coastal	10,124	31.8	1.4	3.0	63.7
Interior	11,344	65.4	0.5	6.6	27.5
Total	21,265	49.7	0.9	4.9	44.4

Table 8. Percent of tokens with each code for the CallHome Corpus, using the dialect categories “coastal” and “interior”.

4.2.4 Data Accuracy

As described above, with such a difficult classification task, it is important to quantify the reliability of the coding data. Therefore, many of the tokens were coded more than once. Coder 1 coded 3,144 tokens twice, coder 2 coded 760 tokens twice, and 2,036 tokens were coded by both coder 1 and coder 2.

Table 9 shows the distribution of [s], [h], [z], and deletion codes for the tokens that were coded more than once. When considering all four categories, Coder 1 had 86% agreement and Coder 2 had 86% agreement. Because there are so few tokens coded as [h] and [z], I am also reporting the data with [s], [h], and [z] combined into the category *retained* in Table 10. When [s], [h] and [z] are combined, the overall agreement increases slightly to 89% and 87%.

Across coders, the agreement was significantly lower – 73% for all codes considered separately and 77% for retention vs. deletion. This is mostly due to a difference in coding criteria by the two coders; Coder 1 labeled only 40% of all tokens as deletion while Coder 2 labeled 53% of tokens as deletion.

		Coder 1				Coder 2			
		[s]	[h]	[z]	deleted	[s]	[h]	[z]	deleted
Coder 1	[s]	1576	7	71	262	830	6	5	274
	[h]		22	0	45	7	1	1	16
	[z]			150	45	58	1	4	85
	deleted				966	84	8	4	652
Coder 2	[s]					327	0	7	87
	[h]						4	0	6
	[z]							1	5
	deleted								323

Table 9. Cross tabulation of codes within and across coders for tokens coded more than once.

		Coder 1		Coder 2	
		retained	deleted	retained	deleted
Coder 1	retained	1826	352	913	375
	deleted		966	96	652
Coder 2	retained			339	98
	deleted				323

Table 10. Cross tabulation of retained vs. deleted codes within and across coders for tokens coded more than once.

The overall agreement level of 73% agreement may be misleading, because a certain level of agreement occurs purely by chance when items are classified into a finite set of distinct categories. Therefore, a measurement of the reliability of the judgments is needed that corrects the proportion of agreement for chance. Such a measurement is Kappa (κ) (Cohen 1960, 1968):

$$k = \frac{p_o - p_c}{1 - p_c}$$

where p_o is the observed proportion of agreement and p_c is the proportion of agreement expected by chance. The value of κ varies between -1 and 1. Negative values of κ indicate there is less agreement than would be expected by chance; a κ value of 0 indicates that the agreement is exactly accounted for by chance, and a positive value indicates that there is more agreement than would be expected by chance. A value of 1 indicates perfect agreement.

κ can be calculated whenever there are multiple trials of a series of judgments, either by an individual coder, or across coders. In Minitab, κ is calculated using the Attribute Agreement Analysis function. Table 11 shows the κ values for each of the coders and across the coders for each of the coding categories. The values of κ across coders for [s] and deleted are 0.57 and 0.53, which means that after removing the agreement that can be attributed to chance, roughly half of the tokens that have these codes are in agreement (moderate agreement). The values of κ across coders for [h] and [z] are 0.04 and 0.01, indicating that their agreement is essentially due to chance. When the retained categories are combined, overall κ increases very slightly from 0.75 for the four coding categories to 0.76 for coder 1, from 0.73 to 0.74 for coder 2, and from 0.50 to 0.53 across coders.

	Kappa		
	Coder 1	Coder 2	Across Coders
[s]	0.78	0.75	0.57
[h]	0.45	0.57	0.04
[z]	0.70	0.13	0.01
deleted	0.76	0.74	0.53
Overall	0.75	0.73	0.50

Table 11. Reliability measurement κ for the CallHome corpus.

Cole *et al.* (1994) and Lander *et al.* (1995) examine the agreement between transcribers for broad phonetic transcriptions of several different languages using both the waveforms and spectrograms. In these experiments, either native or non-native speakers labeled the segments and marked the boundaries between segments. The accuracy was measured across all the segments in a reference string, based on the number of substitutions, deletions, and insertions between them, divided by the total number of segments. Label agreement among native speakers for the full label set (between 42 and 62 symbols) varied between 61% for German and 74% for Spanish, with an average of 68% (Cole et al. 1994). Non-native agreement was lower - ranging from 41% for German to 72% for Japanese with an average of 59% (Lander et al. 1995). Cole *et al.* additionally looked at transcriber agreement on the location of segment boundaries for the segments on which the transcribers agreed, and found that an average of 78% of boundaries were within 11 milliseconds of each other.

It is somewhat difficult to directly compare the overall agreement found in the studies performed by Cole *et al.* (1994) and Lander *et al.* (1995) with the present data because their labeling task included the full set of phonemes in the language while the present study focuses only on the realization of a single phoneme. However, the overall 73% agreement across coders using all four categories compares favorably to the 59% and 68% agreement in their studies.

Wester *et al.* (2001) compare the transcription of several variable phonological rules in Dutch generated by a speech recognizer and expert listeners. The processes they consider involve only the categorical insertion or deletion of segments (/n/-deletion, /r/-deletion, /t/-deletion, schwa-deletion, and schwa-insertion), with no attempt to discriminate between multiple variants of a phoneme. The method they use is very similar to mine; nine expert listeners were shown just the orthographic representation on the screen (no spectrograms were available to them), and each listener could listen as many times as needed. The speech recognizer ASR was used in forced recognition mode.

Wester *et al.* measure κ across listener pairs: range between 0.49 and 0.73, with a median of 0.63, which they characterize as 'moderate agreement'. In the present study, the overall κ measurement across the two coders using all 4 codes is 0.5, which falls within the range that Wester et al. reports.

For the data analysis portion of this study, each token of syllable-final /s/ had to be given a single code, even if had been coded more than once and had been given conflicting codes. In such cases, the following procedure was used. In the case of two conflicting codes, one of which was retained ([s], [h], or [z]), and the other of which was deletion, the single code given that token was the [s], [h], or [z], because if one of the coders heard the segment, it is more likely that it was indeed present. In the case of two conflicting codes among [s], [h], and [z], the single code kept for that token was [s] because the more frequent occurrence of [s] increases the likelihood of [s] over the other two codes.

4.3 Broadcast News

4.3.1 The LDC 1997 HUB-4 Spanish Speech Corpus

The Spanish Broadcast News corpus consists of speech recorded from news broadcasts from three sources: the Voice of America, (17.6 hours of transcribed speech), ECO (7.8 hours of transcribed speech), and Univision (7.3 hours of transcribed speech). The corpus is comprised of 81 separate speech files that were recorded in 1996 and 1997. The Voice of America has a large number of Cuban and other coastal speakers, while Univision and Eco, which originate in Mexico, have mostly interior speakers. The breakdown of speakers by the two broad dialect categories "coastal" and "interior" is shown in Table 12.

Dialect	News Source	Speakers	Syllable-final /s/		
			Total	Word-final	Word-internal
Coastal	Voice of America	67	12,358	9,048	3,310
	Univision	73	781	534	247
	Eco	24	450	329	121
	Total	164	13,589	9,911	3,678
Interior	Voice of America	27	3,981	2,956	1,025
	Univision	338	16,450	12,328	4,122
	Eco	273	20,514	15,156	5,358
	Total	638	40,945	30,440	10,505
All	All	802	54,534	40,351	14,183

Table 12. Number of speakers and number of syllable-final /s/ tokens in the Spanish Broadcast News corpus.

In contrast with the CallHome Spanish corpus, the number of tokens of syllable-final /s/ by each speaker is extremely skewed. As Table 13 indicates, over 400 of the speakers in the corpus have less than 10 tokens of syllable-final /s/ in their speech, while 10 have over 1,000 and another 12 have between 500 and 1,000.

#Tokens	# Speakers
< 10	396
10-49	263
50-99	46
100-499	67
500-999	12
> 1000	10

Table 13. Number of speakers having the number of syllable-final /s/ tokens in each range in the Spanish Broadcast News corpus.

The speech is orthographically transcribed in the Broadcast News corpus in a similar manner as the CallHome Spanish corpus. Timing information is provided to correlate the text and the speech. For each speaker turn, a start time and end time for the turn is indicated, and within a speaker turn, a timing tag was generally included every few seconds. Each speaker has a unique identifier within a speech file. When possible, this identifier is the name of the speaker, allowing the speech of a given speaker to be identified as belonging to the same speaker across speech files. For each speaker, the gender and a general dialect classification are given using the following categories: (1) coastal (Caribbean, lowland); (2) interior (mainland, highland); (3) Peninsular (Spain); (4) non-native (but still speaking Spanish). For this study, Peninsular and non-native speakers are not considered.

The Spanish Broadcast News audio files are of considerably higher quality than the CallHome Spanish audio files. The files in this corpus were all recorded at 16,000 Hz, which means that the spectral content of /s/ is present in the files and can be seen on a spectrogram (although the corpus does contain a number of speakers talking on the telephone – see below).

4.3.2 Broadcast News Data Labeling

I generated the alignments and labels automatically using the HTK Hidden Markov Model Toolkit, a set of software tools designed for building Hidden Markov Model (HMM)-based speech processing tools. Using the Spanish Broadcast News speech data with its accompanying orthographic transcriptions and the Spanish lexicon, I trained the acoustic models via bootstrapping using the procedure listed below. An acoustic model was generated for each of the allophones defined in the Spanish lexicon.

This description of the process is to illustrate some of the key steps in training the HTK acoustic models and using these models to generate phoneme-level transcriptions and alignment data. This is not a complete list of the steps needed to train the models, but only includes those key steps. I wrote Perl scripts to perform each of the steps. A complete tutorial on the use of HTK is included in Young *et al.* (1995).

4.3.2.1 Data Preparation

1. **Division of the speech and transcript files**. A Perl script parsed each of the transcript files. The script identified each utterance or speaker turn in the transcript file, and then created a separate speech and transcript file for each time-aligned portion of speech.

2. **Creation of the dictionary**. In this step, a Perl script first created a complete list of all of the words used in all of the transcripts along with the pronunciation. In order for HTK to run, all of the words in the transcripts must have a pronunciation in the dictionary file, so for words that weren't in the Spanish lexicon, the Perl script automatically generated a pronunciation using the automatic pronunciation rules from the orthographic representation. Words that were not found in the Spanish lexicon described above included mostly fragments of words as well as some proper names and foreign words.

In the HTK bootstrapping procedure, for the initial training of the acoustic models, only one pronunciation (the canonical one) of each word is used. For the case of underlying syllable final /s/, the initial acoustic models were created using a dictionary that listed the surface phoneme listed in the Spanish lexicon ([s] or [z], depending on the context). This means that in the first iteration of acoustic model training, all occurrences of [s] in the canonical pronunciation are assumed to be present. Therefore, the resulting initial acoustic model for the phoneme /s/ attempts to model all realizations of underlying /s/: [s], [h], [z] and ***deleted***. In later iterations of the acoustic model training, during re-alignment of the data, the recognizer decides if an /s/ is pronounced as [s], [h], [z] or deleted, and the retraining of the acoustic models is based on this decision.

3. **Creation of the transcription label files.** A perl script converted the transcript files from the Broadcast News mark-up format to the specific format required by HTK for the word-level *Master Label File* (MLF), which consists of all of the word-level transcriptions. Using the word-level transcriptions and the dictionary created in step (2), I used the HTK tools to generate an initial phone-level transcription.

4. **Coding of the speech into parameters.** In this step, a sequence of feature vectors based on Mel Frequency Cepstral Coefficients (MFCCs) was calculated from the speech waveform in each audio file. The acoustic models generated in the subsequent steps are based on these MFCCs. I used the HTK tools to code each speech utterance and then stored the resulting MFCCs in a separate file.

Several parameters must be set for the coding, and in this case, I set them to generic values (for example, a window size of 25 ms, 12 cepstral coefficients).

4.3.2.2 Creation of Monophone HMM acoustic models

5. **Initialization of the HMMs**. A speech recognizer's acoustic models may be based on *monophones* (one acoustic model per phoneme, regardless of the context) or *triphones* (each phoneme has a number of separate acoustic models, each one for a different preceding and following phonetic context). When training acoustic models using bootstrapping, the monophone acoustic models are trained first, and then triphone acoustic models are trained later (if needed). The first step in creating the monophone acoustic models is to initialize all of the models by setting each acoustic model's HMM values (which are the mean and variance of the Gaussians of the MFCCs) equal to the global mean and variance of the MFCCs across all of the files. Because the goal of the first iteration of training is to generate rough estimates of the HMMs that are subsequently refined, it is not necessary to use all of the speech data in this step, but instead a representative subset of the speech data. I performed this initialization and the re-estimation described in the next step using only about 200 of the individual speech files, out of a total of over 27,000.

6. **Re-estimation of the HMMs**. Training of the acoustic models is performed in an iterative manner. In this step, I used the HTK tools and the phoneme-level transcriptions to first estimate the HMM values for the initialized HMMs, and

then re-estimate the values two more times. Each subsequent re-estimation is based on the previous estimates, so the models improve with each iteration.

7. **Realignment of the training data.** The acoustic models that were estimated prior to this step were based only on the canonical pronunciations (i.e., assuming that every underlying /s/ was pronounced). In this step, I used the HTK recognition tool to perform a *forced recognition.* In the forced recognition mode, the recognizer uses the orthographic transcription to determine the sequence of words that were spoken, and is therefore only aligning the acoustic data with the known utterance. In addition to the orthographic transcriptions, the dictionary containing both the canonical pronunciations and the alternate pronunciations of syllable-final and word-final /s/ were used in this alignment procedure to generate a phoneme-level transcription. For each word with syllable-final /s/, the recognizer automatically chose the pronunciation for each word that best matched the acoustic data. The result is a file with all of the time-aligned phoneme-level transcription, with the underlying syllable-final /s/ transcribed as [s], [h], or 0 (or [z] if the canonical pronunciation contained a [z] – see below).

8. **Re-estimation of the HMMs**. As in step 6, I then performed three iterations of HMM re-estimation, this time using the phoneme-level transcription created in the previous step. Because this transcription incorporates the variable syllable-final /s/ aspiration and deletion rules, the re-estimation of the model for [s] only includes those instances where the recognizer had determined that the underlying /s/ was actually pronounced as an [s]. Likewise, the model for [h]

9. **Realignment of the training data**. After re-estimating the HMMs, I ran another iteration of forced recognition on the speech data, and the recognizer once again determined if each /s/ was pronounced as [s], [h] or 0, based on the improved [s] and [h] acoustic models from step 8. The final monophone alignment generated in this step includes the phoneme-level transcription along with the alignment times for each phoneme. The HTK-generated alignment times are to the closest 10 ms.

After the final monophone alignments were generated, the next step in training a fully-functional speech recognizer would be to copy each monophone HMM into separate triphone HMMs for each phonetic context that the phoneme is found in, and then to re-estimate the HMMs iteratively in a procedure similar to the procedure described above for the monophones. However, this was unnecessary for the current project, and in fact degraded the overall accuracy of the automatically generated coding.

4.3.2.3 Issues encountered during the HTK coding procedure

Telephone speech. I had to run the complete training procedure a number of times as issues with the data and lexicon were discovered. One of the more notable issues discovered was the fact that the Broadcast News transmission includes some speakers speaking over the telephone. In one of the first runs of the training procedure, the final alignment showed that there were a quite a few speakers having almost no tokens of syllable-final /s/ labeled as [s], but with many tokens identified as [h]. Upon listening to these speakers, I discovered that the high percentage of [h] for many of these speakers was due to the telephone quality speech, which does not include high frequency content. Therefore, I had to go back and listen to all speakers with a high percentage of tokens identified as [h] to determine if they were speaking via the telephone. If that speaker's speech was telephone speech, I eliminated all of his or her utterances from the training procedure and subsequent analysis.

Selection of pronunciation alternates available in the dictionary. The selection of the pronunciation alternates listed in the pronunciation dictionary has proven to have a great effect on the performance of the recognizer in the labeling of syllable-final and word-final /s/. During the data collection process, I attempted to allow four alternate pronunciations of syllable-final and word-final /s/: [s], [h], [z], and deletion. In an attempt to get better division between the retention categories [s], [h], and [z], I made several adjustments to the training procedure, including modifying the point at which the dictionary with multiple pronunciations is introduced, the number of estimation and realignment iterations performed, and modifying which pronunciation alternates were available during the realignment of the training data (ex., limiting the environments where [h] and [z] could occur). None of the alternate methods have made a significant improvement in the distinction of [s], [h] and [z]. The recognition of [z] was especially problematic. Despite the fact that performance on word-internal [z] was quite good, I was unable to get HTK to accurately categorize tokens as [z] in word-final position, and so I had to remove [z] as an alternative for recognition in word-final position. Therefore, all of the data reported here is generated as described above – with [s], [h] and deletion available for each syllable-final or word-final /s/, and [z] also available in those locations where the canonical pronunciation was listed with [z] (i.e.,word- internally when followed by a voiced consonant).

Monophone acoustic models vs. triphone acoustic models. I attempted to continue the training procedure through the development of the triphone acoustic models several different times, but the results of the time alignments were worse with the triphone models than with the monophone models. I believe this was because the triphone models take into account the variation that occurs in each context. For example, if the typical realization of an /s/ is stronger when followed by a vowel than by a consonant, and there are separate triphone models representing the /s/ before a consonant as compared to a vowel, the model before a consonant would be for a weaker /s/. When the recognizer is subsequently used in forced recognition mode to determine the absence or presence of an /s/ using these triphone models, the criteria for the cut-off for labeling an /s/ as deleted would be weaker before a consonant than before a vowel.

4.3.3 Categorical Data

Table 14 shows the categorical classifications for the Broadcast News corpus. In this corpus, there is a higher rate of /s/ retained as [s], and a much higher rate of aspiration, when compared to the CallHome corpus.

Dialect	**N**	**[s]**	**[h]**	**[z]**	**deletion**
coastal	13,589	52.1	11.6	1.3	35.0
interior	40,945	71.7	7.0	1.3	20.0
Total	54,534	66.9	8.1	1.3	23.7

Table 14. Percent of tokens with each coding by dialect for the Broadcast News Corpus.

Even though there are much higher rates of aspiration in this corpus than in the CallHome corpus, aspiration still occurs at lower rate than reported in the literature. There are several possible explanations for this. First, although the recordings are of much better quality than the CallHome recordings, it is possible that the original speech did contain a larger proportion of aspiration, but that the frication associated with [h] is not discernable in the recordings due to its low intensity. Lipski (1983a, 1985) identifies the transmission quality in the radio broadcasts as an issue that prevents analysis of three categories [s], [h], and deletion. Instead he uses just two categories, retained and weakened /s/ (aspiration or deletion) in his study of radio broadcasts due to the difficulty of identifying aspiration.

Another possible reason for the low rate of aspiration in this corpus may lie in the speaking style used. As discussed above, Lipski (1983a, 1985) reports much lower rates of aspiration and deletion in news broadcasts than in conversational speech. Since the speech in this corpus is much more carefully pronounced than the informal speech that has largely been the subject of previous studies reporting high rates of aspiration, it is possible that the more formal broadcast news speech may exhibit less aspiration.

A third possible explanation for the low aspiration rate may be due to the nature of the "aspiration." Although the process of aspiration is normally described impressionistically in the literature as a categorical change from the phoneme [s] to the phoneme [h], in fact the result of aspiration may not be pronounced as the segment [h]. Widdison (1995a, 1995b, 1997) performs several perceptual experiments where he replaces a vowel in a non-/s/ environment with a vowel from another word where the vowel is followed by /s/. Although there is no frication of the /s/ in the acoustic signal, such tokens are perceived by native speakers of Spanish as containing an underlying /s/ nearly half the time, and when they are identified as having an underlying /s/, they are always identified as being from informal speech. From this, Widdison concludes that the result of aspiration is actually realized as "breathy voice quality over the preceding vowel margin which largely mimics [h]," not as a separate phoneme [h] (1997:258). If this is the case, it is not surprising that the results in this study would identify much fewer occurrences of [h], because most previous studies are based on impressionistic listening without the use of a spectrogram. If, as Kerswill and Wright (1990:264) point out, the fact that transcriptions are a sequence of segments "predisposes the transcriber to hear a series of discrete, completely articulated segments," human listeners may tend to transcribe an [h] even if what they hear is just some breathiness in the preceding vowel It would therefore not be surprising that the phoneme would not be included in the automatic labeling performed by HTK, or by hand-labeling that was done via listening in conjunction with careful inspection of the spectrogram. In addition, Terrell (1979) corroborates the finding that aspiration does not always result in a well-pronounced [h] sound; his category 'aspiration' includes both what he describes as an [h] (voiceless or voiced) and gemination of the following consonant, even if there is no frication.

4.3.4 Acoustic Measurements

I have calculated several acoustic measurements that may correlate with syllable-final /s/ lenition from the time alignment and phoneme-level transcripts using perl scripts and xwaves/entropic tools. There are two reasons for looking at the acoustic measurements. First, while the automatically generated codes for the two categories [s] and *deletion* have high reliability measures, the code [h] (reflecting partial lenition) has lower reliability (See section 4.3.5.1 below). If acoustic measurements can be found that correlate well with the distinction between [s] and [h], then it is still possible to model the process of aspiration separately from that of total deletion. Secondly, although previous researchers have categorized the weakening of syllable final /s/ into the three categories [s], [h] and deletion, the process is actually phonetically gradual, so continuous variables are more sensitive and are more appropriate to model the phenomenon. Terrell (1977) indicates that his categories "***s***", "***h***" and ***deletion*** actually encompass several phonetic variants. For Terrell, "***s***" includes [s], [z] and the partially voiced [s^z], with the voiced variants only occurring prior to a voiced segment. The category "***h***" includes [h], voiced [h], assibilated [h^s], and a geminated following consonant (with no frication). The category ***deletion*** includes only tokens where there is no trace of the underlying /s/ whatsoever.

Lenition has been described as a decrease in vocal effort (Kirchner 1998) or as gestural blending and overlap (Browman and Goldstein 1992). In this view, there should be both temporal and spectral effects that can be measured acoustically. Lavoie (2001) studies the effect of position and stress on consonants in U.S. English and Mexican Spanish in order to determine which articulatory and acoustic characteristics correlate with weakening. Lavoie finds both gradient and categorical effects of weakening in the consonants that she studies. Articulatorily, Lavoie measures the extent of linguopalatal contact using electropalatography, and she finds that consonants in the weaker positions (word-internal or non-pre-stress) exhibit less linguopalatal contact for a shorter duration than consonants in the stronger positions (word-initial or pre-stress). Acoustically, Lavoie finds that the characteristics of duration, intensity, voicing and the degree of formant structure correlate with the weakening of consonants, with a decrease in duration being the primary acoustic measure of consonant lenition.

In the case of the first stage of weakening of syllable-final /s/ (aspiration), the decrease in vocal effort should correspond to a decrease in the frequency content (a shift from /s/ to /h/) and a decrease in the overall intensity of the /s/. There should also be a decrease in the duration of the /s/. Widdison (1995b) does find that a weakened /s/ is associated with a shorter duration, lower frequency content, and lower amplitude.

Another possible measure of the strength of a syllable-final /s/ lies in the vowel quality of the immediately preceding vowel. Claims have been made that in some dialects of Spanish, there is a phonemic distinction between vowels in open syllables and vowels in closed syllables. (Navarro Tomás 1966, Honsa 1965). This distinction is one of vowel quality, with the open variant of the vowel occurring in closed syllables and vice-versa. Although the claim is that the vowel contrast remains when a syllable-final /s/ is deleted, (and thus would not show an effect of syllable-final /s/ aspiration and deletion) vowel quality is a factor to be investigated to see if it correlates with /s/ lenition.

In the present study, the effect of lenition on these acoustic measures may be somewhat obscured if their correlation with lenition is slight, because unlike other studies such as Lavoie's, the speech used for this study is not carefully elicited 'laboratory speech'. Lavoie is able to carefully control the utterances in her study: each of the words are in carrier phrases, with the consonant under investigation only in intervocalic context, either word-initial or word-medial, and either in a pre-stress or a non-pre-stress position, and she controls the speech rate. By doing this, she is able to control for effects that cannot be controlled in a large spoken corpus such as the Broadcast News corpus – such as the speaker, speech rate, and phonetic context. Thus, in the measurements in this section, only large effects may be seen. In the following chapters, the measurement data that is shown to correlate with lenition will be reported for these independent factors that cannot be controlled for.

4.3.4.1 Duration

I have calculated several measurements of duration directly from the automatically generated time alignments. The duration that I expect to correlate best with /s/ lenition is the duration of the /s/ itself. There is obviously an effect on duration when the /s/ is deleted completely, but there should also be a measurable gradient effect on duration when the /s/ is retained but weakened to [h]. I also calculated the duration of the immediately preceding vowel and the following segment to determine what if any measurable compensatory lengthening there is when /s/ is weakened or deleted.

Figure 1 shows a boxplot of the duration of the [s], [h], or [z] for the tokens that were hand-labeled. I have used the durations generated using HTK, but they are plotted according to the hand-corrected label because the goal is to determine if there is a correlation between the different perceptual categories and the automatically generated acoustic measurements. The duration is not shown for tokens that were hand-labeled as deleted, because their duration is always 0. In Figure 1, the boxes represent the middle 50% of the values (from the 1st quartile to the 3rd quartile). The line inside the box indicates the median value. The vertical "whiskers" coming out of the boxes extend to the lowest and highest values, excluding outliers, and the stars represent outliers.

The boxes in Figure 1 show that there is a difference in the distributions of the duration of [s] and [h]; the median value of tokens labeled [s] is 80 msec while the median value for [h] is 50 msec. There is no overlap in the boxes for [s] and [h], although there is significant overlap of the whiskers. This indicates that there is a tendency for tokens of [h] to be shorter than tokens of [s], but this tendency is not categorical. Thus, duration of /s/ can be used as a gradient measure of the extent of /s/ lenition.

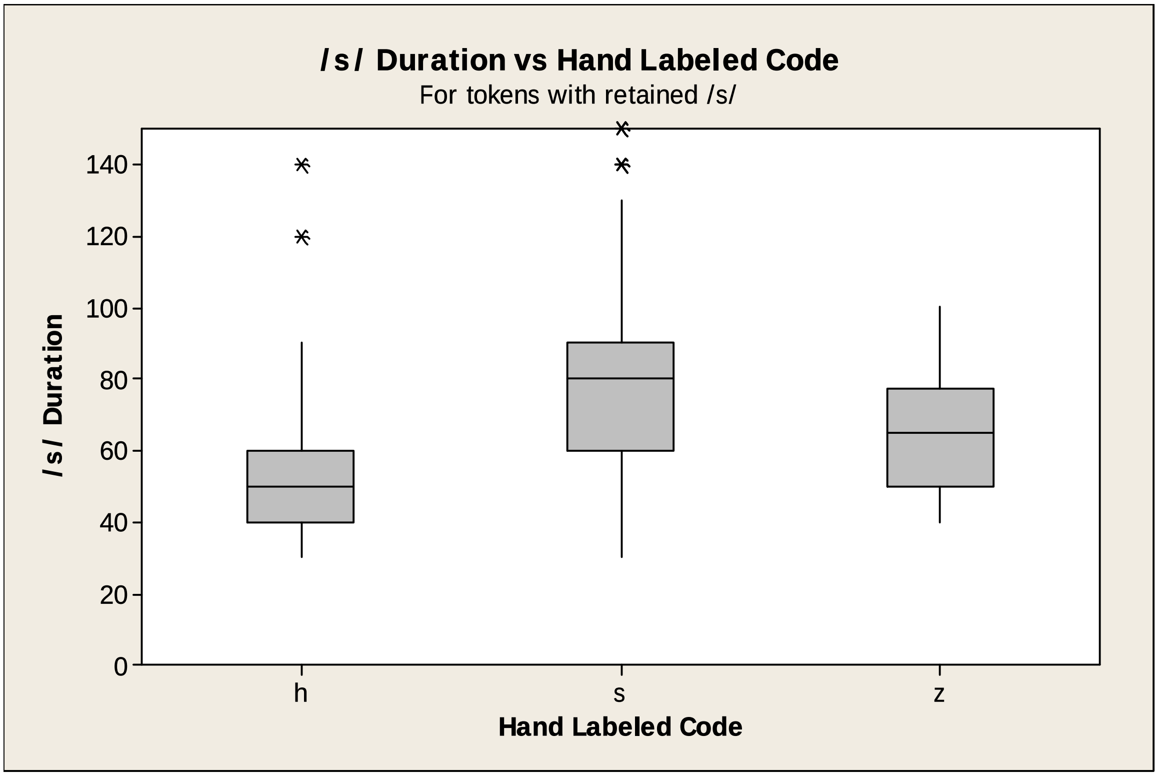

Figure 1. Boxplot of /s/ duration for hand-labeled tokens with retained /s/.

In addition to the duration of the /s/ itself, it has been suggested that the duration of the immediately preceding vowel or the following segment may be affected by syllable-final /s/ lenition due to compensatory lengthening. Lipski (1983a:243) goes as far as asserting that even when there is no trace of the /s/, temporal cues prevent minimal pairs from being confused. While not claiming that temporal cues to the presence of an underlying /s/ always exist, Terrell (1979) does state that compensatory lengthening (gemination) of the following consonant does happen with some regularity. Because he considers a geminated following consonant to be a weakened - but not completely deleted - /s/, he includes tokens displaying gemination of the following consonant in his category *'h'*. In order to test for compensatory lengthening, I have measured the duration of the preceding vowel and the following segment.

Figure 2 shows a boxplot of the duration of the immediately following segment, separated by context and hand-labeled code, for the tokens that were hand-labeled. If tokens labeled *deletion* exhibit gemination of the following consonant, I would expect the deleted category to have much longer following segment durations than the retained categories. This is partially instantiated; in the pre-consonant and word-internal conditions, the box representing the middle 50% of tokens for the category *deletion* is somewhat higher than that for [s], indicating that there may be some compensatory lengthening involved. However, the median value for these two categories only differs by 10 ms (60 ms vs. 50 ms). In the pre-vowel context, the effect of deletion cannot be evaluated with this hand-labeled data due to the small sample size of deleted tokens.

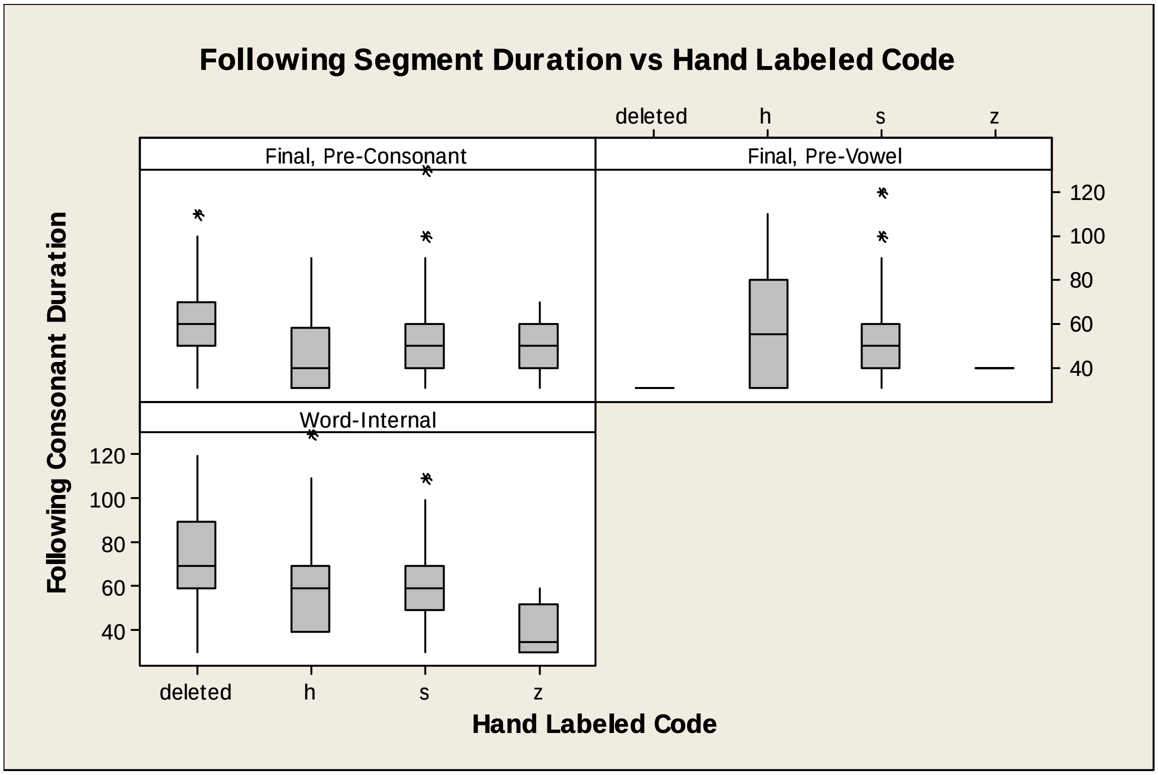

Figure 2. Duration of the immediately following segment, by segment type and hand-labeled code.

Figure 3 shows a boxplot of the duration of the immediately preceding vowel by hand-labeled code. All four of the following contexts behave the same, so this data is not broken out by following segment type. There is essentially no difference in the preceding vowel length for tokens coded [s] as compared with ***deletion***, indicating that there is no compensatory lengthening of the vowel in this data. This contradicts the results of Resnick and Hammond's (1975) comparison of vowel duration preceding a deleted underlying syllable-final /s/ vs. no underlying syllable-final /s/ in Cuban Spanish. They found that there was no difference in vowel length in word-final pre-consonantal position, but that there was a difference in word-internal pre-consonantal position, with vowels between 33% and 40% longer in the words with following (unpronounced) /s/ than in the words with no underlying /s/.

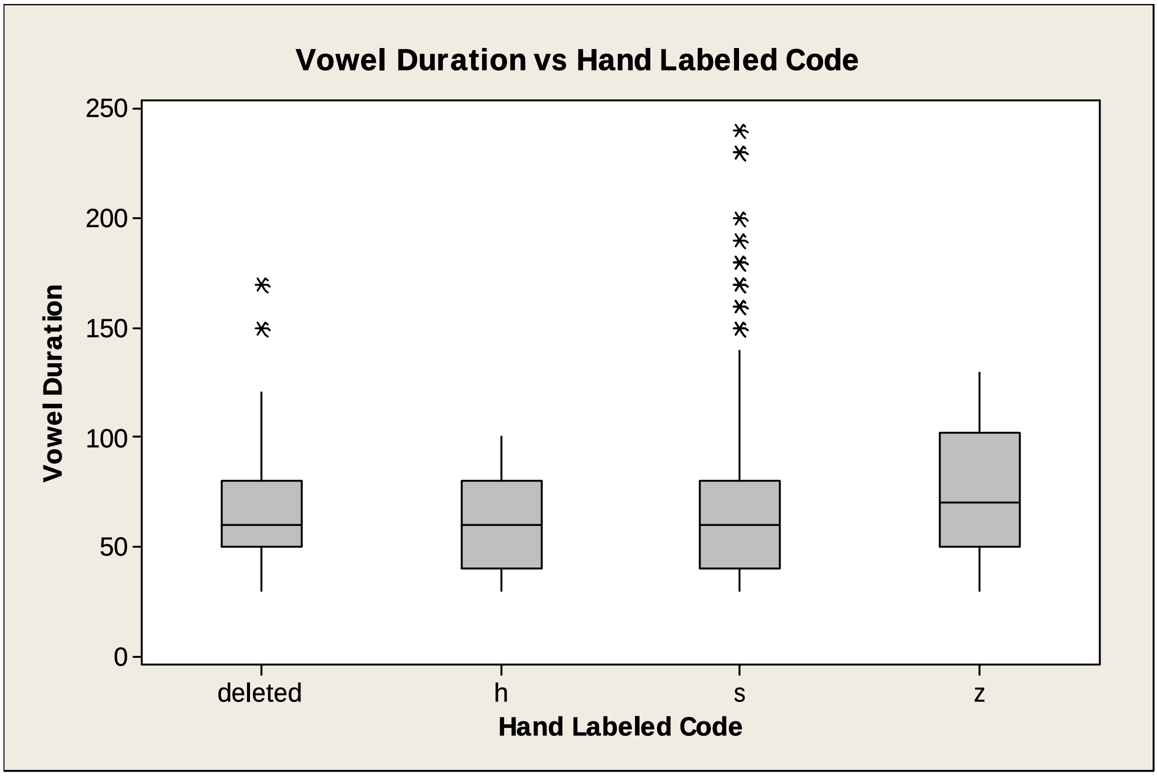

Figure 3. Duration of the immediately preceding vowel by hand-labeled code.

Because the hand-labeled data contains a limited number of tokens, and only 5 speakers, I have also calculated the durations for all of the automatically generated data in Table 15, separated into the categories ***retained*** ([s], [h], and [z]) and ***deleted***. This data confirms the findings for the hand-labeled data. There is little evidence of compensatory lengthening of the preceding vowel in any of the contexts, for either speakers of the interior or coastal dialects. There appears to be a small amount of compensatory lengthening of the following segment in word-internal position in both dialects, and in final/pre-consonantal position for the interior dialect. However, as shown in the "Total" column, the total duration of the preceding vowel, /s/ and following segment is consistently much longer when /s/ is retained than when it is deleted.

Dialect	Context	/s/ status	Preceding Vowel	/s/	Following Segment	Total
Coastal	Word-Internal	retained	68	67	65	200
		deleted	66	0	87	153
	Final, Pre-Consonant	retained	79	77	61	217
		deleted	69	0	67	136
	Final, Pre-Vowel	retained	70	97	68	235
		deleted	73	0	66	140
	Final, Pre-Pause	retained	89	108		197
		deleted	70	0		70
Interior	Word-Internal	retained	59	74	56	189
		deleted	64	0	75	139
	Final, Pre-Consonant	retained	65	70	48	184
		deleted	73	0	61	135
	Final, Pre-Vowel	retained	60	94	57	211
		deleted	60	0	58	118
	Final, Pre-Pause	retained	84	108		191
		deleted	63	0		63

Table 15. Average duration of the immediately preceding vowel, the syllable-final /s/, and the following segment, for all automatically generated data.

From my data, while there is some evidence of a small amount of compensatory lengthening of the following consonant when an /s/ is deleted, it does not appear that any additional insights into syllable-final /s/ lenition can be gained from the duration of either the preceding vowel or the following segment.

4.3.4.2 Spectral Moments of /s/

I have calculated two measurements of the spectral content of the /s/, the first and second spectral moments. The first spectral moment of a fricative, also known as the spectral Center of Gravity (COG), is a measure of the average frequency of the fricative's energy (Harrington and Cassidy 1999). It is calculated by the equation

$$mom_1 = \frac{\sum fI}{\sum I}$$

where I is the amplitude (in dB) and f is the frequency (in Hz). The acoustic measurement of spectral COG has been shown to correlate well with perceptual categories of fricatives (Harrington and Cassidy 1999, Shadle and Mair 1996) as well as with articulation as measured by electropalatalography (Tabain 2001, Hoole *et al.*1989). Since the phoneme [s] is characterized by a concentration of energy at high frequencies, its spectral COG is relatively high. In contrast, the fricative [h] contains a higher concentration of lower frequency content, so its spectral COG is lower.

The second spectral moment is a measurement of the diffusion of spectral energy across frequencies (similar to the variance of the spectral COG). It is calculated by the equation

$$mom_2 = \sqrt{\frac{\sum f^2 I}{\sum I} - mom_1^2}$$

This second spectral moment has been shown to be larger for the weaker, flatter fricatives such as [h] (Shadle and Mair 1996).

For each syllable-final /s/, I calculated the spectral moments across a 20 msec (320 sample) frame centered at the middle of the /s/. To calculate the frequency components, the frame was filtered with a Hamming filter, and the entropic tool for the FFT (Fast Fourier Transform) performed a 512-point FFT. The equations for the first and second spectral moments were then implemented for the spectral components between 1000 and 8000 Hz. As Shadle and Mair (1996) point out, the spectral moments are highly affected by the frequency range used in the calculation. I chose to not include frequencies lower than 1000 Hz to reduce the effect of any lower frequency background noise. Since the sampling rate of this corpus is 16000 Hz, the highest spectral content is 8000 Hz.

Figure 4 shows a boxplot of the Spectral COG of the hand-labeled tokens, separated by the hand-labeled code. As expected, the median Spectral COG of tokens labeled [h] is much lower (3216 Hz) than those labeled [s] (4965 Hz). There is no overlap of the middle 50% of tokens, although there is overlap of the whiskers. Because the Spectral COG does show an effect of the loss of oral constriction from [s] to [h], it is a useful measure of the gradient nature of /s/ lenition.

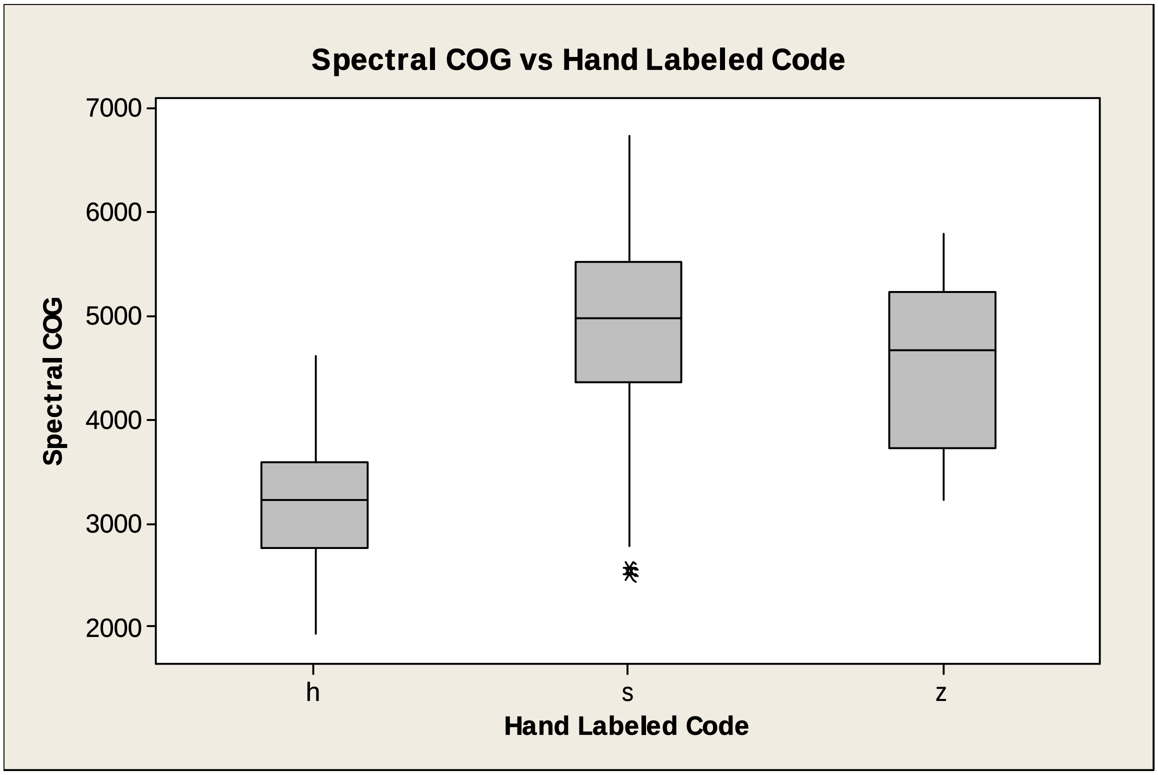

Figure 4. Spectral COG (in Hz) by hand-labeled code, for tokens with retained /s/.

Figure 5 shows a boxplot of the Spectral Second Moment of the hand-labeled tokens, separated by the hand-labeled code. Since the distributions of [h] and [s] are nearly identical, this is not a good measure of /s/ lenition.

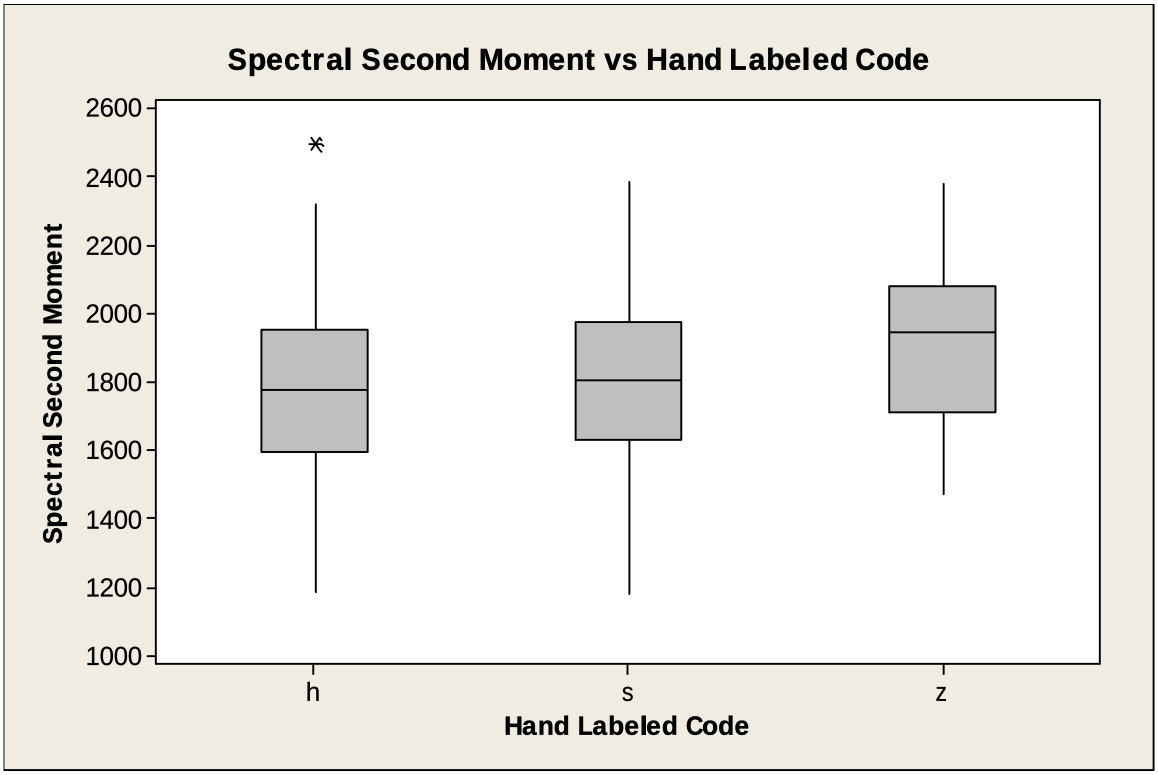

Figure 5. Spectral Second Moment (in Hz) by hand-labeled code, for tokens with retained /s/.

4.3.4.3 Intensity

I have calculated the intensity of the /s/ for all of the tokens with retained /s/. The intensity was calculated using the entropic utility *pwr* to calculate the average log power for the middle 1/3 of the segment. Because the unnormalized power of the /s/ is highly dependent on the overall loudness of the speech signal, the intensity of the syllable-final /s/ must be calculated as a relative measurement (in dB) with respect to some reference. Previous studies have used different portions of the speech signal as the reference to normalize the intensities of consonants, such as the average across the whole word containing the target consonant (Cole, Hualde and Iskarous 1999), the vowel two syllables preceding the target consonant (Lavoie 2000), or the immediately preceding vowel (Manrique and Massone 1981). I have chosen to use the immediately preceding vowel as the reference because the context of the tokens of syllable-final /s/ is highly variable in this speech corpus. Because the tokens of syllable-final /s/ occur in different words, the average power of the word would not be a consistent normalizing factor. Using a vowel that is one or two syllables away from the syllable-final /s/ also would not work, because many of the tokens are in the first syllable or the last syllable of an utterance. To normalize the value, the log power of the reference vowel was subtracted from the log power of the /s/ (because the logarithm of the power is used, this is equivalent to a ratio of the raw powers).

Figure 6 shows a boxplot of /s/ intensity for hand labeled tokens with each of the retained codes [s], [h], and [z]. The higher the intensity value, the more intense the frication. Although [h] has been shown to have a lower power than [s] in Australian English (Harrington and Cassidy 1999), my data does not show such an effect; there is essentially no difference between the values for [s] and the values for [h]. Furthermore, there is a great amount of variability in the intensity values of tokens labeled [s], as shown by the long whiskers with respect to the box and the large number of outliers. This variability suggests that the measure is not a good predictor of [s]. There are several possible causes of the large variation in the intensity for tokens labeled [s]. First, using the immediately preceding vowel to normalize the power may be problematic because the vowel may have a different 'baseline' power depending on the vowel's identity. It is also possible that coarticulation with the following segment causes variation in the intensity of the /s/. For example, Widdison (1995a) notes that the anticipatory raising of the velum when a velar follows an /s/ reduces the airflow and thus its intensity.

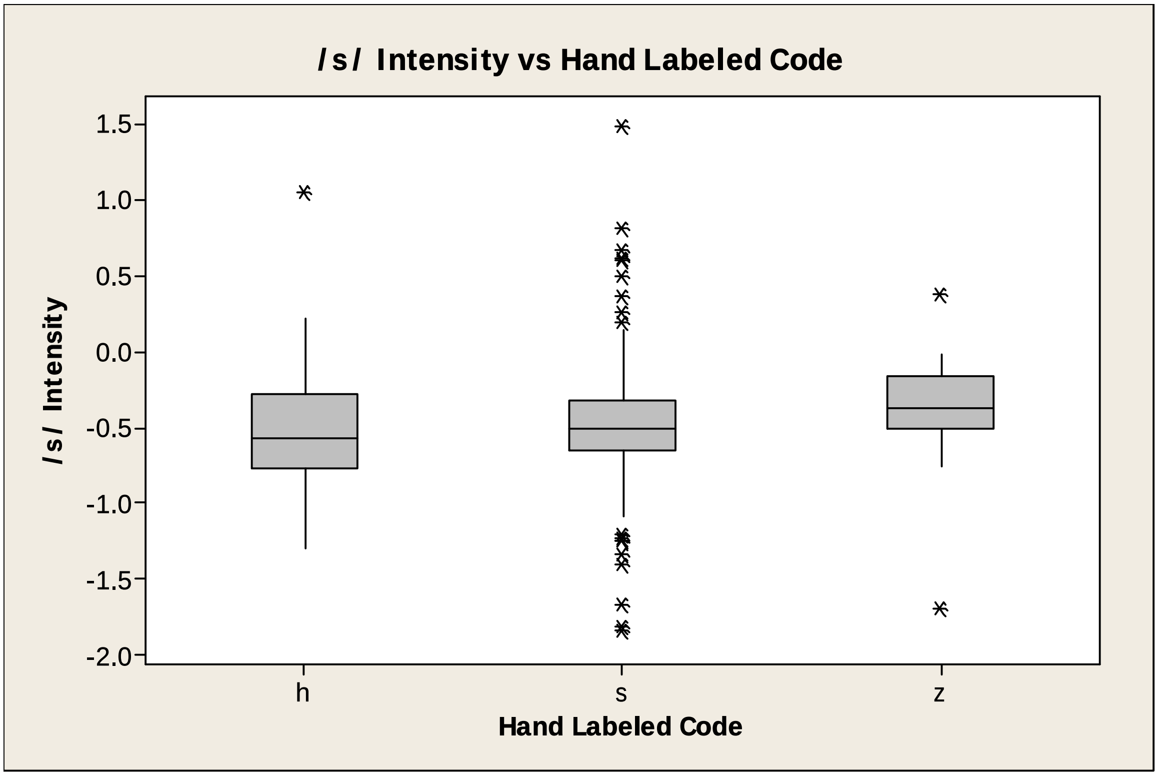

Figure 6. /s/ Intensity (measured in dB) with respect to the immediately preceding vowel.

Because the /s/ intensity measurement cannot distinguish [h] and [s], and because the intensity of [s] has such high variability, this measure does not appear to be a useful measure of lenition of syllable-final /s/. In addition, Figure 16 below comparing the /s/ intensity calculated from the automatically generated time alignments and the hand-labeled time alignments shows that there are a number of tokens with rather large discrepancies between the automatically generated values and the values for the hand-labeled data.

4.3.4.4 Preceding vowel formants

I also measured the first and second formants of the immediately preceding vowel because the vowel quality of some vowels in Spanish has been described as being affected by whether the vowel occurs in an open or closed syllable. I calculated the vowel's formants by averaging the values generated by the xwaves formant tracker over the middle 1/3 of the vowel (as determined by the phoneme-level time alignments). I have chosen to report data only for preceding vowels /e/ and /o/, because these vowels have been the most widely reported to have a phonemic height distinction based on whether they occur in an open or closed syllable. If no effect of syllable-final /s/ lenition is found on the formants of these two vowels, it is unlikely to be present for other vowels.

Figure 7 and Figure 8 show boxplots for the first and second formants of /e/ and Figure 9 and Figure 10 show the formants for /o/, for the hand-labeled tokens. Because formant values are highly dependent on the speaker's vocal track, I have separated the data by speaker. There are no consistent differences in either F1 or F2 of either /e/ or /o/ across the codes [s], [h], and *deletion*. In all four plots, for most of the speakers, the boxes representing the middle 50% of tokens overlap, and the median values are nearly identical. This means that vowel formants cannot be used as a measure of syllable-final /s/ lenition.

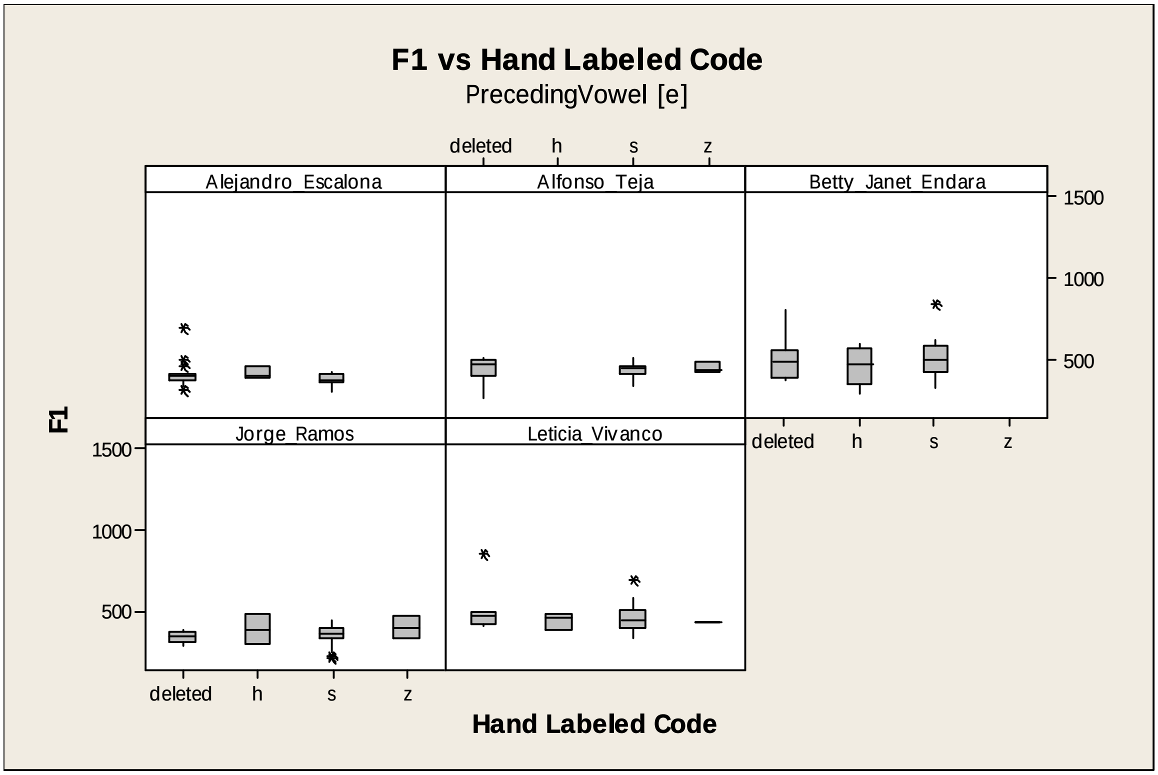

Figure 7. Boxplot of F1 of preceding vowel [e] by speaker for each of the hand labeled codes.

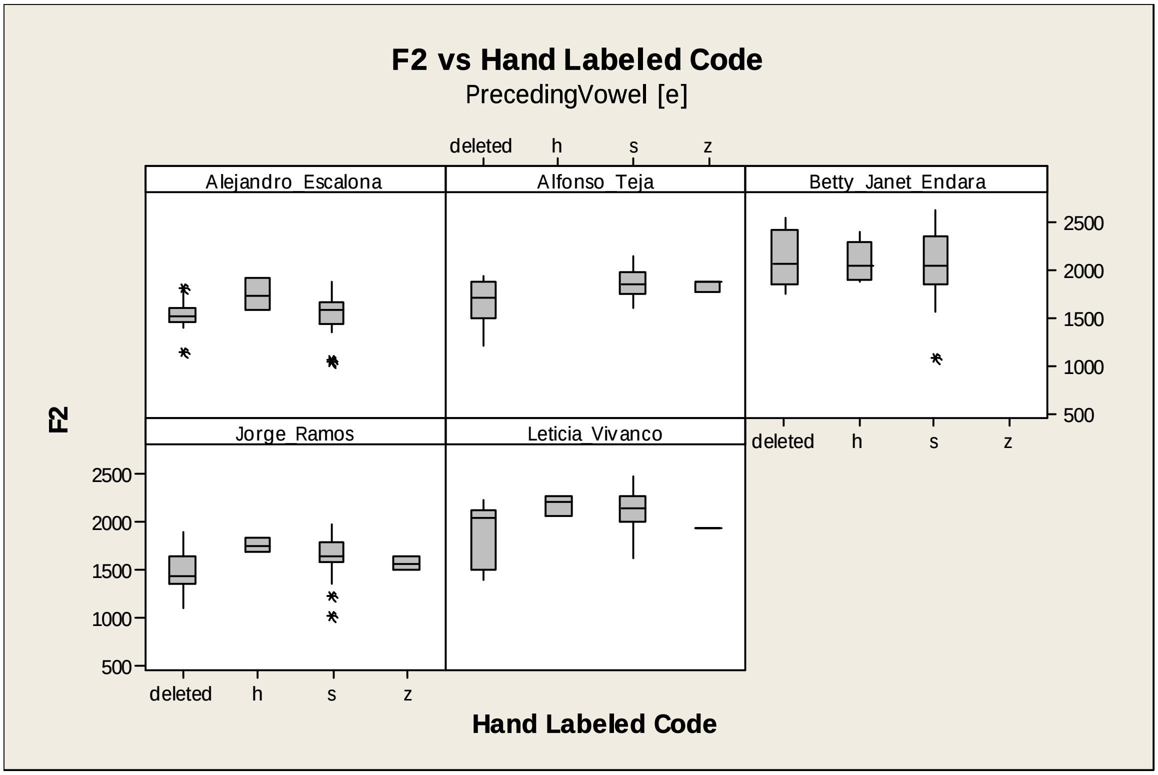

Figure 8. Boxplot of F2 of preceding vowel [e] by speaker for each of the hand labeled codes.

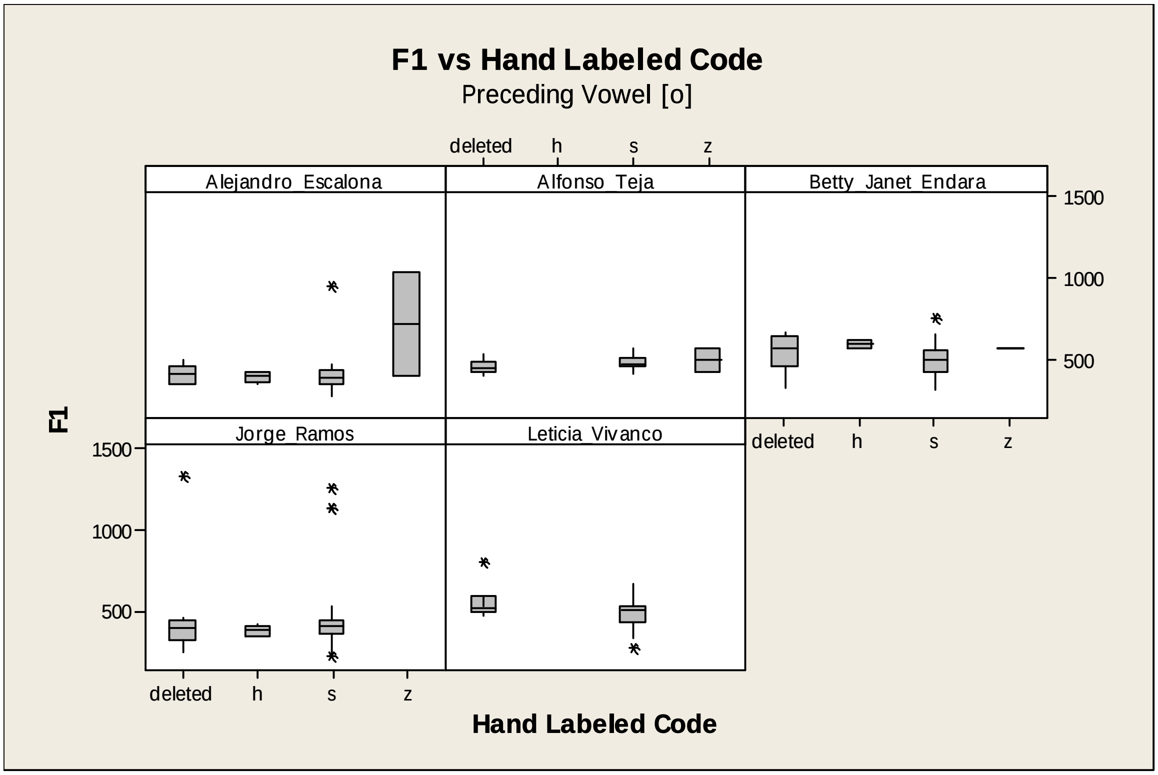

Figure 9. Boxplot of F1 of preceding vowel [o] by speaker for each of the hand labeled codes.

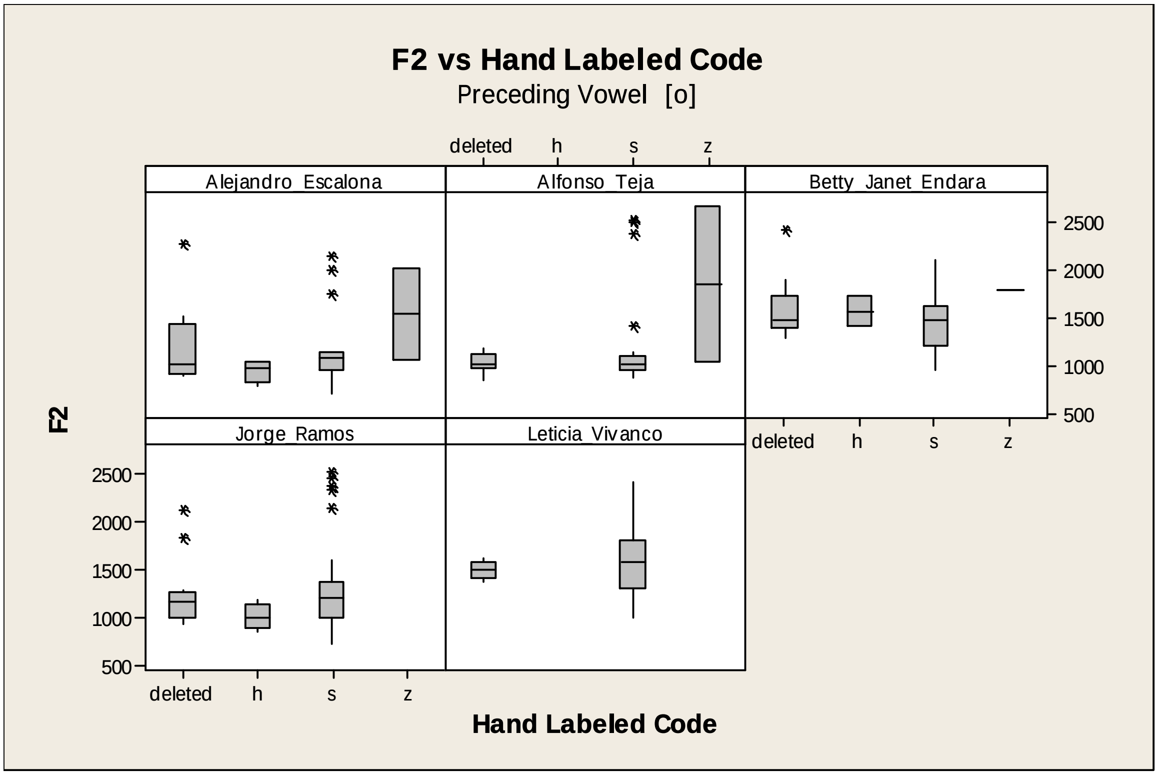

Figure 10. Boxplot of F1 of preceding vowel [o] by speaker for each of the hand labeled codes.

Note that this result does not necessarily mean that these speakers do not have a difference in vowel quality for the vowels /e/ and /o/ in open and closed syllables. Both when the /s/ is retained as [s] and when the /s/ is deleted, the preceding vowel is in an underlyingly closed syllable. It is possible that these speakers retain the "closed syllable" vowel quality even when the /s/ is deleted. Resnick and Hammond (1975) performed a slightly different experiment that was meant to look for a difference in vowel quality between open and underlying closed syllables. They compared vowel quality in Cuban Spanish for vowels preceding an underlying syllable-final /s/ that was deleted (a closed syllable) with the vowel quality for vowels in the same context but with no underlying syllable-final /s/ (an open syllable). They found no measurable difference in the vowel formants.

4.3.4.5 Summary

As the preceding discussion has shown, the combination of duration and Spectral COG measurements should correlate well with syllable-final [s] lenition. Figure 11 shows a scatterplot of /s/ duration and spectral COG for each of the tokens that were hand-labeled. As the boxplots in the preceding sections indicated, and this scatterplot confirms, tokens labeled [h] generally have a lower spectral COG and shorter duration than the tokens labeled [s].

The fact that there is a continuum and actually some overlap between [s] and [h] in both the spectral COG and duration confirms that the process is indeed gradient. Since these two measurements do correlate with a more reduced production of /s/, they will be useful to measure the gradient process, and will allow analysis of lenition within the categories. This should provide additional insights to the investigation than can be found by simply looking at discrete categories.

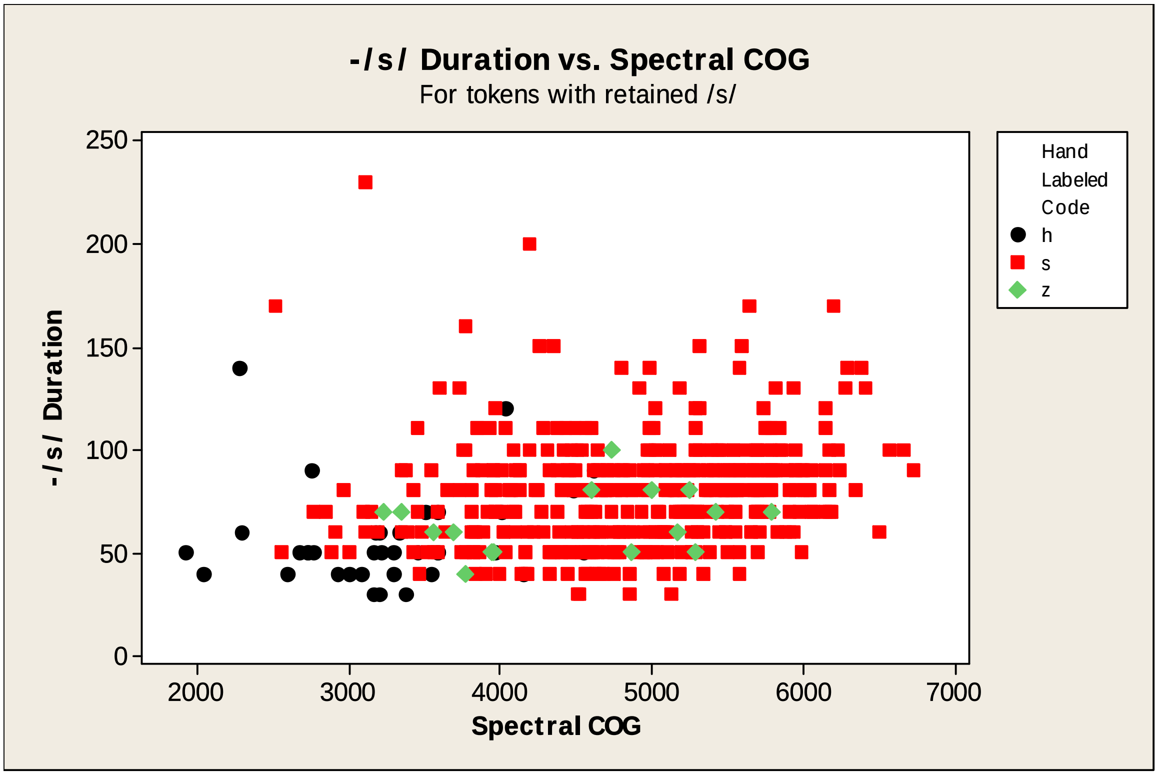

Figure 11. Duration and Spectral COG for hand-labeled tokens from the Broadcast News Corpus.

Table 15 shows the average values for two of the measurements made for the Broadcast News, for all of the tokens labeled [s] or [h]: the average duration of the [s] or [h] in msec., and the average spectral COG. The average duration for the coastal speakers is about the same as for interior speakers, and the average spectral COG is much lower. This is an indication that even for the tokens of underlying syllable-final /s/ that are retained, they are more [h]-like for the coastal speakers, as expected.

Dialect		N	% del	Ave. /s/ duration (ms)	Ave. spectral COG
Coastal	Word-Internal	3,678	40.4	67	3900
	Final, Pre-Consonant	5,102	51.9	77	3973
	Final, Pre-Vowel	3,987	8.1	97	4414
	Final, Pre-Pause	1,822	21.0	107	4202
Interior	Word-Internal	10,505	13.1	73	4946
	Final, Pre-Consonant	16,392	33.8	70	4794
	Final, Pre-Vowel	8,475	4.7	93	5034
	Final, Pre-Pause	5,583	15.4	108	4983

Table 16. Summary of measurement data.

4.3.5 Data Accuracy

4.3.5.1 Categorical Data

After generating labels and time alignments for all of the speech in the Broadcast News corpus using perl scripts and the HTK tools, I hand-corrected approximately 120 tokens from each of 5 speakers. I did this in xwaves using both the waveform and the spectrogram. The automatically generated label file and automatically generated formants were displayed on the screen and corrected as necessary. The characteristics that I corrected were the categorization of the syllable-final /s/ as [s], [h], [z], or deletion, the time-alignment for the /s/, the preceding and following segments, and the formant tracking. Because the intensity and spectral COG measurements are calculated from the numerical values of the signal, they could not be hand-corrected. The overall performance of HTK in generating the data from the Broadcast News is fairly good. Spot listening of the label files generated by HTK indicates fairly accurate alignment.

Table 17 shows a comparison of the HTK-generated codes and the codes that were hand-labeled. The automatically generated codes are very accurate for those tokens labeled [s]. Performance of [h] is the lowest, with only 70.0% agreement. These consist primarily of tokens that were automatically labeled [h] but hand-coded [s]. When [s], [h], and [z] are all combined into one ***retained*** category, overall accuracy is 96.2%.

	Hand-labeled code				% Agreement
	s	h	z	deleted	
s	388	2	7	1	97.5
h	14	35	1	0	70.0
z	1	0	8	0	88.9
deleted	10	6	6	122	84.7

Table 17. Comparison of HTK-generated codes and hand-labeled codes.

Table 18 shows the κ measurements for each of the codes between the automatically labeled data and the hand-labeled data. Both [s] and deletion have very high values of κ, indicating nearly perfect agreement between the automatically generated codes and the hand-labeled data. The other codes of [h] and [z] have lower values of κ, although even for [z] there is moderate agreement; after removing the agreement that can be attributed to chance, half of the codes [z] are in agreement. When the [s], [h] and [z] categories are combined, the overall value of κ increases to 0.89, indicating that after correcting for chance, nearly nine out of ten judgments are in agreement.

	Kappa
[s]	0.87
[h]	0.73
[z]	0.50
deleted	0.89
Overall	0.84

Table 18.Reliability measurement κ for the Broadcast News corpus.

In their comparison of human listener vs. speech recognizer categorization of variable phonological rules in Dutch, Wester *et al.* (2001) report a median κ value of 0.55. In each of the processes they study, the speech recognizer rarely indicates a segment is present when the expert listeners identify it absent. They speculate that this difference occurs because the speech recognizer requires segments to be at least a minimum duration, so when the segment is weakened and has a duration shorter than that minimum, the speech recognizer will identify it as deleted while the listeners still hear it. Wester *et al.* conclude that although there are some consistent differences between the identifications made by expert listeners and those of the speech recognizer, there are many advantages of using the speech recognizer for the investigation of phonological variation. My results also demonstrate that the speech recognizer identifies a slightly higher deletion rate than the hand-labeling does, but the high overall agreement indicates that the automatic labeling of the categorical phonological rule of syllable-final /s/ deletion is very accurate.

4.3.5.2 Measurement Data

Figure 12 shows a comparison scatterplot of duration that was automatically generated from the HTK time-alignments and the hand-labeled duration for the tokens that were automatically labeled [s]. The data is separated into 4 four panels, one for each of the following segment contexts. Figure 13 and Figure 14 show the same data for the tokens that were automatically labeled [h] and [z]. When automatically generating the labels using HTK, the label [z] was only available in the word-internal location when followed by a voiced consonant, which is why the only context that contains tokens is Word-Internal for [z].

Due to the architecture of the speech recognition system, all of the time-alignments generated using HTK must coincide with the edge of a speech window. The parameters that I set for HTK created a series of 25 msec speech windows, overlapping so that the beginning of each window was 10 msec after the beginning of the immediately preceding window. This means that the time-alignments for the beginning and ending of each segment (and therefore also the durations) are only accurate to the closest 10 msec. Since the time alignments made by hand have a finer granularity, the hand-labeled durations and the automatically generated ones should be close but not equal. As shown in Figure 12, Figure 13, and Figure 14, this is the case; the large majority of tokens have automatically generated durations that are close to the hand-labeled durations. Accuracy of the automatically generated durations is lowest in the pre-pausal environment. Many of the tokens in the graphs are just higher than the reference line, indicating that the /s/ durations identified by HTK are slightly longer than the durations using hand-labeling (an average of 6.8 ms longer). This is primarily due to HTK identifying the beginning of the phoneme /s/ just prior to the beginning of frication.

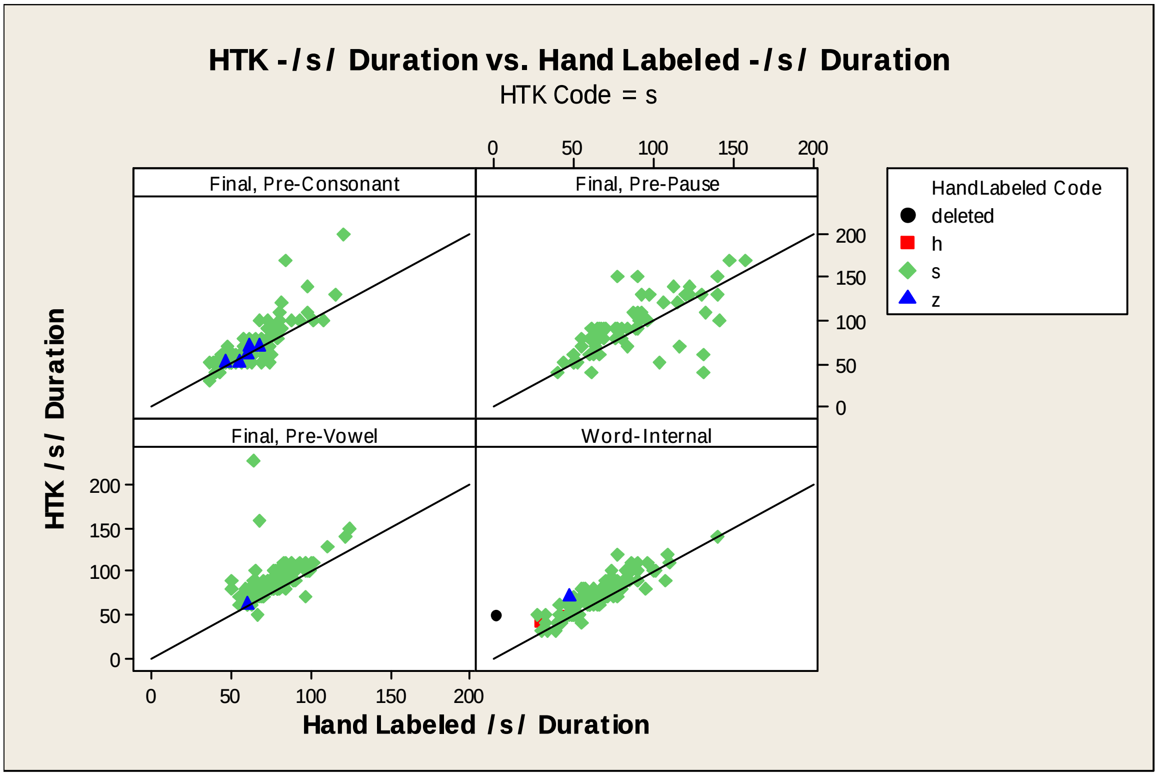

Figure 12. Comparison of label and duration (automatically generated vs. hand-labeled), for automatically generated label [s].

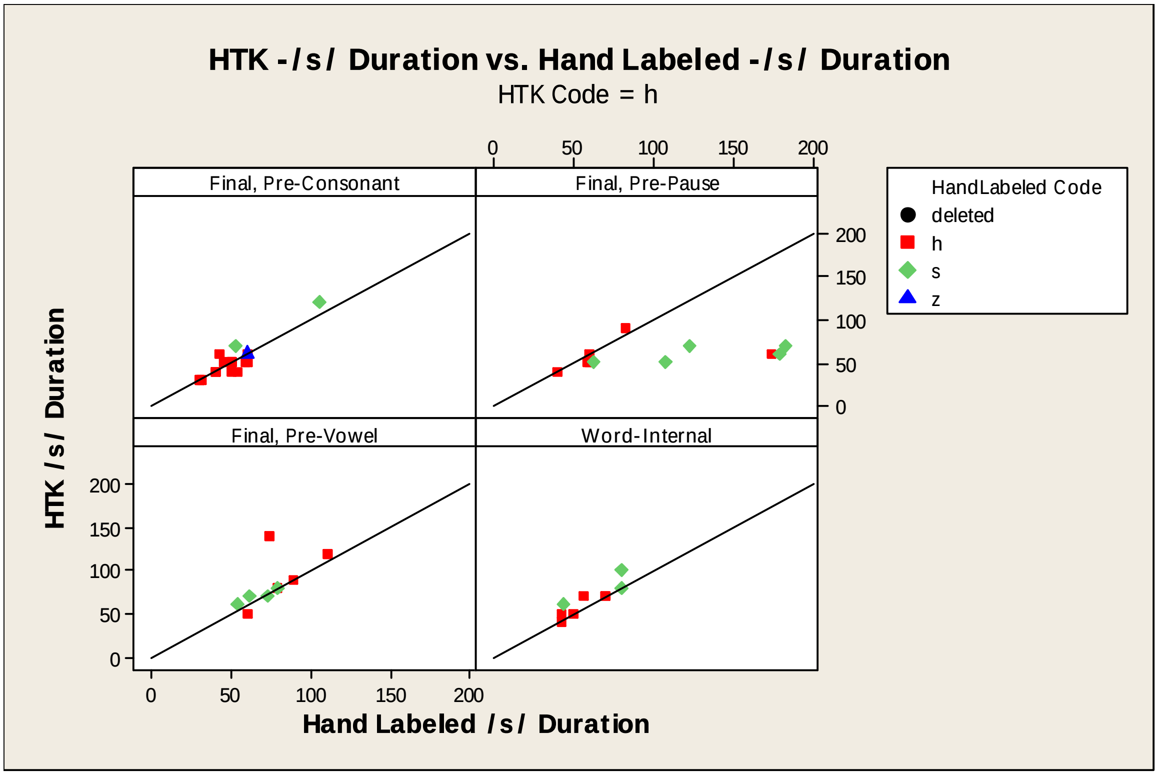

Figure 13. Comparison of label and duration (automatically generated vs. hand-labeled), for automatically generated label [h].

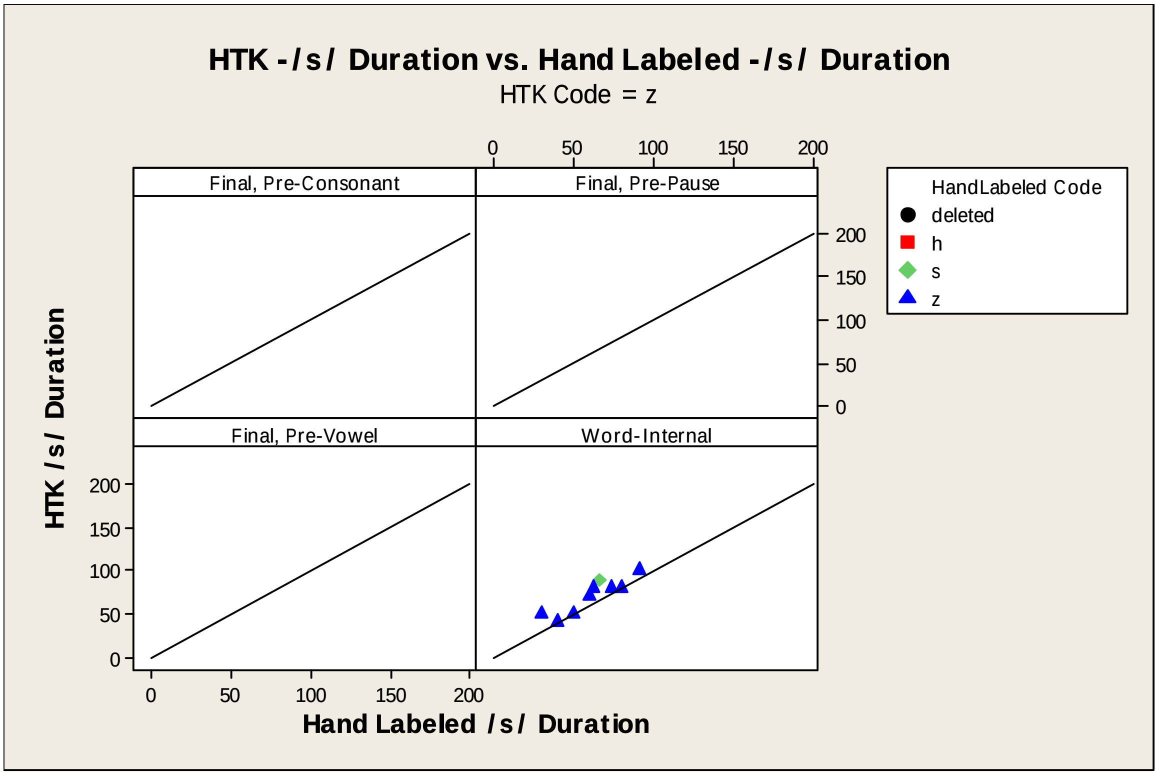

Figure 14. Comparison of label and duration (automatically generated vs. hand-labeled), for automatically generated label [z].

Figure 15 and Figure 16 display scatterplots comparing the measurement data for the Spectral Center of Gravity and the /s/ intensity for tokens that were labeled as retaining -/s/ as [s], [h], or [z]. Both of these are measurements that are calculated from the numerical values of the speech signal during the specified time interval, and they cannot be verified using the waveform or spectrogram. Thus, the difference in the values generated using the HTK alignment and the hand-labeled alignment are due to the difference in the time period used to calculate these measurements, which in turn is based on the alignment of the beginning and end of the segment. Since both the Spectral COG and intensity should be relatively constant during the middle of the segment, and should show significant change only at the beginning and end of the segment, I calculated these values using only the middle of the segment (the middle 1/3 of the segment for intensities and the middle 20 msec for the Spectral COG). In this way, if the time alignments were slightly off, the portion of the segment used for the calculation should still fall within the relatively constant portion. For the spectral COG, the tokens do fall very near the reference line where the HTK and hand-labeled values are equal. The tokens for /s/ intensity do not fall along the line as well, with many of them having higher intensity from HTK than from the hand labeling.

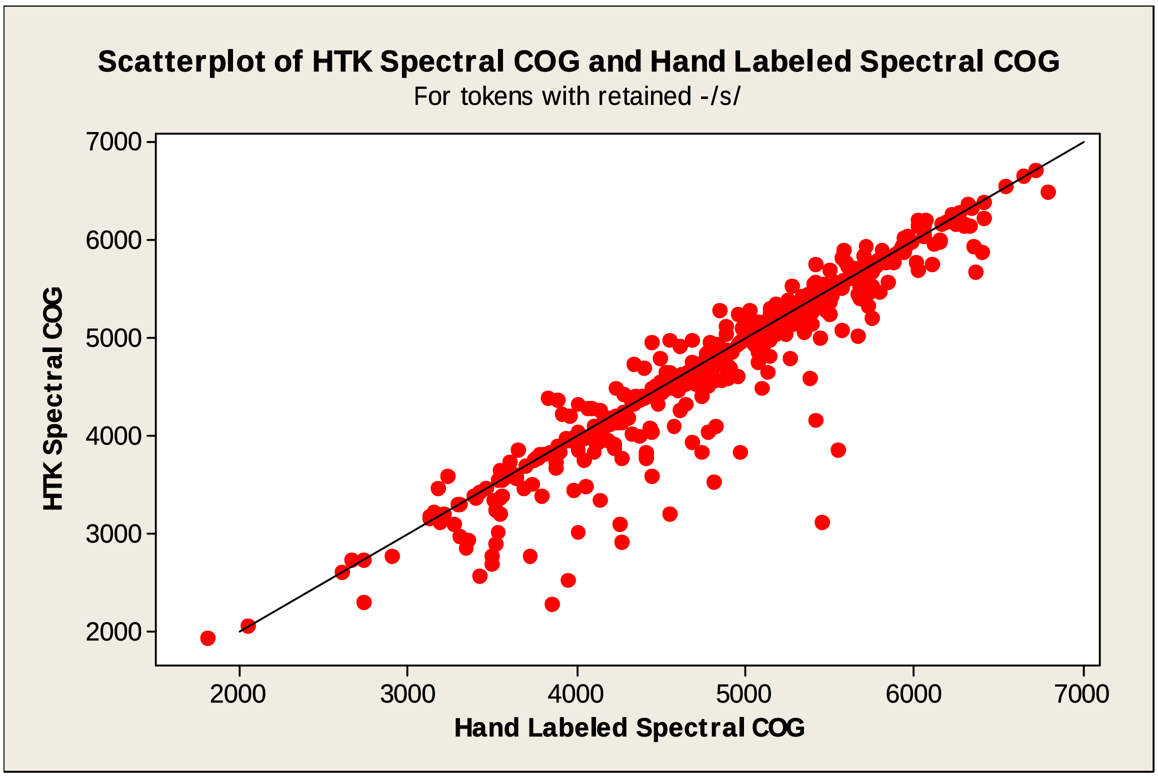

Figure 15. Comparison of spectral COG calculated from automatically generated time alignments and hand labeled time alignments.

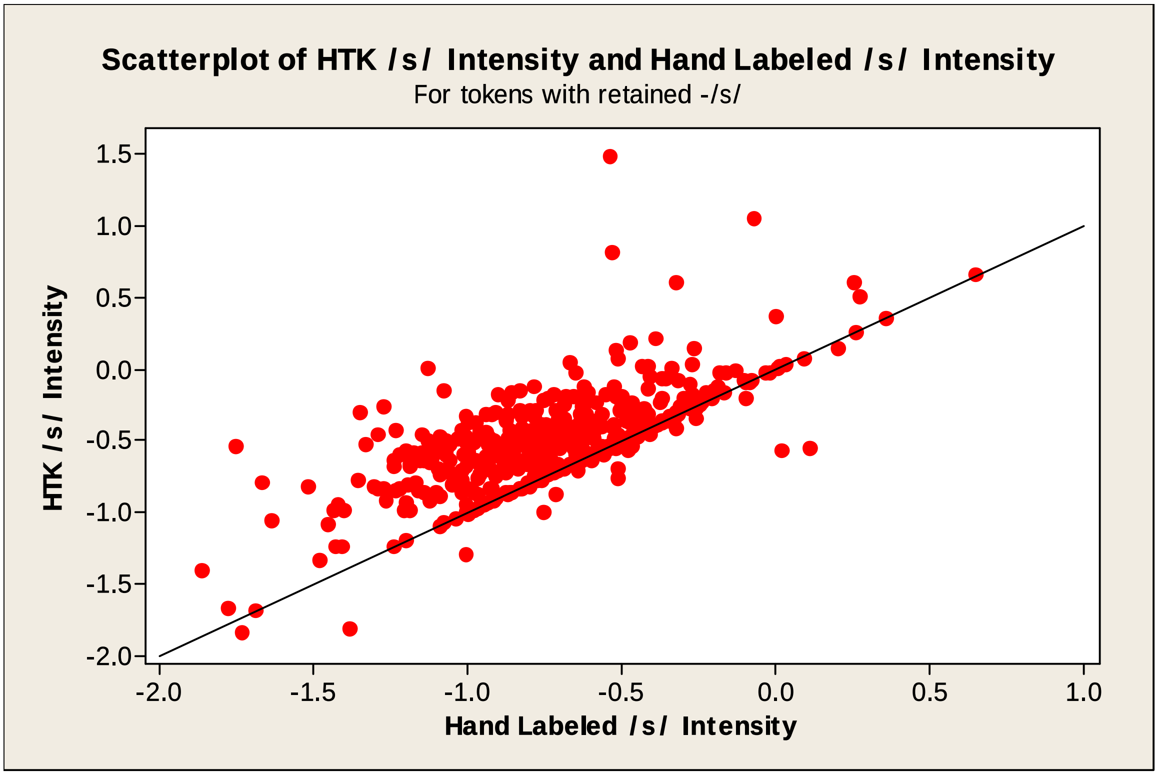

Figure 16. Comparison of /s/ intensity calculated from automatically generated time alignments and hand labeled time alignments.

Table 19 shows a comparison of the average values for each of the acoustic measurements using the automatically generated phoneme-level time-alignments and using the hand-labeled time alignments. Overall, the validity of the automatically generated /s/-intensity values are somewhat questionable, but all of the other values are accurate enough to use for the present study.

Measurement	HTK	Hand-Labeled	Difference
Preceding vowel duration (ms)	66.7	74.3	-7.5
/s/ duration (ms)	76.9	70.1	6.8
Following segment duration (ms)	53.1	53.5	-0.4
/s/ intensity (dB)	-0.49	-0.72	0.23
Spectral COG (Hz)	4739	4802	-63
Spectral 2nd Moment (Hz)	1802	1799	3

Table 19. Comparison of average values of acoustic measurements using HTK and Hand-labeling.

4.4 Summary

As has been described above, both the CallHome and the Broadcast News corpora were coded for [s], [h], [z], and deletion. However, the [h] and [z] categories do not appear to be particularly useful for the present study for two reasons. First, they have significantly lower relative frequencies in comparison to [s], particularly for the CallHome corpus. Secondly, there is low confidence in accurately applying the three different codes [s], [h], and [z] for the retained category, as shown in the lower agreement between human coders for the CallHome corpus, and lower agreement between the automatically generated labels and the hand-labeling for the Broadcast News corpus. But as we have seen, the overall κ values for both corpora are good when all the retained categories are combined. Therefore, in the data analysis in the following chapters, the classifications [s], [h], and [z] will all be combined into a "retained" category, and only the deletion percentages will be reported.

For the CallHome corpus, no measurements were made, which means that the only analysis possible with the data from this corpus is the categorical decision “retained” vs. “deleted”.

For the Broadcast News corpus, additional analysis can be made using the measurement data. The measurements that are accurate enough and correlate well with lenition are the duration of the /s/ and the Spectral COG.

5 Independent Variables

For each token of syllable-final /s/, the values for a number of independent variables to be investigated for their contribution to syllable-final /s/ lenition have been calculated for further statistical analysis. These factors include the linguistic and extra-linguistic factors that have been previously identified by other researchers as well as a number of usage-based factors.

5.1 Linguistic and extra-linguistic factors

5.1.1 Speaker

I have coded several factors related to the speaker for each of the tokens of syllable-final /s/. These are: (1) a unique identifier for each speaker; (2) sex of the speaker; (3) dialect. In addition, I have included for each token of syllable-final /s/, the overall deletion rate of the speaker.

In both of the speech corpora used, the speaker for each speech turn is identified, and within each file, every speaker has a unique identifier in each speech file. In the CallHome corpus, each speaker is only in one file. In the Broadcast News corpus, many of the speakers (particularly the news anchors) are identified by name, and therefore have the same identifier if they appear in more than one speech file. Figure 17 shows the proportion of deletion for each of the 195 speakers in the CallHome corpus with 25 or more tokens of syllable-final /s/, sorted by their proportion of deletion. In the figure, each individual speaker is identified by a number. This figure shows that there is a great variety of individual speaker deletion rates, varying from low-deletion speakers (7%) to extremely high-deletion speakers (91%). There is a significant number of speakers who have higher rates of deletion; 96 out of the total 195 – almost half – have deletion rates of 40% or higher. Figure 18 shows the same data for the 201 speakers in the Broadcast News corpus with 25 or more tokens. The Broadcast News corpus has a similar range of deletion (0%-93%), but there are fewer speakers with the higher deletion rates. Only 43 of the 201 speakers (about 21%) have deletion rates above 40%. The disparity between the two corpora is probably due primarily to the difference in the speech style, with the more formal broadcast speech style having fewer speakers with high deletion rates. There may also be a difference in the dialect mix of the speakers in the two corpora. Both of these figures show that speakers fall on a continuum.

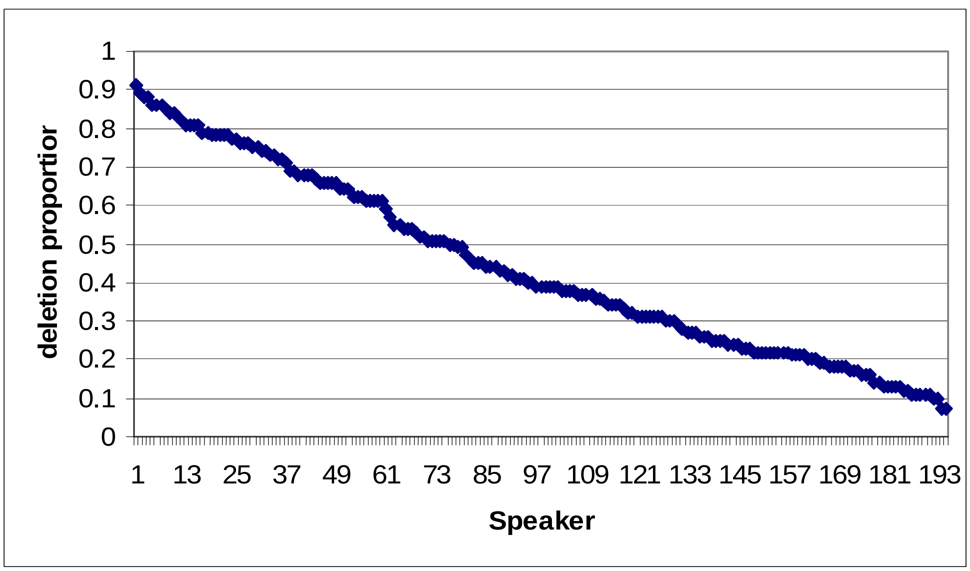

Figure 17. Deletion proportion for 195 speakers in the CallHome corpus.

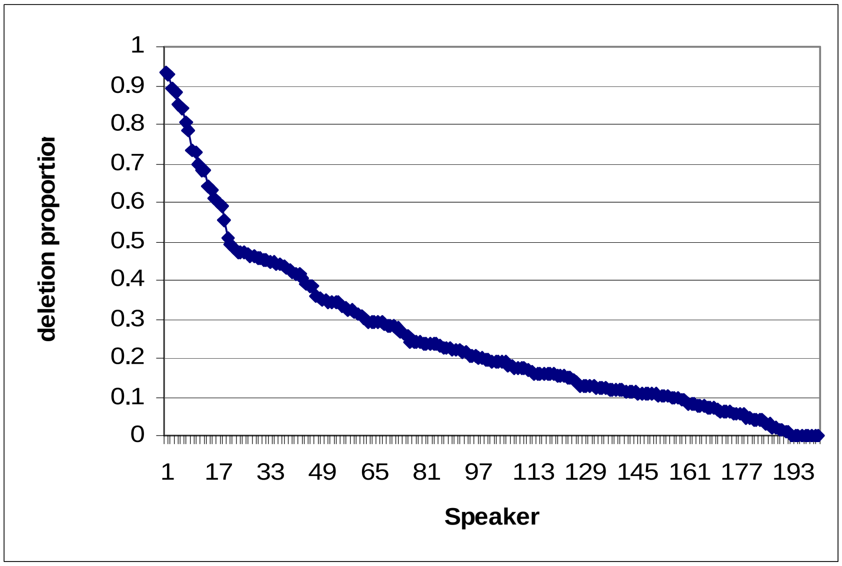

Figure 18. Deletion proportion for 201 speakers in the Broadcast News corpus.

It is not possible to do a statistical analysis of the data set using the speaker's identity as one of the factors because of the large number of speakers (over 200 in the CallHome corpus and over 800 in the Broadcast News corpus). However, for each of the tokens of syllable-final /s/, I have also coded the overall deletion rate of the speaker, and this value could be used as an independent variable in a statistical analysis using data from just the speakers that have a larger number of tokens. This variable cannot be used for the full data set because there are many speakers who have a small number of tokens (in the Broadcast News corpus, almost half of the speakers have 10 or less tokens). For these speakers, due to the small number of tokens, the calculated deletion rate may not be representative of their true deletion rate.

The deletion data separated out by country for the CallHome corpus was shown above in Chapter 4. The data for the two dialect regions *coastal* and *interior* is broken down by sex in Table 20 for the CallHome corpus and in Table 21 for the Broadcast

News corpus. In keeping with the description in the literature, male speakers in the interior dialects in both corpora delete /s/ somewhat more often than females. For the coastal dialects, the males delete more in the CallHome corpus, while the females display a very slightly higher deletion rate (35.3% vs. 34.9%) than the males in the Broadcast News corpus. This difference is probably not significant. The duration data for the Broadcast News corpus exhibits the same tendencies as the deletion data, with the durations of /s/ for males and females about the same in the coastal dialects and somewhat shorter durations for the males than the females in the interior dialects.

Dialect	Sex	N	% deletion
Coastal	Male	4,666	64.4
	Female	5,255	63.1
Interior	Male	5,222	28.3
	Female	6,122	26.8

Table 20. Percent deletion for the CallHome corpus by dialect type and sex.

Dialect	Sex	N	% deletion	/s/ duration	Spectral COG
Coastal	Male	9,416	34.9	86	3993
	Female	4,173	35.3	86	4439
Interior	Male	29,967	21.1	81	4856
	Female	10,978	17.0	85	5097

Table 21. Percent deletion and measurement data for the Broadcast News corpus by dialect type and sex.

Table 22 shows the deletion rate and the measurements for the 10 speakers in the Broadcast News corpus who have more than 1,000 tokens of syllable-final /s/. Their deletion rates vary from 9% to 45%. It is interesting to note that all of the interior speakers have a significantly higher rate of deletion in the word-final position than the word-internal position, whereas only one of the coastal speakers has a higher rate of deletion in the word-final position (Betty Janet Endara, 43% vs. 38%). The Spectral COG measurements of the interior speakers are consistently higher than those of the coastal speakers. There does not appear to be a difference in the average /s/ durations between the two dialects.

Speaker	Dialect	Total		Word -final				Word-internal			
		N	% del	N	% del	Ave. /s/ dur. (ms)	Ave. spectral COG	N	% del	Ave. /s/ dur. (ms)	Ave. spectral COG
Jorge Ramos	interior	5868	24	4395	27	73	4710	1473	14	62	4805
Alfonso Teja	interior	3547	14	2604	19	86	5503	943	2	76	5534
Cesar Cardoza	coastal	2416	45	1720	41	78	4221	696	54	58	3689
Leticia Vivanco	interior	2112	19	1542	23	78	5121	570	9	72	5226
Luis Felipe Arce	interior	1643	17	1238	19	88	4536	405	10	66	4258
Betty Janet Endara	coastal	1607	42	1167	43	81	4433	440	38	64	3935
Jacobo Zabludovsky	interior	1505	18	1053	22	102	5346	452	7	79	5454
Moises Lopez	coastal	1214	28	918	28	97	4104	296	28	62	3794
Armando Guzman	interior	1096	9	845	11	79	4840	251	3	64	4865
Aristides Quinteros	coastal	1025	23	773	22	87	3823	252	28	61	3617

Table 22. Percent deletion and measurements for speakers in the Broadcast News corpus having more than 1,000 tokens of syllable-final /s/.

5.1.2 Following segment

As described in section 2.2.1 above, the following segment has been shown to be one of the key factors influencing the extent of syllable-final /s/ lenition. For the statistical analysis in Chapter 5, I have coded not only the broad phoneme category of the following segment as has been customary in the literature (vowel, consonant, pause), but also the type of segment (glide, liquid, nasal, voiced stop, voiceless fricative, voiceless stop, pause, vowel), the place of articulation for consonants (labial, coronal, velar), and the actual identity of the segment.

Table 23 shows the deletion percentage for following segment type and position in the word for the CallHome corpus and Table 24 shows the data for the Broadcast News data. As expected, there is significantly less deletion when a word-final /s/ is followed by a vowel than by a consonant or pause and when retained, the average /s/ duration before a vowel is longer than before any of the consonants. The Spectral COG is higher before a vowel than before any of the consonants. For both of the corpora, tokens with a following pause have higher deletion rates than those with a following vowel, but lower deletion rates than tokens followed by most of the consonant types. Word-final /s/ preceding a pause has a much higher average duration than in any other context. As for the consonants, there is great variation in the extent of lenition, depending on the manner of articulation. In general, syllable-final /s/ in word-internal position deletes at a lower rate than in word-final position. In agreement with Valdivieso and Magaña (1991) a following voiceless stop causes less deletion than a following voiceless fricative. In word-final position, there is the least deletion before a voiceless stop, and a high rate of deletion before a liquid, which is similar to Brown and Torres Cacoullos' (2003) findings.

The effect of a following liquid is different depending on whether it is word-internal (where it has low deletion) or word-final (where it has high deletion).

Position in the Word	Following Segment Type	N	% deletion
Word-Internal	Liquid	10	30.0
	Nasal	157	61.8
	Voiced Stop	66	45.5
	Voiceless Fricative	18	50.0
	Voiceless Stop	6,288	31.5
Word-Final	Glide	306	62.7
	Liquid	785	71.1
	Nasal	1,396	63.8
	Voiced Stop	1,828	58.4
	Voiceless Fricative	181	59.1
	Voiceless Stop	3,649	51.1
	Pause	2,919	42.5
	Vowel	3,662	38.3

Table 23. Percentage deletion in the CallHome corpus separated by position in the word and following segment type.

Position in the Word	Following Segment Type	N	% deletion	/s/ duration	Spectral COG
Word-Internal	Liquid	375	12.3	82	4398
	Nasal	537	16.6	83	4694
	Voiced Stop	567	33.3	66	4339
	Voiceless Fricative	146	17.8	79	4853
	Voiceless Stop	12,558	20.0	72	4771
Word-Final	Glide	165	17.0	83	4315
	Liquid	2,009	44.5	79	4282
	Nasal	3,165	35.2	76	4595
	Voiced Stop	7,831	57.6	67	4540
	Voiceless Fricative	906	23.5	72	4778
	Voiceless Stop	7,583	19.2	71	4766
	Pause	7,395	16.8	108	4800
	Vowel	11,297	5.5	94	4884

Table 24. Percentage deletion and measurement data in the Broadcast News corpus separated by position in the word and following segment type.

Table 25 shows the percent deletion for consonants separated by the place of articulation for the CallHome corpus. Table 26 shows the percent deletion and measurement data for the Broadcast News corpus. For both of the corpora, a following velar exhibits a lower rate of deletion than a following labial. The role of a following coronal is conflicting in the two corpora. In the CallHome corpus, a following coronal has the lowest rate of deletion, while it has the highest rate of deletion in the Broadcast News corpus. This appears to be primarily due to the effect of a following [d]. For Broadcast News, a following [d] has a much higher rate of deletion than the other voiced stops. In word-final position, [d] has about 60% deletion (compared to about 40% for [b] and [g]), and in word-internal position, it has about 40% deletion (compared to about

10% for [b]). In contrast, the CallHome corpus has a lower rate of deletion before a [d] than the other voiced stops (about 55.2% for word-final [d] in comparison to over 60% [b] and [g]). The discrepancy in the rate of deletion of [d] identified in the two corpora is probably a result of the difference in method used to perform the coding. When /s/ is followed by a [d], the result is often a single segment that contains features of both segments. Perceptually, this single segment retains some features of the /s/, so there would be a tendency to code it as being retained in the CallHome corpus. However, if the two segments coalesce into a single segment, the speech recognizer would be more likely to label the [s] as deleted. Morris (2000) claims that a syllable-final /s/ coalesces with any following voiced stop ([b], [d], and [g]), but it is not surprising that the effect is greatest for a following [d] because [s] and [d] share the place of articulation.

	N	% deletion
Coronal	8,526	40.2
Labial	3,264	56.5
Velar	2,588	52.4

Table 25. Percent deletion by the place of articulation of the following consonant for the CallHome corpus.

Following Consonant	N	% deletion	/s/ duration	Spectral COG
Coronal	21,413	36.9	70	4635
Labial	8,825	24.8	72	4701
Velar	5,439	17.8	77	4839

Table 26. Percent deletion and measurement data by the place of articulation of the following consonant for the Broadcast News corpus.

5.1.3 Preceding vowel

Table 27 and Table 28 show the percent deletion and measurements by the preceding vowel for the CallHome and Broadcast News corpora. There is a large range of deletion rates, particularly for the CallHome corpus – from 34.3% for /u/ to 51.8% for /o/. In both corpora, tokens with a preceding /i/ have relatively low deletion rates and tokens with a preceding /a/ or /o/ have relatively high deletion rates. The behavior of /u/ is different in the two corpora: in the CallHome corpus, it has the lowest rate of deletion among the preceding vowel categories, while for the Broadcast News corpus, it has a relatively high rate of deletion.

These results are partially in agreement with Brown and Torres Cacoullos (2003) findings for preceding vowel in word-final context. They find that /i/ and /n/ have a low deletion rate and /a/ has a high deletion rate. Unlike their results, I found /o/ has the highest deletion for both corpora, while they found /o/ had moderate deletion.

Preceding Vowel	N	% deletion
a	4,527	49.7
e	10,791	40.1
i	903	38.0
o	4,516	51.8
u	435	34.3

Table 27. Percent deletion in the CallHome corpus, by preceding vowel.

Preceding Vowel	N	% deletion	/s/ duration	Spectral COG
a	10,798	23.0	86	4833
e	20,393	23.7	82	4817
i	3,960	16.5	81	4798
o	17,666	25.8	83	4637
u	1,352	24.8	80	4552

Table 28. Percent deletion, average /s/ duration and average Spectral COG in the Broadcast News corpus, by preceding vowel.

Although it appears that there is a relatively strong effect of the preceding vowel on lenition of the following /s/, it is important to note that the data for this factor is not evenly distributed with other factors. For example, 43% of the word-internal tokens have a preceding vowel [e] in the Broadcast News corpus, while [e] represents only 29% of the word-final data. Also, certain preceding vowels occur more frequently with certain following segments, for example 33% of tokens with a preceding [e] have a following [t], while only 4% of tokens with a preceding [o] have a following [t]. Because the data is skewed, the multivariate analysis in the following chapter is necessary. The multivariate analysis will measure the impact of each factor taking into account the distribution of the data.

5.1.4 Stress

For stress, I have coded the stress of the syllable immediately following the /s/ as well as the syllable that the /s/ is in the coda of. Table 29 and Table 30 show data for the factor of following stress. Although there should only be two values for following stress - *stressed* and *unstressed* - I have included a third category, *monosyllabic*. In the pronunciation lexicon, every word has one syllable marked as being stressed. This means that in the case of monosyllabic words, the single syllable is always listed as being stressed, although many monosyllabic words are actually unstressed clitics. Because there are so many tokens with a following monosyllabic word, I have kept them separate in order to avoid skewing either of the other categories.

In addition to following stress, I have the data in Table 29 and Table 30 by the following segment context because the literature indicates following stress is particularly important in the case of a following vowel, with an unstressed vowel favoring lenition (Lipski 1984a). The data in Table 29 agrees with the literature; in the CallHome corpus, there is more deletion before an unstressed vowel than before a stressed one, with a following vowel in a monosyllabic word patterning with a following unstressed vowel. For the Broadcast News, the results are different. There is a slightly higher deletion rate preceding a stressed vowel than an unstressed one, the tokens with a following vowel in a monosyllabic word show similar deletion rates as the tokens followed by an unstressed vowel.

There is an effect of following stress on the duration when the /s/ occurs before a vowel, with a longer average duration for a following stressed vowel (101 msec) than a following unstressed vowel (89 msec). In normal speech, a word-final /s/ followed by a vowel would be resyllabified to be in the onset of the following syllable. This duration data can be compared to two studies of /s/ duration in the onset of stressed and unstressed syllables. First, my duration data correlates very well with Lavoie's (2000) finding that /s/ in Mexican Spanish is longer in the onset of word-medial syllable when the syllable is stressed (101 msec) than when it is unstressed (84 msec). However, in another study, Manrique and Massone (1981) report the opposite effect in the Spanish of Buenos Aires.

They find that /s/ in the onset of an unstressed syllable is longer (187 msec) than in a stressed syllable (148 msec).

With regard to the pre-consonant contexts, Table 29 and Table 30 show an effect of following stress type in word-final position but not word-internally. Both corpora have a lower rate of deletion of word-final /s/ before a stressed syllable than an unstressed syllable, and the deletion rate is higher before a monosyllabic word than either a stressed or an unstressed syllable. For the Broadcast News corpus, there is not much effect of following stress in the pre-consonant contexts on either duration or Spectral COG.

Context	Following Stress	N	% del
Word-Internal	Stressed	3,618	33.2
	Unstressed	2,845	31.2
Final, Pre-Consonant	Monosyllabic	3,043	63.0
	Stressed	3,240	52.3
	Unstressed	1,556	56.2
Final, Pre-Vowel	Monosyllabic	1,690	41.6
	Stressed	1,240	37.6
	Unstressed	1,038	41.0

Table 29. Percent deletion for the CallHome corpus by following stress.

Context	Following Stress	N	% deletion	/s/ duration	Spectral COG
Word-Internal	Stressed	5,430	18.5	75	4767
	Unstressed	8,730	21.3	71	4727
Final, Pre-Consonant	Monosyllabic	7,501	46.3	74	4659
	Stressed	6,012	31.5	73	4674
	Unstressed	7,981	35.5	69	4607
Final, Pre-Vowel	Monosyllabic	4,078	5.5	99	4925
	Stressed	1,855	7.2	101	4863
	Unstressed	5,529	5.2	89	4846

Table 30. Percent deletion and measurement data for the Broadcast News corpus by following stress.

Table 31 shows the percent deletion by the syllable stress type for the CallHome corpus and Table 32 shows the same data for the Broadcast News corpus. As with the following syllable stress type, I have included the value *monosyllabic* because it is not possible to determine if monosyllabic words are stressed or unstressed from the pronunciation lexicon. For both corpora, the factor values *stressed* and *pre-stressed* appear to pattern together with low deletion rates, and the *monosyllabic* and *post-tonic* have high deletion rates. This is different than what Brown and Cacoullos (2003) report; they find that word-final /s/ in a stressed syllable is more likely to be weakened. For the Broadcast News corpus, the /s/ duration is much longer in the post-tonic position even though the deletion rate is highest. A large contributor to the duration of post-tonic tokens is the fact that many are phrase-final. If the phrase-final tokens are removed, the average duration of post-tonic /s/ is only slightly longer than the other categories (83 msec).

Syllable Stress	N	% del
Monosyllabic	5,079	45.6
Stressed	4,127	36.4
Pre-Stress	4,251	33.1
Post Tonic	7,636	54.2

Table 31. Percent deletion by syllable stress type in the CallHome corpus.

Syllable Stress	N	% deletion	/s/ duration	Spectral COG
Monosyllabic	9,693	24.9	79	4714
Stressed	9,018	22.3	79	4756
Pre-Stress	8,390	18.9	74	4752
Post Tonic	27,393	25.2	89	4769

Table 32. Percent deletion and measurement data by syllable stress type in the Broadcast News corpus.

5.1.5 Grammatical Status

I have coded the grammatical status of words with several different levels of granularity. In the most general way, the grammatical status can be split into four categories: (1) *word-internal*: the syllable-final -/s/ is word-internal, and therefore has no grammatical role; (2) *lexical*: the -/s/ is word-final and part of the lexical item, not a morpheme; (3) *plural*: the word-final -/s/ makes up the plural morpheme (for any part of speech); (4) *verb*: the word-final -/s/ is part of a verbal ending. Both the *plural* and *verb* categories can be further split up. I have coded them separately in different variables, so that in the multivariate analysis, the finer-grained factor values will only be used if they improve on the statistical model. I have divided up the plural category into the sub-categories that have been identified above in Chapter 2: (1) first modifiers (articles *los, las, unos,* determiners *esas, esos, estas, estos,* quantifiers *todos, todas, menos, muchas,* etc.); (2) pronouns (*ellos, les, nos, vos, nosotros, ustedes*) (3) adjectives; and (4) nouns. There are three main verbal sub-categories: (1) *es;* (2) first person plural ending *–mos;* and (3) second person singular ending *–s*.

Table 33 shows the percent deletion by the grammatical category of the syllable-final /s/ for the CallHome corpus, with the categories *plural* and *verb* subdivided as described above. Table 34 shows the same data for the Broadcast News corpus. For both the CallHome corpus and the Broadcast News corpus, word-internal -/s/ has a lower deletion rate than any of the word-final categories. Among the word-final categories, *plural* /s/ (with all the sub-categories combined) overall has the most deletion. Lexical /s/

tends to favor deletion slightly more than the *verbal* /s/ category. This is in agreement with numerous studies that have found more deletion in plural markers than in word-final lexical /s/ (Poplack 1980a, 1980b, 1981; Hundley 1987; Cepeda 1995; Terrell 1977).

Among the plural sub-categories, in the CallHome corpus, all of the categories have very close deletion rates (ranging from 52.2% to 54.7%), except for *numbers*, which is much lower (40.0%). In the Broadcast News corpus, there is greater variation in the deletion rates - ranging from 21.4% for adjectives to 32.4% for pronouns. These findings are surprising, because other studies have found different results. Poplack (1980a, 1980b, 1981), Cepeda (1995) and Terrell (1977) all report low deletion rates among determiners in comparison to nouns and adjectives.

For the verb sub-categories, both corpora show the lowest deletion rate for *es*, followed by the second person singular form *–s,* and the first person plural form *–mos* with the most deletion. This is in agreement with Terrell's (1977) data; in his chart, he specifies that the verb *es* has low deletion, the second person singular form has moderate deletion, and that *–mos* has high deletion.

Terrell (1977) also claims that there are low deletion rates among numerical modifiers. This finding is supported by the data from the CallHome corpus but not from the Broadcast News corpus. A large contributor to the high deletion rate of word-final /s/ in numbers in the Broadcast News corpus is due to the distribution of the numbers used in that corpus. Of the tokens in the *number* sub-category, over one-third appear in either the morpheme *–cientos* (hundreds), or the words *miles* (thousands) or *millones* (millions), all of which have relatively high deletion rates.

Grammatical Category	N	% del
All Plural	**6136**	**52.0**
First Modifier	2,067	52.2
Noun	2,050	54.7
Adjective	685	53.1
Pronoun	670	53.1
Number	664	40.0
All Verbs	**5,059**	**47.8**
es	1759	42.8
-mos	822	52.4
2nd person sing.	2458	47.8
Lexical	**3,429**	**50.2**
Word-Internal	**6,464**	**32.3**

Table 33. Percent Deletion in the CallHome corpus by grammatical category.

Grammatical Category	N	% del	Ave. [s] duration (ms)	Ave. spectral COG
All Plural	**28,971**	**25.5**	**85**	**4735**
First Modifier	8,010	26.4	76	4665
Noun	10,888	25.9	89	4782
Adjective	6,783	21.4	90	4754
Pronoun	1,150	32.4	82	4635
Number	2,140	29.2	82	4747
All Verbs	**5,443**	**23.7**	**88**	**4817**
es	1,370	21.5	86	4910
-mos	1,214	29.9	79	4598
2nd person sing.	2,845	22.3	93	4854
Lexical	**4,683**	**24.7**	**93**	**4818**
Word-Internal	**14,170**	**20.2**	**72**	**4743**

Table 34. Percent deletion and measurement data in the Broadcast News corpus by grammatical category.

Another consideration related to grammatical category is whether a word is a function word or a content word. Phillips (2001) claims that content words and function words behave differently with respect to lenition processes because function words are less analyzed by the speaker into their constituent phonemes. Following Phillips (2001:128), I am using the term *function word* to indicate a word that receives low sentence stress. This includes: articles (*los, las, unos)*, determiners (*esas, esos, estas, estos*), quantifiers (*todos, todas, menos, muchas*, etc.), pronouns (*ellos, les, nos, vos, nosotros, ustedes*), numbers, the verb form *es,* and certain common lexical items such as *más, entonces, pues.*

In Table 35 and Table 36, I have split up the tokens of word-final /s/ into these two categories. For the CallHome corpus, content words show more deletion than function words, but for the Broadcast News corpus, function words show more deletion. As we will see in the next section, this is a result of an interaction with the number of syllables, and for a given number of syllables, function words have higher rates of deletion than content words in both corpora.

	N	% deletion
Content Words	7,307	51.1
Function Words	7,419	48.4

Table 35. Percent deletion of word-final /s/ in the CallHome corpus, separated by content words and function words.

	N	% deletion	/s/ duration	Spectral COG
Content Words	26,405	23.9	91	4779
Function Words	13,946	26.9	79	4718

Table 36. Percent deletion and measurement data of word-final /s/ in the Broadcast News corpus, separated by content words and function words.

5.1.6 Number of syllables

As described above in section 2.2.4, the number of syllables in a word has an effect on -/s/ weakening, with greater rates of deletion for words with more syllables. However, it is not clear whether the distinction is simply between monosyllabic words in comparison to polysyllabic words, or if there is also an effect of the number of syllables in the word (2 syllables vs. 3 syllables vs. 4 syllables). I have therefore coded two separate variables, one for just the distinction between monosyllabic and polysyllabic words, and another one for the actual number of syllables.

The data for the CallHome corpus separated by the number of syllables is shown in Table 37, and the corresponding data for the Broadcast News corpus is in Table 38. I have additionally separated the data by function words and content words because function words tend to have fewer syllables than content words. It is somewhat surprising that there are some tokens of function words in the category of 4 or more syllables. These are mostly numbers, such as *novecientos* ('nine hundred'), *ochocientos* ('eight hundred'), etc. For both of the corpora, the number of syllables has a clear effect on the deletion rate, both for content words and for function words. The primary distinction does appear to be the contrast between monosyllabic and polysyllabic words. There is an average of more than 10% greater deletion rate for words with 2 or more syllables in comparison with words that have only one syllable. The deletion rate among polysyllabic words is relatively constant, with the exception of content words in the Broadcast News corpus, where there is nearly 5% more deletion among words with 2 syllables than those with 3 syllables.

	Number of Syllables	N	% deletion
Content Words	1	612	36.9
	2	3,715	51.4
	3	2,060	53.6
	4+	824	54.6
Function Words	1	5,632	44.9
	2	971	60.1
	3	780	59.7
	4+	36	33.3

Table 37. Percent deletion by the number of syllables in the word and word class, for word-final tokens in the CallHome corpus.

	Number of Syllables	N	% deletion	/s/ duration	Spectral COG
Content Words	1	820	17.1	93	4749
	2	7,358	27.0	91	4774
	3	9,772	22.2	91	4759
	4+	8,429	23.9	90	4811
Function Words	1	10,851	24.4	80	4727
	2	1,649	34.3	81	4711
	3	1,007	36.6	79	4708
	4+	439	39.9	72	4520

Table 38. Percent deletion and measurement data by the number of syllables in the word, for word-final tokens in the Broadcast News corpus.

5.1.7 Speech rate

Although it has not been discussed in the Spanish /s/ lenition literature, speech rate has been shown to have a significant impact on other lenition processes (see for example Fosler-Lussier and Morgan 1999). The speech rate should have an impact on the duration of segments as well as the deletion rate. I have therefore calculated the speech rate for each of the tokens of syllable-final /s/. Following Fosler-Lussier and Morgan (1999), I have chosen to measure the speech rate in syllables per second. Other methods that Fosler-Lussier and Morgan consider for measuring speech rate are words per second or segments per second, but they find that syllables per second is the best measure. Different words can vary greatly in their duration. Different syllables also have slightly different intrinsic durations, but they are much more consistent than whole words. Individual segments have high deletion rates in comparison with whole syllables, so it is more problematic to calculate a speech rate based on a count of the segments because it is first necessary to determine how many segments were actually spoken.

My goal was to get a calculation of the local speech rate (the speech rate right at the time of the /s/), because this should have a greater effect on lenition than a more global speech rate. However, in order to ensure that small inaccuracies in the time alignments do not cause large effects on the calculated speech rate, it is necessary to average the speech rate across a number of syllables. For the CallHome corpus, the speaker turns are fairly long and often include more than one sentence and several pauses, which makes the speaker turn too long of a time period over which to average the speech rate. Instead, I calculated the speech rate using automatically generated approximate word-level time alignments. I used the time interval including up to the five words before and up to the five words after the word in question, for a total of up to 11 words. If a change in speaker turn or a long pause occurred within this 11-word window, the window was shortened to begin or end at the speaker turn or pause. The duration across the entire window was divided by the total number of syllables in the words spoken during that time.

The transcriptions that accompany the Broadcast News corpus have more frequent time alignments (the utterances are generally broken down by the transcriber into phrases of 30-50 syllables), and generally contain a time alignment for each pause in the speech. Because of this, I was able to calculate the speech rate across the breath group by dividing the time between alignments in the transcription file by the total number of syllables in that interval.

The data for the CallHome corpus is shown in Table 39 and the data for the Broadcast News is shown in Table 40. I have removed the phrase-final tokens because the speech rate calculated across the breath group may not be representative of the local speech rate at the end of a phrase. The average speech rate for the CallHome tokens is slightly higher than for the Broadcast News corpus (6.18 syllables/second compared to 5.93 syllables/second). As expected, both corpora demonstrate more deletion as the speech rate increases. The duration data for the Broadcast News corpus shows that the duration of the /s/ also is shorter as the speech rate increases, as would be expected. Somewhat surprisingly, the spectral COG is nearly constant across speech rates, indicating little change in articulation. The deletion data is in keeping with Fosler-

Lussier and Morgan's (1999) results for English telephone speech. They find that the overall segment deletion rate increases as the speech rate increases, from 9.3% for their slowest category to 13.6% for their fastest category.

Speech rate	N	% deletion
0<X<5	4,765	39.1
5<X<6	3,937	42.6
6<X<7	3,802	45.1
7<X	6,033	50.5

Table 39. Percent deletion by speech rate (syllables/second) for tokens of -/s/ that are not phrase-final the CallHome corpus.

Speech rate	N	% deletion	/s/ duration	Spectral COG
X<5	7,803	22.6	89	4711
5<X<6	15,285	21.6	82	4763
6<X<7	16,563	25.3	75	4766
X>7	7,525	32.4	69	4705

Table 40. Percent deletion and measurement data by speech rate (syllables/second) for tokens of -/s/ that are not phrase-final in the Broadcast News corpus.

5.1.8 Lexical identity

In addition to all of the extralinguistic and linguistic factors already listed, I have also included a factor for the lexical identity. Just as with the factor for the speaker identity, there are too many different words to use this factor in a statistical analysis of the entire corpus. There are over 6,000 different words with syllable- or word-final /s/ in the Broadcast News corpus and over 2,500 different words in the CallHome corpus.

Figure 19 shows the deletion proportion for each lexical item that occurs in the CallHome corpus, sorted by the deletion proportion, for words that have at least 20 tokens in that corpus.

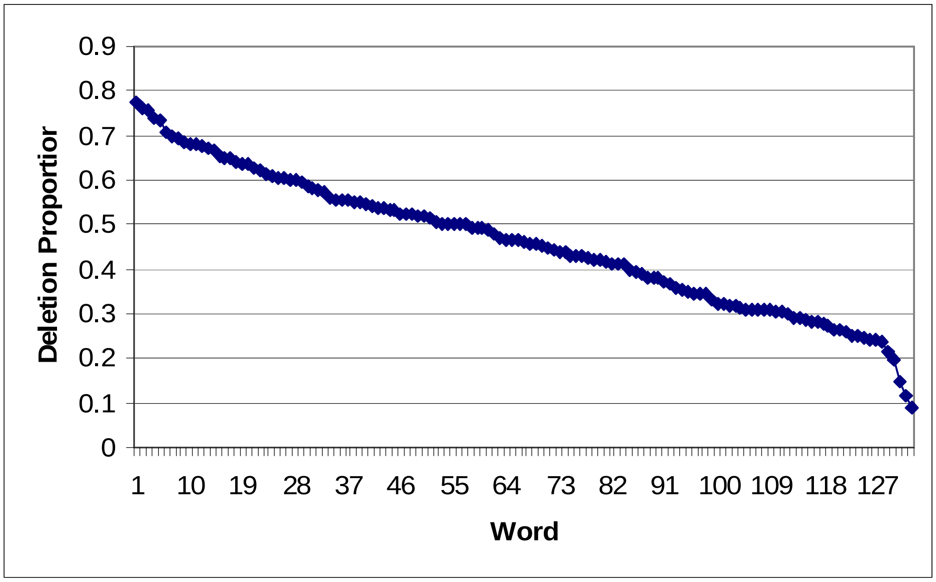

Figure 19. Deletion proportion by word in the CallHome corpus.

Figure 20 is the same data for the Broadcast News corpus. Both figures show that there is a wide range of deletion proportions. The Call Home corpus has more words that have high deletion. The Broadcast News corpus has 21 words that have more than 50% deletion: *cientos, decenas, buenas, regresemos, jefes, portavoz, todas, principios, través, buenos, oyentes, queremos, miembros, miles, informes, nosotros, centígrados, todos, antes, muchas,* and *toneladas.* Just as with the speaker deletion rates, for each token of syllable-final /s/, I have also coded the overall deletion rate of the word across the entire corpus. This variable can only be used on a subset of the data that includes only the words that are relatively frequent. For the words that appear only a handful of times, the calculated deletion rate may not be representative of that word's true deletion rate, particularly if all of the tokens of that word were spoken by the same speaker.

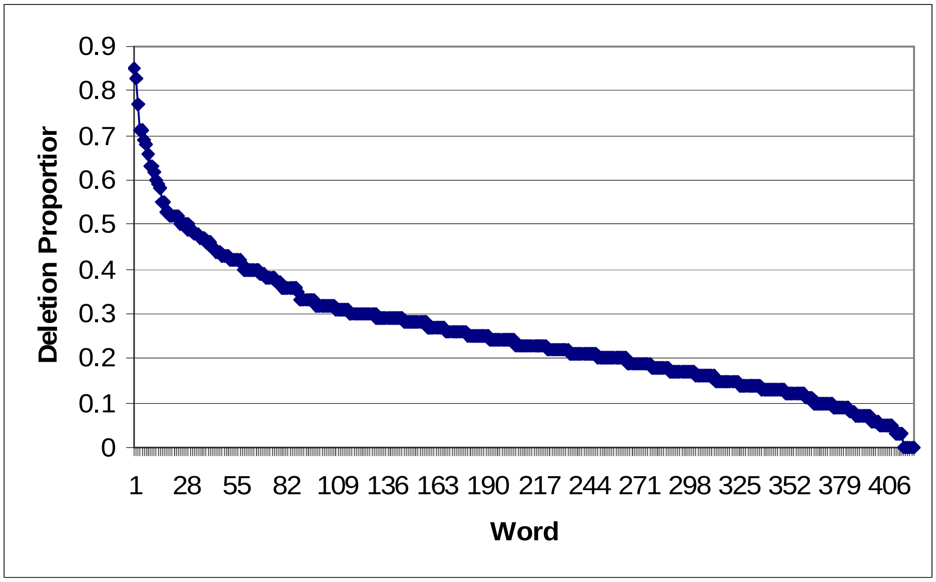

Figure 20. Deletion proportion by word in the Broadcast News corpus.

5.2 Usage-based Independent Variables

A number of usage-based variables have been presented in Chapter 3. The primary variable that has been investigated by other researchers is word frequency. I have coded this variable as well as several other usage-based variables: measures of word predictability (relative frequency of word sequences and conditional probabilities), measures of the speaking situation (repetition number of the word and local deletion rates), and two other usage-based measures that might affect the lexical representation of the word (degree of ambiguity and phonetic context proportion). Another measure of word predictability that has been claimed to affect pronunciations is the semantic relatedness to other words in the conversation, but this is beyond the scope of the present study.

5.2.1 Calculation of usage-based variables from the corpora

For overall word frequency, I was able to use the word count data from the LDC Spanish Lexicon described in section 4.1. The advantage of using the word frequency data from the pronunciation lexicon is that it contains word counts for corpora with very high word counts (a total of over 60 million words). With such high word counts, the word frequencies listed are likely to be representative of the actual word frequencies in the language. However, for the other usage-based variables that depend on how often and in what context words appear (relative frequencies, conditional probabilities, degree of ambiguity and phonetic context proportion), I had to calculate the variables myself from the transcripts of speech corpora. In order to increase the total word count to be used in these calculations, in addition to the transcripts from the two corpora that were used as the data sources for the present syllable-final /s/ study, I also included the transcripts from the LDC Hub5 Spanish Telephone Speech Corpus (CallFriend). The total word count from these three corpora is 1,067,626 (199,550 words from CallHome, 550,651 from CallFriend, and 317,425 from Broadcast News).

The CallFriend corpus consists of 106 unscripted telephone calls between native speakers of Spanish. The speech data was originally recorded to support an Automatic Language Recognition project, and the transcriptions were created to support the same LVCSR Spanish project as the CallHome corpus. The speech in this corpus was collected in a similar manner to the CallHome corpus; the speakers were able to call whomever they wanted within North America, Puerto Rico, and the Dominican Republic, and they were allowed to speak on any topic of their choosing. Each phone call lasted up to 30 minutes. As much of the phone call as possible was transcribed orthographically, although speech that was difficult to transcribe was skipped, resulting in somewhere between 10 and 30 minutes of speech being transcribed orthographically for each phone call. The resulting transcripts have a higher word count than either of the other corpora.

To make the calculations for the usage-based variables, I wrote a perl script to parse each line of each transcript and make a count of the number of times each word (unigram), each two-word (bigram) sequence, and each three-word (trigram) sequence occurs. From these counts, I then calculated the word frequencies, relative frequencies, conditional probabilities, degree of ambiguity, and proportion of phonetic contexts as described in the following sections.

5.2.2 Word frequency

For word frequency, I have coded five different measures, four based on the counts given in the Spanish pronunciation lexicon, and one calculated from the transcripts of the three corpora. As described above in section 4.1, the lexicon includes word counts for five corpora: a telephone speech corpus (CallHome), a radio broadcast corpus, and 3 different text corpora. If word frequency affects lenition due to the word being spoken more, the effect should be based on the word counts for the speech corpora as opposed to written text corpora. However, the total word count of the text corpora is so much larger than that of the speech corpora—and particularly the informal speech corpus (CallHome). However, the speech corpora may not have a sufficient total word count to make the individual word counts representative of the actual word frequencies for the whole language. This problem is extreme for the CallHome corpus (the word counts for this corpus in the pronunciation lexicon are based on less than 150,000 words), and many words from the Broadcast News do not appear in the CallHome corpus. Because the text corpora may in fact better represent the word frequencies for the less frequent words, I have coded a separate variable for the word frequency for each type of speech separately: (1) informal speech; (2) radio speech; (3) text. In addition, I have calculated the average of these three frequencies. Each of these word frequency variables is measured as the number of times the word appears per 1,000,000 words.

In addition to the word frequencies given in the pronunciation lexicon, I have also calculated word frequencies from transcripts of the CallHome, Broadcast News and CallFriend corpora. I first calculated the probability of each word:

$$p(w_i) = \frac{C(w_i)}{\sum_j (w_j)} = \frac{C(w_i)}{N}$$

where $C(w_i)$ is the count of how many times the word w_i appears and N is the total number of words in the three corpora.

I then multiplied this probability by 1,000,000 to normalize it to the frequency per million words. Converting the measurement from word probability to the frequency count per million words is for convenience; it allows the numbers to be more easily compared to data on word frequency by other researchers, and also makes the numbers more readable since so many of the tokens have word frequency less than 100 (word probabilities less than .0001). In this section, I will report results only based on the average word frequency from the word counts in the pronunciation lexicon, but for the analysis in Chapter 6, there are 5 separate variables of word frequency for investigation.

Although word frequency is normally discussed in terms of the number of times a word appears per million words, word frequency is not linearly distributed. There are many words that have relatively low frequencies, and few words that have extremely high frequencies. Table 41 shows the distribution of unique words as well as the total number of tokens, separated by the cutoffs that are powers of 10: 10, 100, 1,000, and 10,000 for the Broadcast News corpus. Although there are many more unique words in the lowest frequency category, there are approximately equal numbers of total tokens in the different categories. In order to capture the fact that the tokens are more evenly distributed by the power of 10 of the frequency, the appropriate independent variables for the statistical analysis are based on the logarithm of the word frequency.

Frequency	Unique Words	Total Tokens
X<10	4,555	10,786
10<X<100	1,422	15,870
100<X<1,000	176	13,703
1,000<X<10,000	20	8,858
X>10,000	2	5,317

Table 41. Distribution of unique words and total number of tokens by word frequency for the Broadcast News corpus.

The data by word frequency is shown in Table 42 for the CallHome corpus and in Table 43 for the Broadcast News corpus. I have separated the data into the categories Word-Internal, Function Words, and Content Words because they behave differently, and also they are not equally distributed among the frequency categories. Not surprisingly, for each of the corpora, the Content Words have a lower proportion of tokens in the highest frequency category (more than 1000 words/million) than the Function Words do.[2]

For word-internal tokens, there appears to be a frequency effect in both corpora. This effect is stronger in the Broadcast News corpus, with nearly twice as much deletion in the highest frequency category in comparison with the lowest frequency category. There is also an effect on duration in word-internal tokens; the highest frequency words have an /s/ that is on average 11 msec shorter than in the lowest frequency words.

The effect of frequency on word-final /s/ is not as clear. For function words, there are lower deletion rates in the highest frequency category (greater than 1000 tokens/million words) than in the next highest frequency category (between 100 and 1000 tokens/million words). This may be due to interactions with other variables such as the number of syllables in the word, with the more frequent words also tending to have fewer syllables. For content words, the two middle frequency categories contain the majority of the tokens. In both corpora, the higher of these two frequency categories has a slightly lower deletion rate than the lower of the frequency categories. This is also puzzling, and may be due to interactions with other variables that the statistical analysis can account for.

---

[2] In fact, only 22 words with a syllable-final /s/ occur with a frequency greater than 1000 words/million: *años, desde, después, dos, entonces, es, esta, está, están, este, esto, estoy, hasta, las, los, más, nos, pues, sus, todos, tres,* and *vamos*. Of these, nearly all are either word-internal or occur in monosyllabic words. The only exceptions are ***años, después*** (which has both a word-internal and a word-final /s/)*, entonces, todos,* and *vamos.*

	Frequency	N	% deletion
Word-Internal	X<10	667	30.0
	10<X<100	1,355	31.5
	100<X<1000	1,753	31.9
	X>1000	2,764	33.7
Content Word	X<10	1,326	50.5
	10<X<100	2,257	53.2
	100<X<1000	3,367	49.2
	X>1000	357	58.3
Function Word	X<10	60	41.7
	10<X<100	180	50.0
	100<X<1000	1,262	54.0
	X>1000	5,917	47.2

Table 42. Deletion rate for the CallHome corpus by word frequency.

	Frequency	N	% deletion	/s/ duration	Spectral COG
Word-Internal	X<10	3,613	14.7	77	4769
	10<X<100	4,243	17.4	74	4745
	100<X<1000	3,437	24.1	70	4712
	X>1000	2,890	26.5	66	4744
Content Word	X<10	7,064	21.5	93	4812
	10<X<100	10,621	25.3	90	4778
	100<X<1000	8,016	23.4	89	4761
	X>1000	704	34.4	87	4645
Function Word	X<10	113	28.3	84	4773
	10<X<100	1,002	34.1	77	4627
	100<X<1000	2,250	31.8	80	4726
	X>1000	10,581	25.2	80	4724

Table 43. Deletion rate and measurement data for the Broadcast News corpus by word frequency.

5.2.3 Relative frequencies

In addition to the frequency of individual words (***unigrams***), I have also coded the relative frequencies of ***bigrams*** (two-word sequences) and ***trigrams*** (three-word sequences). Because the LDC pronunciation lexicon has no bigram or trigram information, the only bigram and trigram relative frequency information I have is what I calculated from the three corpora. Just as with the relative frequencies of individual words, I first calculated the probability of each bigram or trigram, and then multiplied this probability by 1,000,000 to normalize it to the frequency per million bigrams or trigrams.

I have calculated two bigram relative frequencies. The first is based on the bigram probability with the preceding word. The probability of the word sequence $w_{i-1}w_i$, where w_i is the word with the syllable-final /s/ under investigation and w_{i-1} is the immediately preceding word is given by:

$$p(w_{i-1}w_i) = \frac{C(w_{i-1}w_i)}{N}$$

where C($(w_{i-1}w_i)$ is the count of the bigram $w_{i-1}w_i$ and N is the total number of bigrams. The second bigram relative frequency is based on the bigram probability with the following word:

$$p(w_i w_{i+1}) = \frac{C(w_i w_{i+1})}{N}$$

I have also calculated two trigram relative frequencies. The first is based on the trigram probability with the two preceding words:

$$p(w_{i-2}w_{i-1}w_i) = \frac{C(w_{i-2}w_{i-1}w_i)}{N}$$

The second is the trigram probability with the immediately preceding and the immediately following word:

$$p(w_{i-1}w_i w_{i+1}) = \frac{C(w_{i-1}w_i w_{i+1})}{N}$$

I have included in Table 44 and Table 45 the deletion and measurement data by the relative frequency of the bigram with the following word. The other bigram and trigram relative frequencies will not be addressed in this section, although they will all be considered in the statistical model in Chapter 6. The effect of the frequency of co-occurrence with the following word should be the greatest among the bigram and trigram frequencies because a word-final /s/ actually comes in contact with the following word. In order to show the data in a table, I have divided the data into Low, Medium and High categories of relative frequency. The cutoffs between the categories were chosen in order to have approximately equal numbers of tokens in each category.

The results in Table 44 and Table 45 are conflicting. For the Broadcast News corpus, there appears to be an effect of the bigram relative frequency – for both word-internal and word-final tokens – with the higher frequency bigrams having higher deletion rates, shorter /s/ duration and lower Spectral COG. However, for the Call Home corpus, there is little effect of frequency, and for the word-final tokens, the highest bigram relative frequency category has the lowest deletion rate. This does not necessarily mean that no effect exists; there may be interactions with other variables.

	Relative Frequency	N	% deletion
Word-Internal	Low	2,521	31.1
	Medium	2,324	34.2
	High	1,694	31.8
Word-Final	Low	4,999	50.4
	Medium	4,571	51.2
	High	5,156	47.8

Table 44. Deletion rate by bigram relative frequency with the following word ($w_i w_{i+1}$) for the CallHome corpus.

	Relative Frequency	N	% deletion	/s/ duration	Spectral COG
Word-Internal	Low	5,209	16.9	74	4820
	Medium	5,602	21.0	73	4702
	High	3,372	23.8	69	4687
Word-Final	Low	11,438	20.9	88	4825
	Medium	15,745	25.3	87	4737
	High	13,168	28.1	86	4722

Table 45. Data by bigram relative frequency with the following word ($w_i w_{i+1}$) for the Broadcast News corpus.

5.2.4 Conditional probability

I have coded several conditional probabilities for each token syllable-final /s/. While the relative frequency measurements indicate how likely a word or group of words is in a global sense, the conditional probabilities measure how likely the word in question is given another word or words that co-occur. As with the relative frequencies, I have coded conditional probabilities with respect to two sizes of word sequences, bigrams and trigrams. There are three conditional bigram probabilities that I have calculated. The first is the *forward conditional bigram probability of the word*, which measures the likelihood of the word in question given the preceding word. It is calculated by the number of times both words appear together divided by the number of times the preceding word occurs.

$$p(w_i|w_{i-1}) = \frac{C(w_{i-1}w_i)}{C(w_{i-1})}$$

The second conditional bigram probability is the *forward conditional bigram probability of the following word*, the likelihood of the following word given the current word:

$$p(w_{i+1}|w_i) = \frac{C(w_i w_{i+1})}{C(w_i)}$$

The third conditional bigram probability is the *reverse conditional bigram probability of the word*, the probability of the word in question, given the following word:

$$p(w_i|w_{i+1}) = \frac{C(w_i w_{i+1})}{C(w_{i+1})}$$

I have also coded a fourth measure of bigram probability called the *mutual information* of the word in question with the following word, which is a measure of how likely the two words are to occur together in comparison to how likely each of the words are individually:

$$m.i. = \frac{C(w_i w_{i+1})}{C(w_i)C(w_{i+1})}$$

Unlike the conditional probabilities, the mutual information is high only when the first word is likely given the second word and the second word is likely given the first word. According to Gregory ***et al***. (1999), mutual information is a good measure because a larger value for mutual information indicates stronger cohesion between the two words.

I have calculated two trigram probabilities of the word with a syllable final /s/. The first is the ***forward conditional trigram probability of the word***, the probability of the word given the two previous words, $p(w_i|w_{i-2}w_{i-1})$. The second is the ***centered conditional trigram***, the probability of the word given the one preceding and one following word, $p(w_i|w_{i-1}w_{i+1})$.

Table 46 and Table 47 show the data for one of the conditional probability measures, the forward conditional bigram probability (the probability of the following word given the current word). Only tokens that are word-final but not utterance final are included. The tokens are split into two groups based on the relative frequency, since low frequency words will have on average higher conditional probabilities. the reason low frequency words have higher conditional probabilities is due to the fact that the data I have used to calculate the conditional probabilities is not representative of the shole language due to the low word count used, and the fact that the data I am evaluating is the same that I used to calculate the condidtional probabilities (the CallHome, Broadcast News, and CallFriend corpora). For example, ***recompensados*** appears only once, it appears with the following word ***por***, with probability 1the highest it can be. As with the relative frequency data, the criteria for the categories were selected in an attempt to have roughly equal numbers of tokens in the high frequency and low frequency categories, and in the three conditional probability categories. Because the distribution of tokens by word frequency is different for the two corpora, I have different criteria for the categories. For the CallHome corpus, I used a cutoff between high frequency and low frequency words of 1,000, and for Broadcast News, I used 100. For the conditional probabilities, I also selected the cutoffs to get categories with roughly the same number of tokens.

The data in Table 46 and 47 shows that deletion rates increase as the forward conditional bigram probability increases for both corpora, and for both high frequency and low frequency words. This effect is stronger in the Broadcast News corpus, where the deletion rate is nearly 10% higher for the high conditional probability category compared to the medium conditional probability category (32.0% vs. 23.3% for low frequency words and 27.6% vs. 37.5% for high frequency words). There is also an effect on Spectral COG, with words with a higher conditional probability having a lower Spectral COG. Interestingly, there is little effect of conditional probability on /s/ duration.

Word frequency	Conditional Probability	N	% deletion
Low	Low	1,542	52.9
	Medium	2,381	53.7
	High	2,474	54.8
High	Low	3,094	47.5
	Medium	2,046	49.5
	High	469	53.7

Table 46. Deletion rate by word frequency and forward conditional bigram probability for the CallHome corpus.

Word frequency	Conditional Probability	N	% deletion	/s/ duration	Spectral COG
Low	Low	960	21.0	82	4844
	Medium	7,101	23.3	85	4793
	High	5,824	32.0	84	4762
High	Low	11,777	24.1	80	4719
	Medium	4,765	27.6	82	4763
	High	2,566	37.5	78	4642

Table 47. Deletion rate and measurement data by word frequency and forward conditional bigram probability for the Broadcast News corpus.

5.2.5 Repetition

I have coded the repetition number for a given speaker in a given speech file for each word containing a syllable-final /s/. While the number of times a words has already been spoken is not exactly a word predictability measure as the relative frequencies and conditional probabilities are, it is related to word predictability because a word that has already been used in a given discourse is more likely than one that has not been previously used. If, as we have seen in the previous sections, words or word sequences that are more predictable have higher rates of -/s/ lenition, then it would make sense that tokens of -/s/ that appear in a second or subsequent repetition should have higher deletion rates than tokens of -/s/ in words that appear the first time.

The data in Table 48 and Table 49 show the percent deletion and measurement data by the repetition number for the CallHome and Broadcast News corpora. I have separated word-internal, content word and function word tokens of -/s/ because I hypothesized that the effect of this variable should be stronger among content words than function words. However, there is no clear pattern of the effect of repetition in this data. For the CallHome corpus, the deletion rates vary for each type of word without a clear trend. For the Broadcast News corpus, there appears to be a trend of lower deletion rate with increasing repetition for both types of word-final /s/ (content words and function words), but the opposite trend for word-internal /s/. Although not shown in this chart, dividing the word-final /s/ further into lexical, plural, and verbal /s/ all show the same trend of higher deletion rates for the first repetition when compared to later repetitions.

	Repetition Number	N	% deletion
Word-Internal	1	4,556	32.4
	2	1,075	34.7
	3	393	30.3
	4+	510	29.6
Content Word	1	5,981	50.9
	2	977	51.7
	3	232	50.9
	4+	117	59.8
Function Word	1	3,962	50.0
	2	1,483	46.5
	3	716	44.8
	4+	1,245	47.7

Table 48. Percent deletion by repetition number for the CallHome corpus.

	Repetition Number	N	% deletion	/s/ duration	Spectral COG
Word-Internal	1	12,481	19.8	73	4738
	2	1,277	22.9	70	4795
	3	284	22.5	72	4788
	4+	141	24.0	71	4680
Content Word	1	23,259	24.4	91	4778
	2	2,304	21.1	91	4797
	3	521	19.4	90	4775
	4+	321	15.0	96	4773
Function Word	1	8,423	29.5	80	4726
	2	2,460	24.8	79	4705
	3	1,164	23.7	80	4733
	4+	1,863	20.6	76	4689

Table 49. Percent deletion and measurement data by repetition number for the Broadcast News corpus.

5.2.6 Local deletion rate

I have calculated two types of local deletion rates, one within the NP (for plural /s/), and one that includes tokens of syllable-final /s/ that are further away. The local deletion rate can be considered a usage-based variable because it is consistent with an exemplar model that includes an auditory buffer (Johnson 1997:151). Since the buffer retains a record of the acoustic parameters over a certain length of time, more recently accessed exemplars are more salient in memory. When a speaker wants to say a word with syllable-final /s/, the next token of syllable-final /s/ will be disproportionately affected by the pronunciation in more recently accessed exemplars.

The variable for deletion within the NP measures whether plural /s/ was deleted or retained in the preceding words within the NP. I have coded two variables for deletion within the NP, one with attribute data indicating exactly what sequence of retention or deletion occurred (for example *sso*, indicating the first two elements retained the plural /s/ and the third deleted it). The other variable I have coded includes just the proportion of deletion in the tokens preceding. Using the proportion of deletion rather than the exact sequence has the advantage that it creates a numerical value, and also it avoids the large number of attribute values (*sso*, *sos*, and *oss* would all have values of 0.33). Because the deletion within the NP is only meaningful for a portion of the tokens, it cannot be used in an analysis of all the data combined. I have therefore coded the local deletion rate for all tokens, based on up to the last five tokens of the speaker.

The breakdown of this variable is shown in Table 50 for the CallHome corpus and Table 51 for the Broadcast News corpus. Both corpora show much higher deletion rates for tokens following other tokens of /s/ that have been deleted than following tokens of /s/ that have been retained, and the measurement data form the Broadcast News corpus shows shorter durations and lower Spectral COG for higher local deletion rates. This effect is not merely an effect of 'high deletion' speakers in comparison with 'low deletion' speakers, as is demonstrated in Table 52. Table 52 shows the data for just one speaker, Jorge Ramos, who has the most tokens in the Broadcast News corpus. While the magnitude of the effect is lower than the effect in the data across all speakers, he also has a higher deletion, shorter duration and lower spectral COG when the local deletion rate is higher.

Local deletion rate	N	% deletion
0.0<X<0.2	3,529	21.8
0.2<X<0.4	4,680	29.6
0.4<X<0.6	4,185	41.0
0.6<X<0.8	3,784	54.8
0.8<X<1.0	4,876	69.9

Table 50. Deletion rate as a function of the local deletion rate for the CallHome corpus.

Local deletion rate	N	% deletion	/s/ duration	Spectral COG
0.0<X<0.2	17,959	14.5	86	4915
0.2<X<0.4	16,519	19.8	82	4796
0.4<X<0.6	10,233	28.1	81	4614
0.6<X<0.8	5,130	38.1	79	4404
0.8<X<1.0	3,391	54.9	78	4229

Table 51. Deletion rate as a function of the local deletion rate for the Broadcast News corpus.

Local deletion rate	N	% deletion	/s/ duration	Spectral COG
0.0<X<0.2	1,566	21.7	72	4775
0.2<X<0.4	2,134	21.7	70	4772
0.4<X<0.6	1,373	25.5	69	4690
0.6<X<0.8	570	29.6	68	4633
0.8<X<1.0	184	29.3	67	4629

Table 52. Deletion rate as a function of the local deletion rate for the speaker Jorge Ramos in the Broadcast News corpus.

5.2.7 ***Degree of ambiguity of the word***

I have coded two independent variables that measure the degree of ambiguity of the word for plural /s/. The first of these is whether or not the singular and the plural would be ambiguous if the word-final /s/ is deleted. As an example, for the articles *los* and *las*, deletion of the /s/ in *los* does not lead to ambiguity because the singular is *el*, while deletion of the /s/ in *las* does lead to ambiguity between *las* and *la*. According to Wright (1997), the pronunciation of individual words can be affected by whether there are other words with similar pronunciations (***similarity-neighborhood density***); when words are harder to identify due to them being in a denser neighborhood, they are more carefully pronounced and are less likely to be reduced. In the case of plural /s/, according to this theory, if deleting a plural /s/ would lead to ambiguity, there should be pressure to retain the /s/. The data for this variable is shown in Table 53 and Table 54. The results are conflicting; in the Broadcast News corpus, the results are as predicted, with slightly lower deletion among plurals that cause ambiguity than those that do not. However, in the CallHome corpus, the deletion rate is higher among plurals that cause ambiguity than those that do not.

Status of deletion of /s/	N	% deletion
Ambiguous	3,339	53.9
Unambiguous	2,736	50.0

Table 53. Deletion data for plural /s/, based on whether deletion of /s/ would lead to ambiguity for the CallHome corpus.

Status of deletion of /s/	N	% deletion	/s/ duration	Spectral COG
Ambiguous	17,281	24.2	86	4735
Unambiguous	11,563	27.4	84	4733

Table 54. Deletion and measurement data for plural /s/, based on whether deletion of /s/ would lead to ambiguity for the Broadcast News corpus.

The second variable that measures the degree of ambiguity is how frequently a word appears in its singular form as compared to the plural form, expressed as the singular-to-plural ratio. Even if a plural form would be technically ambiguous with the singular form in the case of a deleted /s/, it is possible that such ambiguity is minimal if the word is overwhelmingly used in the plural. The data for this variable is shown in Table 55 and Table 56. There is no clear pattern in this data for either of the corpora. Thus, neither of the variables looking at ambiguity caused by deletion of /s/ appear to contribute either to deletion or lenition of /s/. In Chapter 6, these variables will be evaluated for plural tokens of /s/.

Singular-to-Plural ratio	N	% deletion
X<0.5	1,952	49.5
0.5<X<1	1,273	50.1
1<X<2	549	57.7
X>2	2,301	54.1

Table 55. Deletion data for plural /s/ by the word singular-to-plural ratio for the CallHome corpus.

Singular-to-Plural ratio	N	% deletion	/s/ duration	Spectral COG
X<0.5	11,069	25.2	89	4780
0.5<X<1	6,741	23.8	82	4654
1<X<2	2,504	25.8	89	4785
X>2	8,530	27.3	83	4725

Table 56. Deletion and measurement data for plural /s/ by the word singular-to-plural ratio for the Broadcast News corpus.

5.2.8 Phonetic context proportion for individual words

For each token of syllable-final /s/, I have coded the proportion that the /s/ in that particular word occurs in each of the phonetic contexts (pre-consonant, pre-vowel, pre-pause) in the three corpora used to generate statistics. So, for example, across all three corpora, the word *los* occurs before a consonant 65% of the time, making the value of the factor proportion consonant 0.65 for all tokens of *los*. In comparison, the word *las* occurs before a consonant 80% of the time. If, as Bybee (2001a) claims, usage patterns shape the lexical representation of the word and fine grained phonetic detail is stored in the lexicon, the phonetic context that words most frequently appear in should have an effect. Words that appear primarily in pre-consonantal position (ending in the context that most favors deletion) should have higher deletion rates than words that appear more often before a vowel, if all the other factors are equal. This means that if all other factors were equal, we would expect *las* to have slightly more advanced lenition than *los* because it is in the favoring context more often than *los.*

Because the effect of the phonetic context is hypothesized to operate on the underlying representation of the word, it should be seen in each phonetic context, consonant, vowel, pause; when a word whose lexical representation has been altered due to it being mostly followed by a consonant, that effect should also be seen when it is followed by a vowel. The data in Table 57 and Table 58 is somewhat difficult to interpret. These tables show the word-final data for each following segment type (consonant, vowel, pause) separated into two categories, "High" word consonant context and "Low" word consonant context, using a cutoff of 0.6 between the two categories. This cutoff was chosen because the median word consonant context for the Broadcast News is 60% and for CallHome it is 67%.

For the Broadcast News Corpus, it appears that there may be a small effect of the word being primarily followed by a consonant, in the pre-consonant and pre-pause positions. In the pre-vowel positions for the CallHome corpus, words with high consonant contexts actually have a lower average deletion rate than one with low consonant contexts.

Following segment type	Word consonant context	N	% deletion
Consonant	< 0.6 (Low)	3,018	57.6
	> 0.6 (High)	4,821	57.0
Vowel	< 0.6 (Low)	2,663	41.5
	> 0.6 (High)	1,305	37.5
Pause	< 0.6 (Low)	2,197	44.1
	> 0.6 (High)	722	37.8

Table 57. Percent deletion by following segment type and how often the word appears in pre-consonant position, for the CallHome corpus.

Following segment type	Word consonant context	N	% deletion	/s/ duration	Spectral COG
Consonant	< 0.6 (Low)	7,171	36.2	75	4706
	> 0.6 (High)	14,323	39.0	70	4611
Vowel	< 0.6 (Low)	7,234	5.8	96	4888
	> 0.6 (High)	4,228	5.4	91	4858
Pause	< 0.6 (Low)	6,199	16.6	108	4801
	> 0.6 (High)	1,196	17.9	108	4795

Table 58. Percent deletion by following segment type and how often the word appears in pre-consonant position, for the Broadcast News corpus.

6 *Data Analysis*

The data in the preceding chapter demonstrates that there are a number of independent factors that contribute both to the categorical deletion of syllable-final /s/ and the gradient lenition as measured using /s/-duration and the spectral COG. In the previous chapter, the data has been presented for individual independent variables one at a time (or in some cases, two independent variables). A more rigorous statistical analysis is needed in order to evaluate which of these factors provide a statistically significant contribution to syllable-final /s/ lenition, measure the relative contributions of the factors, and determine if there are any interactions between the independent variables.

I will first present statistical models for deletion, duration, and spectral COG for all of the syllable-final /s/ data from the Broadcast News corpus together. I will then separate the data into some of the major subsets of data that may behave differently from each other: content words vs. function words, pre-consonant vs. pre-vowel, plural /s/, and the verbal ending -*mos*.

Performing the same statistical analyses first without including any of the usage-based variables, and then adding the usage-based variables, allows an assessment of whether the usage-based variables improve the overall fit of the model.

6.1 *Model for all tokens of syllable-final /s/*

6.1.1 *Categorical deletion data*

I have performed a series of analyses of the data using Minitab 14, a statistical software package that has the capability of sophisticated data analysis and numerous statistical tests. For the categorical deletion data, the appropriate type of model is the binary logistic regression (Agresti 1990, Christensen, 1990). The binary logistic regression is used when the dependent variable has only two possible values and at least one of the independent variables contains discrete attribute data (data that has several different values and is non-numeric). The binary logistic regression can model both continuous and discrete independent variables.

To determine the optimal *base* model (the model that does not include any usage-based variables), I began with the binary logistic regression model that included all of the independent variables non-usage-based, and then determined which ones (if any) do not contribute significantly to the model. I then removed the variables one at a time that didn't contribute to the model. In some cases, I also combined factor values for independent variables with attribute values. Using a method similar to Amastae (1989), I was careful not to combine factor values simply because they intuitively belonged together, but also checked that the combination is supported by the data as having similar factor weights and not being statistically significantly different in their contributions to the model.

The output of the binary logistic regression for all tokens of syllable-final /s/ in the Broadcast News corpus is shown in Table 59. The output of a binary logistic regression provides two main numbers for each independent variable or factor value: the ***p-value*** and the *Odds Ratio*. The p-value for a given independent variable or factor value is a measure of whether that variable is a statistically significant contributor to the dependent variable. A low p-value indicates that the variable contributes significantly to the model at some confidence level (usually $p < 0.05$ or $p<0.01$, indicating that there is less than a 5% or 1% chance that the variable does not contribute to the model).

The Odds Ratio is a measurement of the strength of the independent variable or factor value. For numerical independent variables, the Odds Ratio indicates how much more likely the event (deletion of /s/) is to occur for a unit increase in the independent variable. For independent variables containing discrete factor values, there is a separate Odds Ratio for each of the factor values (with the exception of the first factor value, which serves as the reference), and it indicates how much more likely the event is to occur for that factor value in comparison to the reference. So, for example, in Table 59, for the base model, if the speaker's dialect is "interior", deletion is only 37% as likely as if the speaker's dialect is "coastal". An Odds Ratio close to 1 indicates that there is little effect of that factor on the probability of deletion occurring. Odds Ratios less than 1 indicate a lower probability of deletion, and Odds Ratios above 1 indicate a higher probability of deletion.

The statistic used to assess the validity of the whole model is the *Pearson Goodness-of-fit*, which measures how well the model fits the data. Lower p-values indicate that the model does not fit the data well. In general, a p-value less than the confidence level (less than 0.05 or 0.01) indicates that there is sufficient evidence that the model does not fit the data adequately.

From the base model in Table 59, we can see that all of the independent variables investigated contribute significantly to syllable-final /s/ deletion. A speaker from the interior ***dialect*** is only about one-third as likely to delete a syllable-final /s/ as a coastal speaker, and a ***male*** speaker is slightly more likely to delete than a female. These results are in keeping with what is expected from the literature. *Speech rate* also affects deletion in the expected way; for every syllable per second faster, the model predicts 17% more deletion. *Following segment* is divided into several different independent variables. For the computation of the binary logistic regression, independent variables cannot have colinearities; no independent variables containing attribute data may have a factor value that implies a factor value of another independent variable. Because a following segment type of *pause* implies a following segment place of articulation of *none*, it is not possible to perform the binary logistic regression using separate independent variables for both following segment type and following segment place of articulation. I therefore split the variable *following segment place of articulation* into a series of *indicator variables*, one for each possible factor value (*labial, coronal, velar*). Each indicator variable is set to 1 if the independent variable has that value and 0 otherwise. I also included a separate indicator variable for a following [d] because it behaves differently than would be predicted by it being a coronal voiced stop; even after taking into consideration that it is a voiced stop and not labial or velar, it is still 1.5 times more likely to cause an /s/ to delete. As the data in Table 59 indicates, a following voiced stop favors deletion; deletion is more than twice as probable when a voiced stop follows as compared to a voiceless stop. A following vowel is the least favoring environment (only 12% as likely to cause an /s/ to delete as compared to a following stop).

For *grammatical status*, verbal and plural /s/ are more likely to delete than lexical or word-internal /s/. For this base model, with all the tokens of syllable-final /s/ included, separating the plural and verbal categories was not significant. In addition to the grammatical status (*lexical/verbal/plural*), a separate variable for function words is also significant, with function words having higher deletion rates than content words. The two independent variables related to the number of syllables show that the important criterion is monosyllabic words in comparison to words with more than one syllable. *Monosyllabic words* are about half as likely to delete an /s/, but the addition of each subsequent syllable makes it slightly less likely that the /s/ will be deleted. The results of *stress* are as expected—an /s/ preceding an unstressed syllable is more likely to be deleted, and an /s/ that is in the coda of a stressed syllable is less likely to be deleted.

However, one must be cautious when using the results of the base model because the Pearson Goodness-of-fit statistic (p = 0.000) indicates that the model does not fit the data well at all.

The second group of columns shows the results of the binary logistic regression with the usage-based independent variables that contribute significantly to the model. As is shown in the differences between the base model and the model including usage-based variables, the p-values and Odds Ratios of independent variables and factor values can change, depending on which independent variables are included in the analysis (as well as how the factor is divided into different factor values). The model including the usage-based variables has a Pearson Goodness-of-fit statistic p = 0.121, indicating that this model has much better fit than the base model.

While nearly every one of the usage-based independent variables is statistically significant when evaluated individually (with respect to the base model), only five of the variables contribute significantly when taken together: *word frequency* as measured by the average across the sources in the LDC pronunciation lexicon, *the bigram relative frequency with the following word*, the *consonant proportion* for individual words, the *repetition number*, and the *local deletion rate*. The word frequency and bigram relative frequency both show the expected tendency; when a word or bigram is more frequent, the model predicts a higher deletion rate. The consonant proportion also behaves in the way predicted by a usage-based account; words that appear more frequently in pre-consonant context are predicted to have higher rates of deletion, even after taking into account the context that the token actually is in. The effect of repetition is surprising—it contributes significantly to the model, but it predicts that words that have been repeated more times are less likely to have a deleted [s]. The effect of the local deletion rate is very strong, with an Odds Ratio of 12.65. Adding the local deletion rate also lessens the effect of dialect, sex, and speech rate somewhat.

	Base		With Usage Variables	
Predictor	**p**	**Odds Ratio**	**p**	**Odds Ratio**
Dialect (Ref: 'coastal')				
interior	0.000	0.37	0.000	0.50
Male	0.000	1.18	0.000	1.12
Speech Rate	0.000	1.17	0.000	1.13
Preceding Vowel (Ref: 'a/e')				
i/w/y	0.000	0.76	0.000	0.80
o	0.000	1.20	0.000	1.22
u	0.000	1.34	0.000	1.47
Following Segment Type (Ref: 'Voiceless Stop')				
Glide	0.007	0.56	0.009	0.55
Liquid	0.000	1.64	0.000	1.71
Nasal	0.000	1.58	0.000	1.67
Pause	0.000	0.67	0.000	0.65
Voiced stop	0.000	2.58	0.000	3.07
Voiceless fricative	0.036	1.19	0.001	1.32
Vowel	0.000	0.12	0.000	0.10
Following [d]	0.000	1.50	0.000	1.39
Following labial	0.000	0.71	0.000	0.72
Following velar	0.000	0.51	0.000	0.48
Grammatical Status (Ref: 'Lexical')				
Internal	0.000	0.67	0.000	0.71
Plural	0.071	1.09	0.031	1.11
Verb	0.000	1.26	0.011	1.16
FunctionWord	0.000	1.46	0.000	1.26
SyllableCount	0.000	0.93	0.718	1.01
Monosyllabic	0.000	0.52	0.000	0.65
Following Stress (Ref: 'stressed/phrase final')				
unstressed/ monosyllabic	0.000	1.32	0.000	1.39
Stressed syllable	0.000	0.77	0.000	0.77
Word Frequency			0.000	1.15
Bigram Frequency - Following Word			0.000	1.11
Consonant Proportion			0.001	1.29
Repetition number			0.000	0.93
Local deletion rate			0.000	12.65

Pearson Goodness-of-fit:

p = 0.000 p = 0.121

Table 59. Binary logistic regression for deletion data.

6.1.2 Continuous duration data

The two other dependent variables under investigation, /s/ duration and spectral COG, cannot be modeled using the logistic regression because they have continuous values. For continuous dependent variables, the appropriate type of model is the multiple regression (Lunneborg 1994). Unlike with the logistic regression, the multiple regression cannot have independent variables that contain attribute data; all of the independent variables must be numerical. In order to perform the multiple regression, all of the independent factors consisting of attribute data have been converted into numerical data using *indicator variables* as described above.

The results outputted from a multiple regression are similar to those from the logistic regression. The equivalent to the Odds Ratio in the logistic regression is the *estimated coefficient.* The estimated coefficient indicates how much the dependent variable (/s/ duration or spectral COG) is affected by a unit increase in the independent variable, if all the other independent variables are kept constant. Each estimated coefficient has an associated p-value, just as the Odds Ratios do in the logistic regression. When the p-value is less than the chosen confidence level ($p<0.05$ or $p<0.01$), this indicates that there is evidence of a significant relationship between the independent variable and the dependent variable.

The statistic used to evaluate the linear regression model as a whole is R^2. R^2 is the proportion of the variation in the dependent variable (/s/-duration or spectral COG) that is explained by the model. R^2 varies between 0 and 1, with higher values of R^2 indicating that the model fits the data better.

Table 60 shows the results of the multiple regression for the dependent variable /s/ duration. Nearly all of the factors considered in the deletion data are also statistically significant in the base model of /s/ duration. There is no effect of a following [d] (other than that predicted by the factors *voiced stop* and *coronal*), so it is not included as a separate factor in this analysis. *Voiceless fricatives* and *voiceless stops* have the same effect, and so they have been combined. The factors *dialect* and *monosyllabic* have marginal p-values (p=.047 and p=.023), indicating that their contribution to the model is slight. I have kept them in the analysis in order to keep the independent variables used in this analysis as similar as possible to those for analysis of deletion.

Seven usage-based factors have a significant effect on /s/ duration. The strongest of these factors is the *local deletion rate*, which has a coefficient of -18.21, indicating that the predicted duration is nearly 20 msec shorter when the 5 previous tokens of /s/ have all been deleted as compared to when the previous tokens have all been retained. As with the categorical deletion data, adding the factor of local deletion rate has a great effect on the estimated coefficients for the dialect and speech rate. With the exception of the *forward conditional bigram probability of the preceding word*, the word predictability factors that contribute significantly to the model all are in the expected direction: higher word predictability yields shorter durations. For *word frequency*, the word frequency variables reflecting the word counts from each of the different sources each contribute significantly to the model when considered separately, but the variable using the word counts from the radio transcripts is the best predictor for the model (better than using the word counts from the telephone or text sources, or the average of the three types of sources). For the other relative frequencies, the ***bigram frequency with the following word*** and the ***trigram frequency with the preceding two words*** both contribute significantly, with a ten-fold increase in the relative frequency predicting that the duration would be shortened by approximately 4 msec.

Two conditional probabilities are statistically significant: the *reverse conditional bigram probability of the following word*, and the *forward conditional bigram probability of the preceding word*. The reverse bigram probability of the following word predicts a shorter /s/ duration when the word containing the /s/ is more probable, as expected. The forward conditional bigram probability predicts the opposite for the probability of the word given the preceding word; when the word is more probable, the predicted effect is a somewhat longer /s/ duration, although the magnitude of the forward conditional bigram probability of the preceding word is lower (5.73) than that of the reverse conditional bigram probability of the following word (-13.79).

While only marginally significant ($p = 0.02$) the effect of *consonant proportion for individual words* is in the expected direction, with a larger consonant proportion yielding a shorter duration.

Interestingly, the factor of *repetition* is not significant, indicating that words that have already been spoken do not have shorter /s/ durations, as would have been expected.

The R^2 value for the model as a whole indicates that the predictors explain 22.6% of the variance in duration in the base model and 23.3% when the usage-based factors are added. This indicates that the model with the usage-based factors fits the data only slightly better than the base model.

	Base		With Usage Variables	
Predictor	p	Coefficient	p	Coefficient
Dialect: interior	0.047	0.76	0.013	-2.48
Male	0.000	-1.62	0.000	-3.52
Speech Rate	0.000	-7.58	0.000	-47.98
Preceding Vowel (Ref: 'o/u')				
a/e	0.000	4.01	0.000	10.63
i/w/y	0.000	6.87	0.000	8.43
Following Segment (Ref: Voiceless stop or fricative)				
Glide	0.000	-15.70	0.000	-5.28
Liquid	0.000	13.37	0.000	13.68
Nasal	0.000	9.04	0.000	11.99
Pause	0.000	39.14	0.000	45.67
Voiced stop	0.042	-1.36	0.000	-3.64
Vowel	0.000	30.60	0.000	46.51
Following labial	0.000	2.29	0.004	2.90
Following velar	0.000	11.02	0.000	15.85
Grammatical Status (Ref: 'Lexical')				
Internal	0.004	-2.32	0.195	-1.30
Plural	0.000	-4.38	0.000	-6.94
Verb	0.000	-4.25	0.000	-4.91
FunctionWord	0.000	-4.75	0.000	-3.61
SyllableCount	0.009	0.46	0.001	-3.18
Monosyllabic	0.023	1.96	0.001	3.19
Following Stress	0.000	6.77	0.000	16.71
Post-Stress or Stressed syllable	0.000	2.50	0.000	4.23
Word Frequency - Radio			0.002	-3.13
Bigram Frequency - Following Word			0.000	-4.66
Reverse Conditional Bigram Probability - Following Word			0.000	-13.79
Conditional Bigram Probability - Preceding Word			0.000	5.73
Trigram Frequency - Two Preceding Words			0.000	-4.27
Consonant Proportion			0.020	-2.32
Local deletion rate			0.000	-18.21
R^2	22.6%		23.3%	

Table 60. Multiple regression for /s/ duration data.

6.1.3 Continuous Spectral COG data

Table 61 shows the results of the multiple regression for the spectral COG data. In the base model, the independent variables *function word* and *monosyllabic* are not statistically significant at the $p<0.05$ level, and *syllable count* was only marginally so. I have opted to include them in the analysis in order to keep this model as equivalent as possible to those for deletion and /s/ duration. Because these variables are included in both the base model and the model with usage-based variables, including these variables in the analysis should not affect the data.

Four of the usage-based factors are significant at the $p<0.05$ level. As with the models for deletion and /s/ duration, the *local deletion* has the greatest effect, with an estimated coefficient of –579.8. Three measures of word predictability contribute to the model, *word frequency* as measured using the word counts from the radio source, *bigram frequency of the word with the following word*, and *mutual information of word with the following word*. All three of these measures have the expected tendency, with more predictable words and bigrams having a lower predicted spectral COG. However, the magnitudes of the estimated coefficients for these variables are small in comparison to some of the other variables (particularly in comparison to the place of articulation of the following segment), indicating that the impact of word predictability on the articulation of the /s/ is not very great.

	Base		With Usage Variables	
Predictor	p	Coefficient	p	Coefficient
Dialect: interior	0.000	820.54	0.000	759.85
Male	0.000	-287.39	0.000	-274.01
Speech Rate	0.003	-9.97	0.692	1.36
Preceding Vowel (Ref: 'o/u')				
a/e	0.000	219.14	0.000	217.16
i/w/y	0.000	278.15	0.000	270.34
Following Segment (Ref: Voiceless stop or fricative)				
Glide	0.000	-677.82	0.000	-671.35
Liquid	0.000	-355.89	0.000	-355.82
Nasal	0.000	-90.66	0.000	-92.39
Pause	0.000	139.62	0.000	149.86
Voiced stop	0.000	-180.38	0.000	-192.23
Vowel	0.000	284.42	0.000	303.60
Following labial	0.016	30.11	0.020	28.95
Following velar	0.000	163.30	0.000	166.95
Grammatical Status (Ref: 'Lexical')				
Internal	0.025	-40.74	0.182	-24.54
Plural	0.000	-44.69	0.002	-40.33
Verb	0.007	-42.61	0.244	-18.55
FunctionWord	0.225	-19.17	0.612	-8.24
SyllableCount	0.048	7.94	0.563	-2.66
Monosyllabic	0.172	26.76	0.039	40.45
Following Stress	0.000	36.29	0.000	39.54
Post-Stress or Stressed syllable	0.000	45.03	0.000	47.92
Word Frequency - Radio			0.002	-15.56
Bigram Frequency - Following Word			0.002	-15.17
Bigram Mutual information			0.021	-6.92
Local deletion rate			0.000	-579.79
R^2	23.4%		26.0%	

Table 61. Multiple regression for spectral COG data.

6.2 Model for Content Words and Function Words

The results of the analyses in the preceding section support the usage-based hypothesis that higher word predictability leads to more advanced lenition. However, both the binary logistic regression and the multiple regression calculate models that assume that there are no interactions between the independent variables (the effect of each independent variable is assumed to be the same regardless of the value of the other independent variables). For the present data, there is reason to believe that this may not be the case. By splitting the data into smaller subsets and modeling the subsets separately, we are able to see how the independent variables contribute in each of the different subsets.

The split in the word-final /s/ data to be investigated is into the categories *content words* and *function words*. Jurafsky *et al.* (2001) report that content words and function words do not behave the same with respect to usage-based factors; there are greater effects of conditional probability for function words than for content words, while content words are more strongly influenced by relative frequency. Jurafsky *et al.* make this observation based on the study of two types of reduction (word-final /t,d/ deletion for content words; vowel reduction for function words). The data in the present study allows for the comparison of the effect of usage-based variables on content words and function words.

6.2.1 Deletion in Content Words

Table 62 shows the results of the binary logistic regression model for deletion of /s/ in content words. The Odds Ratios in the base model for the factors that are statistically significant have comparable values here to the ones for the full data-set. The base model for content words has fewer independent factors than the base model for all of the data because the factors *preceding vowel* and *syllable stress* are not significant. Note that the grammatical category *Plural* consists only of adjectives and nouns for content words, and these two sub-categories do not behave statistically significantly differently from each other. The two sub-categories of ***Verb***, the 2nd person singular and the 1st person plural, are significantly different from each other, with the 1st person plural ending *–mos* being nearly 3 times more likely to delete the final /s/.

The model including usage-based variables indicates that four of the usage-based variables are statistically significant: ***word frequency*** (as measured in the radio corpus), ***bigram frequency of the word with the following word***, ***repetition number***, and ***local deletion rate***. The values of the Odds Ratio for each of these factors are very similar to the values for the full data set, with higher word probability (as measured by word frequency and the bigram frequency) predicting slightly higher /s/ deletion. The main difference between the full data set and the content word data set is that the consonant proportion is not significant for the content words, but it is significant for the full data set. The Pearson Goodness-of-fit value for the model that includes the usage-based variables (p=0.931) indicates that the model fits the data well.

	Base		With Usage Variables	
Predictor	**p**	**Odds Ratio**	**p**	**Odds Ratio**
Dialect (Ref: 'coastal')				
interior	0.000	0.50	0.000	0.69
Male	0.001	1.13	0.066	1.08
Speech Rate	0.000	1.12	0.000	1.07
Following Segment Type (Ref: 'Voiceless Stop')				
Glide	0.006	0.39	0.011	0.41
Liquid	0.000	1.89	0.000	2.04
Nasal	0.000	2.28	0.000	2.36
Pause	0.004	0.74	0.000	0.67
Voiced stop	0.000	3.04	0.000	3.53
Voiceless fricative	0.142	1.21	0.026	1.35
Vowel	0.000	0.13	0.000	0.10
Following [d]	0.000	1.69	0.001	1.58
Following labial	0.000	0.72	0.000	0.71
Following velar	0.000	0.57	0.000	0.51
Grammatical Status (Ref: 'Lexical')				
Plural	0.000	1.30	0.000	1.34
Verb	0.353	1.07	0.491	1.06
1st person plural ending *-mos*	0.000	2.98	0.000	2.56
Syllable Count	0.000	0.94	0.925	1.00
Monosyllabic	0.000	0.63	0.000	0.63
Following Stress (Ref: 'stressed/ phrase final')				
Unstressed/ monosyllabic	0.000	1.24	0.000	1.26
Word Frequency - Radio			0.001	1.09
Bigram Frequency - Following Word			0.001	1.11
Repetition number			0.028	0.93
Local deletion rate			0.000	10.92

Pearson Goodness of fit:

p = 0.000 p = 0.931

Table 62. Binary logistic regression for deletion data for word-final /s/ in content words.

6.2.2 Deletion in Function Words

Table 63 shows the results of the binary logistic regression for word-final /s/ in function words. The base model for word-final /s/ in function words does not include the factors *following labial*, *syllable count* or *grammatical status* because these factors are not statistically significant. *Syllable count* is not significant because the function words are generally either one or two syllables. Grammatical status is not significant, probably due to the much smaller number of words in the categories *verb* and *plural*. The Odds Ratios for *following segment type* in both the base model and the model with usage-based variables differ quite a bit from those for models using the full data set and the content words. Because this factor is controlled for by being included in both the base model and the model with the usage-based variables, it should not affect the results of the usage-based variables.

As for the usage-based variables, word frequency is not statistically significant when the other usage-based variables are also included. This is not surprising given the fact that this category includes most of the very frequent words. Instead of the individual word frequency, *the bigram frequency with the following word* is significant, and its Odds Ratio is higher in this model than the one for the full data set. This indicates that token frequency is important for the function words, just as it is with content words, but that the tokens of importance are bigrams rather than individual words. There are two usage-based factors that have surprising results. First, there is an effect of *repetition number* for the function words, indicating that a function word that is repeated within a discourse is less likely to be deleted upon subsequent repetitions. Second, there is a strong effect of *consonant proportion*, with function words that have higher consonant proportions being predicted to have lower deletion rates (which is just the opposite of that predicted by the usage-based exemplar model).

	Base		With Usage Variables	
Predictor	**p**	**Odds Ratio**	**p**	**Odds Ratio**
Dialect (Ref: 'coastal')				
Interior	0.000	0.37	0.000	0.51
Male	0.000	1.26	0.003	1.17
Speech Rate	0.000	1.20	0.000	1.17
Preceding Vowel (Ref: 'a/e')				
i/w/y	0.287	1.20	0.384	1.17
o	0.000	1.30	0.000	1.44
u	0.000	1.64	0.000	2.16
Following Segment Type (Ref: 'Voiceless Stop')				
Glide	0.221	1.43	0.332	1.35
Liquid	0.000	3.64	0.000	3.48
Nasal	0.000	1.94	0.000	2.04
Pause	0.001	1.46	0.686	1.05
Voiced stop	0.000	3.12	0.000	3.94
Voiceless fricative	0.004	1.42	0.000	1.65
Vowel	0.000	0.23	0.000	0.18
Following [d]	0.000	1.85	0.000	1.57
Following velar	0.009	0.82	0.000	0.74
Monosyllabic	0.000	0.77	0.000	0.71
Following Stress (Ref: 'stressed/phrase final')				
unstressed/ monosyllabic	0.000	1.24	0.000	1.37
Stressed syllable	0.000	0.57	0.000	0.54
Bigram Frequency - Following Word			0.000	1.24
Consonant Proportion			0.004	0.47
Repetition number			0.000	0.93
Local deletion rate			0.000	14.12

Pearson Goodness of fit:

p = 0.000 p = 0.480

Table 63. Binary logistic regression for deletion data for word-final /s/ in function words.

6.2.3 /s/ duration in content words

Table 64 shows the results of the multiple linear regression for /s/ duration in content words. The coefficients for the *interior dialect* and *male* variables in the base model are statistically significant but actually have the opposite effect than expected;

contrary to the full data set, for the content words, interior speakers and females have slightly shorter /s/ duration than coastal speakers and males. This may be because these speakers tend to shorten a word-final /s/ in content words rather then deleting the /s/ altogether. For *grammatical status*, /s/ in both plural and verbal endings is shorter than lexical /s/. There is no statistically significant difference in /s/ duration between the sub-categories of plural (adjective and noun, which are the only two types of plural in the category of *content word* (adjective and noun). The first person plural verbal ending –*mos* has shorter /s/ than other verbal endings. There is no effect of whether the /s/ is in a stressed syllable or not.

Five of the usage-based factors are significant at the $p<0.05$ level. As expected, the *local deletion rate* has the greatest effect, with an estimated coefficient of –38.8. The other usage factors have relatively low coefficients. Three measures of word predictability contribute to the model: the *word frequency* as measured using the word counts from the radio sources, the *bigram frequency of the word with the following word*, and the *conditional trigram probability of the word given the preceding two words*. The word frequency and the bigram frequency with the following word both have effects in the expected direction, with higher frequency words and bigrams having shorter durations. The effect of the conditional trigram probability is in the opposite direction, with more predictable words having a slightly longer duration. The fifth usage-based variable that is significant, *consonant proportion*, indicates that words that occur more frequently before a consonant have a shorter duration, even after taking into account the following segment type.

It is interesting that in the model that includes usage-based variables, the *syllable count* and *monosyllabic* variables are not significant at all (p=0.539, p=0.853). This may be due to the addition of word frequency in the model; frequent words generally have fewer syllables.

	Base		With Usage Variables	
Predictor	**p**	**Coefficient**	**p**	**Coefficient**
Dialect: interior	0.017	-1.49	0.576	0.36
Male	0.003	1.75	0.231	0.74
Speech Rate	0.000	-9.19	0.000	-6.93
Preceding Vowel (Ref: 'o/u')				
a/e	0.000	4.83	0.000	4.58
i/w/y	0.078	3.18	0.014	4.90
Following Segment (Ref: Voiceless stop or fricative)				
Glide	0.004	-12.07	0.000	-18.18
Liquid	0.000	11.35	0.178	2.70
Nasal	0.000	11.32	0.730	-0.53
Pause	0.000	34.36	0.000	39.14
Voiced stop	0.070	-2.65	0.000	-20.26
Vowel	0.000	27.66	0.000	44.96
Following labial	0.034	-2.94	0.000	5.36
Following velar	0.000	6.95	0.000	17.08
Grammatical Status (Ref: 'Lexical')				
Plural	0.000	-3.41	0.000	-3.88
Verb	0.042	-2.32	0.395	-1.04
1st p pl verb	0.000	-7.91	0.000	-12.19
SyllableCount	0.051	0.52	0.539	-0.19
Monosyllabic	0.012	-5.03	0.853	0.40
Following Stress	0.000	4.67	0.000	5.70
Stressed syllable	0.000	6.82	0.000	6.87
Word Frequency - Radio			0.015	-0.98
Bigram Frequency - Following Word			0.000	-2.91
Consonant Proportion			0.005	-3.41
Conditional Trigram Probability - 2 Preceding Words			0.000	1.39
Local deletion rate			0.000	-38.81
R^2	18.9%		30.6%	

Table 64. Multiple regression for /s/ duration data for word-final /s/ in content words.

6.2.4 /s/ duration in function words

The results of the multiple linear regression for /s/ duration in function words are shown in Table 65. For grammatical status, the only statistically significant distinction is plural /s/ in comparison with verbal /s/ or lexical /s/ (but note that *es* is the only function word with verbal /s/). In the base model, the variable *syllable count* is not significant, although *monosyllabic* is (with monosyllabic words having longer /s/ duration). This is probably due to the low number of function words with more than two syllables.

In the analysis including usage-based variables, the usage-based variables that are significant are: *word frequency, bigram frequency of the word with the following word, mutual information of the word with the following word, repetition number, conditional trigram probability of the word based on the two preceding words,* and th*e local deletion rate*. Two of the measures of word predictability, the bigram frequency of the word with the following word and the mutual information of the word with the following word, both predict shorter durations when the word is more probable. The coefficient of the word frequency and the conditional trigram probability predict shorter durations when the word is more probable. However, the effect of word frequency is only marginally significant (p=0.036), and its magnitude is less than either the bigram frequency and the bigram mutual information, indicating that individual word frequency is not as important as the is co-occurrence with the following word. The effect of repetition number on /s/ duration is in keeping with the results from deletion, with subsequent repetitions of a function word being predicted to have a slightly longer duration than the first repetition.

	Base		With Usage Variables	
Predictor	**p**	**Coefficient**	**p**	**Coefficient**
Dialect: interior	0.000	9.05	0.000	3.23
Male	0.000	-6.17	0.000	-4.58
Speech Rate	0.000	-5.92	0.000	-5.32
Preceding Vowel (Ref: 'o/u')				
a/e	0.000	4.67	0.000	5.53
i/w/y	0.071	4.41	0.016	5.78
Following Segment (Ref: Voiceless stop or fricative)				
Glide	0.000	-31.83	0.000	-32.48
Liquid	0.000	-11.37	0.000	-8.91
Nasal	0.007	-3.15	0.084	-1.98
Pause	0.000	30.03	0.000	31.92
Voiced stop	0.000	-23.37	0.000	-22.27
Vowel	0.000	36.12	0.000	37.42
Following velar	0.000	11.01	0.000	11.85
Grammatical Status (Ref: 'Lexical' or 'Verb')				
Plural	0.032	-2.03	0.004	-2.92
Monosyllabic	0.000	3.03	0.596	-0.74
Following Stress	0.000	8.14	0.000	8.55
Stressed syllable	0.000	12.31	0.000	14.06
Word Frequency - Average			0.036	1.66
Bigram Frequency - Following Word			0.000	-4.03
Bigram Mutual Information			0.000	-2.07
Repetition			0.000	0.63
Conditional Trigram Probability - Two Preceding Words			0.000	1.96
Local deletion rate			0.000	-38.39
R^2	28.8%		34.6%	

Table 65. Multiple regression for /s/ duration data for word-final /s/ in function words.

6.2.5 Spectral COG in content words

The results of the multiple regression of the Spectral COG in content words are shown in Table 66. A number of the independent variables that are significant in the model for the full data set are not significant for the content words: *speech rate, syllable count*, *monosyllabic*, *following stress,* and *stressed syllable* are not significant.

Three of the usage-based variables contribute significantly to the model of spectral COG for content words: ***bigram frequency of the word with the following word***, ***repetition number***, and ***local deletion rate***. No measure of individual word frequency contributes significantly. While the local deletion rate has a great effect (with a coefficient of –524.65), the other two variables have very low magnitudes (-20.66 and -18.04, smaller effect than any of the other variables).

	Base		**With Usage Variables**	
Predictor	p	Coefficient	p	Coefficient
Dialect: interior	0.000	735.24	0.000	676.11
Male	0.000	-311.69	0.000	-297.30
Preceding Vowel (Ref: 'o/u')				
a/e	0.000	238.64	0.000	238.08
i/w/y	0.000	231.02	0.000	230.19
Following Segment (Ref: Voiceless stop or fricative)				
Glide	0.000	-563.73	0.000	-569.11
Liquid	0.000	-329.78	0.000	-336.72
Nasal	0.022	-67.24	0.064	-54.31
Pause	0.000	163.93	0.000	188.72
Voiced stop	0.000	-158.62	0.000	-164.41
Vowel	0.000	305.63	0.000	333.05
Following labial	0.046	53.37	0.040	55.24
Following velar	0.000	205.52	0.000	222.67
Grammatical Status (Ref: 'Lexical')				
Plural	0.006	-41.10	0.008	-39.68
Verb	0.040	-41.69	0.124	-31.11
1st p pl verb *-mos*	0.015	-76.16	0.087	-53.57
Bigram Frequency -Following Word			0.002	-20.66
Repetition			0.028	-18.04
Local deletion rate			0.000	-524.65
R^2	22.6%		25.0%	

Table 66. Multiple regression for spectral COG in word-final /s/ in content words.

6.2.6 Spectral COG in function words

The results of the multiple regression for the spectral COG in word-final /s/ in function words are in Table 67. Five of the usage-based variables contribute significantly to the model: *bigram frequency of the word with the following word, mutual information of the word with the following word, conditional trigram probability of the word based on the two preceding words, repetition number,* and the *local deletion rate.* The word probability variables produce a similar effect on this model for the spectral COG for function words as they do for /s/ duration; the estimated coefficients for bigram frequency and mutual information measures predict lower spectral COG for more predictable words (indicating more lenition in the articulation of /s/), while the estimated coefficient for the conditional trigram probability predicts higher spectral COG for more predictable words. The effect of the local deletion rate is twice as high among function words than among content words (-1140.02 for function words as compared with -524.65 for content words).

	Base		With Usage Variables	
Predictor	**p**	**Coefficient**	**p**	**Coefficient**
Dialect: interior	0.000	957.81	0.000	796.59
Male	0.000	-314.99	0.000	-270.65
Speech Rate	0.000	-51.45	0.000	-35.52
Preceding Vowel (Ref: 'o/u')				
a/e	0.000	120.93	0.000	123.68
i/w/y	0.108	98.61	0.034	125.19
Following Segment (Ref: Voiceless stop or fricative)				
Glide	0.000	-950.60	0.000	-911.10
Liquid	0.000	-777.37	0.000	-731.27
Nasal	0.000	-255.09	0.000	-244.77
Pause	0.033	99.64	0.001	166.06
Voiced stop	0.000	-466.66	0.000	-455.35
Vowel	0.000	589.28	0.000	596.61
Following labial	0.001	91.59	0.003	74.72
Following velar	0.000	245.70	0.000	253.79
Grammatical Status (Ref: 'Lexical')				
Plural	0.020	-69.45	0.002	-93.56
Verb	0.033	78.47	0.006	100.63
Following Stress	0.027	39.09	0.004	50.58
Stressed syllable	0.000	314.14	0.000	333.31
Bigram Frequency - Following Word			0.000	-60.00
Bigram Mutual Information			0.022	-25.29
Conditional Trigram Probability - Two Preceding Words			0.010	20.02
Repetition			0.000	16.81
Local deletion rate			0.000	-1140.02
$\mathbf{R^2}$	30.3%		37.2%	

Table 67. Multiple regression for spectral COG in word-final /s/ in function words.

6.3 Word-final /s/ followed by a consonant or vowel

Another split in the data to be considered for word-final /s/ is whether each token is followed by a consonant or by a vowel. If frequent multi-word sequences are stored in the lexicon as a single unit, then the effect of frequency should be different for these two categories. When a two-word sequence is stored as a single lexical entry, and the second word begins with a vowel, the word-final /s/ should act more as a word-internal prevocalic /s/, and should therefore not be deleted as often. There may also be a different effect of individual word frequency when followed by a consonant or a vowel, if, as suggested by Bybee (2001b) in her discussion of French liaison, very high frequency words may have more than one variant stored in the lexicon (one used when followed by a consonant and another when followed by a vowel).

6.3.1 Deletion for word-final /s/ followed by a consonant

The data from the model of deletion for word-final /s/ followed by a consonant is shown in Table 68. Six of the usage-based variables are significant in this model: *word frequency*, *bigram probability of the word with the following word*, *consonant proportion*, *conditional bigram probability of the word given the preceding word*, *repetition number*, and *local deletion rate*. With the exception of the conditional bigram probability, which is not significant for the full data set, the effect of all of these variables is about the same for the full data set.

	Base		With Usage Variables	
Predictor	**p**	**Odds Ratio**	**p**	**Odds Ratio**
Dialect (Ref: 'coastal')				
interior	0.000	0.40	0.000	0.53
Male	0.000	1.21	0.000	1.16
Speech Rate	0.000	1.29	0.000	1.25
Preceding Vowel (Ref: 'a/e')				
i/w/y	0.146	0.84	0.469	0.91
o	0.000	1.14	0.000	1.18
u	0.000	1.59	0.000	1.94
Following Segment Type (Ref: 'Voiceless Stop')				
Liquid	0.000	2.61	0.000	2.67
Nasal	0.000	2.05	0.000	2.13
Voiced stop	0.000	3.16	0.000	3.84
Voiceless fricative	0.000	1.38	0.000	1.55
Following [d]	0.000	1.78	0.000	1.49
Following labial	0.002	0.83	0.007	0.84
Following velar	0.000	0.67	0.000	0.60
Grammatical Status (Ref: 'Lexical', 'Verb')				
Plural	0.003	1.13	0.000	1.22
Verbal ending *-mos*	0.000	2.14	0.000	1.78
FunctionWord	0.000	1.42	0.283	1.07
SyllableCount	0.023	0.96	0.063	1.04
Monosyllabic	0.000	0.71	0.000	0.72
Following Stress (Ref: 'stressed/phrase final')				
unstressed/ monosyllabic	0.000	1.36	0.000	1.45
Stressed syllable	0.001	0.81	0.000	0.74
Word Frequency – Average			0.000	1.19
Bigram Probability - Following Word			0.000	1.24
Consonant Proportion			0.000	1.52
Conditional Bigram Probability - Preceding Word			0.000	0.93
Repetition number			0.000	0.93
Local deletion rate			0.000	10.90

Pearson Goodness of fit:

p = 0.360 p = 0.871

Table 68. Binary logistic regression for deletion data for word-final /s/ followed by a consonant.

6.3.2 Deletion for word-final /s/ followed by a vowel

The data for deletion of word-final /s/ followed by a vowel is shown in Table 69. In the base model, there is no effect of the variables *male, speech rate* or *stressed syllable*, so these have not been included. There is also no distinction between the contributions of different following segments (all vowels in this case), so no variables regarding following segment have been included.

A number of the usage-based variables are significant, although they do not improve the overall fit of the model; the Pearson Goodness-of-fit measurement is p=0.000 for both the base model and the model including usage-based variables. The usage-based variables which are significant are: *bigram probability of the word with the following word*, *reverse conditional bigram probability of the word given the following word*, *conditional trigram probability of the word given the two preceding words*, and *local deletion rate*. None of the measures of word frequency are statistically significant. The estimated coefficient for the bigram probability of the word with the following word indicates that more frequent bigrams are predicted to have more deletion. This is somewhat surprising because if frequent bigrams are stored as one unit in the lexicon, and the second word in the sequence begins with a vowel, I would predict that more frequent bigrams would have lower deletion rates. Furthermore, the effect of the bigram mutual information (which is a measure of the cohesion of the two words) does not have a significant effect in this model. However, the reverse conditional bigram probability of the word given the following word is significant, and does have the expected effect, with more probable words having a lower predicted deletion rate. None of the context proportions (consonant proportion, vowel proportion, or pause proportion) are statistically significant.

	Base		With Usage Variables	
Predictor	**p**	**Odds Ratio**	**p**	**Odds Ratio**
Dialect (Ref: 'coastal')				
interior	0.000	0.55	0.213	0.89
Preceding Vowel (Ref: 'a/e')				
i/w/y	0.917	0.97	0.937	1.03
o	0.000	1.67	0.000	1.61
u	0.395	1.35	0.042	2.12
Grammatical Status (Ref: 'Lexical')				
Plural	0.000	0.58	0.027	0.68
Verb	0.157	1.26	0.544	1.12
FunctionWord	0.000	2.30	0.000	2.29
SyllableCount	0.003	0.82	0.037	0.87
Monosyllabic	0.000	0.41	0.000	0.45
Following Stress (Ref: 'stressed/phrase final')				
unstressed/ monosyllabic	0.071	0.83	0.316	0.89
Bigram Frequency - Following Word			0.002	1.19
Reverse Conditional Bigram Probability - Following Word			0.008	0.64
Conditional Trigram Probability - Two Preceding Words			0.004	0.70
Local deletion rate			0.000	36.01

Pearson Goodness of fit:

p = 0.000 p = 0.000

Table 69. Binary logistic regression for deletion data for word-final /s/ preceding a vowel.

6.3.3 /s/ duration for word-final /s/ followed by a consonant

Table 70 shows the results of the multiple regression of /s/ duration for word-final /s/ that is followed by a consonant. For the usage-based variables, there are 5 measures of word probability that contribute: *word frequency, bigram frequency of the word with the following word, bigram mutual information of the word with the following word, bigram* *frequency of the word with the preceding word,* and *conditional trigram probability of the word given the preceding and following words.* Of these, all but the bigram frequency of the word with the preceding word predict a shorter duration of /s/ when the word is more predictable. Two other usage-based variables are also significant: the *pause proportion* of the word, and the *local deletion rate.* When a word appears more frequently before a pause, the word-final /s/ is predicted to have a longer duration, even though in this data set, they are all preceding a consonant.

For both the base model and the model with usage based variables, the values of R^2 are much smaller than for the full data set (10.6% for the base model of word-final /s/ before a consonant in comparison with 22.6% for the base model of the full data set). This indicates that the model for this subset of the data does not explain as much of the variation in duration as the model for the full data set does. This is not surprising because in the full data set, there is great variation in duration due to the phonetic context (consonant/vowel/pause), and the variables for the following segment can therefore account for a large percentage of the variation. When considering only word-final tokens of /s/ followed by a consonant, there is not as much total variation, and the variables for the following segment do not account for as great a percentage of the variation.

	Base		With Usage Variables	
Predictor	**p**	**Coefficient**	**p**	**Coefficient**
Dialect: interior	0.000	-4.40	0.000	-5.99
Male	0.000	-4.14	0.000	-4.03
Speech Rate	0.000	-6.70	0.000	-6.47
Preceding Vowel (Ref: 'o/u')				
a/e	0.001	1.85	0.000	1.98
i/w/y	0.042	3.49	0.339	1.68
Following Segment (Ref: Voiceless stop or fricative)				
Liquid	0.000	11.34	0.000	11.79
Nasal	0.000	8.46	0.000	9.33
Voiced stop	0.001	-2.11	0.007	-1.82
Following velar	0.000	8.85	0.000	9.06
Grammatical Status (Ref: 'Lexical' or 'Verbal')				
Plural	0.001	-2.13	0.000	-2.61
Verbal ending *-mos*	0.001	-6.12	0.014	-4.52
FunctionWord	0.000	-4.79	0.127	-1.33
SyllableCount	0.021	0.62	0.799	-0.08
Following Stress	0.000	4.62	0.000	4.52
Stressed syllable	0.009	2.48	0.016	2.34
Word Frequency – Radio			0.000	-1.60
Bigram Frequency - Following Word			0.000	-4.24
Bigram Mutual Information			0.000	-1.16
Bigram Frequency - Preceding Word			0.006	0.97
Conditional Trigram Probability - Preceding & Following Words			0.000	-2.90
Pause Proportion			0.000	10.55
Local deletion rate			0.000	-9.22
R^2	10.6%		12.0%	

Table 70. Multiple regression for /s/ duration for word-final /s/ followed by a consonant.

6.3.4 /s/ duration for word-final /s/ followed by a vowel

Table 71 shows the data for the multiple regression for /s/ duration for word-final /s/ that is followed by a vowel. For the usage-based variables, there are 3 measures of word probability that contribute: *bigram frequency of the word with the following word*, the *reverse conditional bigram probability of the word given the following word*, and the *trigram frequency of the word with the two preceding words.* As with the deletion data before a vowel, word frequency is not significant. The estimated coefficients of the *bigram frequency of the word with the following word* and the *reverse conditional bigram probability of the word given the following word* both predict shorter /s/ durations when the word has higher predictability. The trigram frequency of the word with the two preceding words predicts a longer duration when the word is more predictable. There are two other usage-based variables that are significant: the *consonant proportion* (words that appear more often before a consonant have shorter /s/, even though they appear before a vowel) and *local deletion rate* (when the local deletion rate is higher, the /s/ duration is lower).

	Base		**With Usage Variables**	
Predictor	p	Coefficient	p	Coefficient
Dialect: interior	0.002	-1.95	0.000	-3.23
Male	0.000	-3.14	0.000	-2.51
Speech Rate	0.000	-7.13	0.000	-7.02
Preceding Vowel (Ref: 'o')				
a/e	0.000	5.12	0.000	4.59
i/w/y	0.029	4.87	0.018	5.50
u	0.000	8.49	0.016	5.62
Grammatical Status (Ref: 'Lexical')				
Plural	0.000	-2.79	0.002	-2.39
Verbal ending *-mos*	0.000	-11.82	0.000	-9.05
FunctionWord	0.000	-6.33	0.000	-4.05
SyllableCount	0.007	0.88	0.668	-0.15
Following Stress	0.000	9.24	0.000	8.87
Stressed syllable	0.000	8.25	0.000	7.16
Bigram Frequency - Following Word			0.000	-1.90
Reverse Conditional Bigram Probability - Following Word			0.000	-1.57
Trigram Frequency - Two Preceding Words			0.000	2.08
Consonant Proportion			0.017	-3.51
Local deletion rate			0.000	-9.10
R^2	11.3%		13.4%	

Table 71. Multiple regression for /s/ duration data for word-final /s/ preceding a vowel.

6.3.5 Spectral COG for word-final /s/ followed by a consonant

The model for spectral COG for word-final /s/ that is followed by a consonant is shown in Table 72. There are seven usage-based variables that are significant: *word frequency, bigram probability of the word with the following word, bigram mutual information, conditional trigram probability of the word given the two preceding words*, the *consonant proportion, repetition number*, and *local deletion rate*. The effects of all of these variables are similar to the effects of the same variables in the data sets already reviewed. The coefficients of the word frequency, the bigram probability of the word with the following word, and the bigram mutual information predict lower spectral COG when the word is more predictable, while the conditional trigram probability of the word given the two preceding words predicts just the opposite. The coefficient of the consonant proportion variable predicts lower spectral COG when the word appears more often before a consonant.

	Base		With Usage Variables	
Predictor	**p**	**Coefficient**	**p**	**Coefficient**
Dialect: interior	0.000	1036.61	0.000	862.85
Male	0.000	-328.28	0.000	-298.93
Speech Rate	0.000	-101.43	0.000	-82.63
Preceding Vowel (Ref: 'o/u')				
a/e	0.001	49.31	0.000	59.73
i/w/y	0.000	197.68	0.000	177.01
Following Segment (Ref: Voiceless stop)				
Liquid	0.000	-636.40	0.000	-621.78
Nasal	0.000	-283.24	0.000	-267.33
Voiced stop	0.000	-524.52	0.000	-513.17
Voiceless fricative	0.005	-97.78	0.002	-106.97
Following labial	0.000	153.03	0.000	130.13
Following velar	0.000	330.46	0.000	334.20
Grammatical Status (Ref: 'Lexical')				
Plural	0.000	-103.16	0.000	-118.50
Verbal ending *-mos*	0.000	-286.62	0.000	-166.71
FunctionWord	0.000	-97.98	0.331	23.44
SyllableCount	0.000	36.83	0.829	1.94
Monosyllabic	0.000	98.70	0.000	100.31
Following Stress	0.000	118.21	0.000	118.74
Stressed syllable	0.000	124.35	0.000	149.33
Word Frequency - Average			0.000	-75.47
Bigram Frequency - Following Word			0.000	-68.49
Bigram Mutual Information			0.001	-20.55
Conditional Trigram Probability - Two Preceding Words			0.000	19.26
Repetition			0.000	19.75
Consonant Proportion			0.000	-223.78
Local deletion rate			0.000	-1157.96
R^2	27.3%		34.2%	

Table 72. Multiple regression for spectral COG data in word-final /s/ preceding a consonant.

6.3.6 Spectral COG for word-final /s/ followed by a vowel

The model for spectral COG for word-final /s/ that is followed by a vowel is shown in Table 73. In the base model, several factors are not included because they are not significant: *grammatical status* (there is no distinction between *lexical, noun,* and *verb*, but there is a difference for the verbal ending *–mos*), *syllable count*, and *following stress.* When taken separately, *function word* and *monosyllabic* are both significant, but when taken together, neither of them are. I have included *function word* in this analysis because it has a lower p-value, indicating that it is more significant.

There are four usage-based variables that are significant: *word frequency*, *bigram probability of the word with the following word*, *conditional trigram probability of the word given the two preceding words*, *repetition number*, and *local deletion rate*. As with the previous analyses, although all four of these variables are statistically significant, the magnitude of the effect of all the variables other than the local deletion rate is relatively low.

	Base		With Usage Variables	
Predictor	**p**	**Coefficient**	**p**	**Coefficient**
Dialect: interior	0.000	643.89	0.000	591.58
Male	0.000	-350.37	0.000	-337.13
Speech Rate	0.002	18.73	0.000	26.83
Preceding Vowel (Ref: 'o/u')				
a/e	0.000	282.84	0.000	279.79
i/w/y	0.000	217.14	0.000	222.37
Verbal ending -mos			0.204	-40.32
FunctionWord	0.003	-39.08	0.111	-22.70
Stressed syllable	0.000	134.32	0.000	140.30
Bigram Frequency - Following Word			0.007	-18.63
Conditional Trigram Probability - 2 Preceding Words			0.019	12.75
Repetition			0.035	-9.57
Local deletion rate			0.000	-448.78
R^2	22.7%		24.9%	

Table 73. Multiple regression for spectral COG in word-final /s/ preceding a vowel.

6.4 Plural -s

I have performed a separate analysis of deletion, /s/ duration, and spectral COG of the plural marker *–s* using a total of 26,831 tokens. For the plural, in addition to the usage-based variables evaluated in the preceding sections, I have included four other plural-specific usage-based variables:

- Deletion of preceding tokens in the NP
- Plural disambiguation: whether or not the word would be ambiguous for number if the word-final /s/ is deleted
- Plural disambiguation in the NP: whether or not the NP would be ambiguous if the word-final /s/ of all elements in the NP are deleted
- Singular-to-plural ratio: ratio of the number of tokens found in the singular as compared to the plural

6.4.1 Deletion of plural -s

Table 74 shows the result of the binary logistic regression for plural /s/. The *part of speech* is separated into four categories: *noun, adjective, pronoun,* and *first modifier*. The variable *function word* is not included in this analysis because it is redundant with the part of speech; the nouns and adjectives are content words, and the pronouns and first modifiers are function words.

Six of the usage-based variables that have been evaluated for all of the data sets contribute to the model: *word frequency, bigram probability of the word with the following word, conditional trigram probability of the word given the two preceding* *words*, *consonant proportion*, the *repetition number*, and *local deletion rate*. The word frequency and the bigram probability of the word with the following word predict that more probable plural words are more likely to have a deleted /s/. The conditional trigram probability predicts that plural words that are more predictable based on the two preceding words are less likely to have a deleted /s/. The consonant proportion predicts that plural words that appear more frequently before a consonant are more likely to have a delete /s/. As with the other subsets evaluated, the effect of the repetition number is that words in their first occurrence are more likely to have deleted /s/ than in subsequent repetitions.

Of the additional usage-based variables for plurals (deletion with the NP, disambiguation between plural and singular for the word or within the NP, and the singular-to-plural ratio), only the singular-to-plural ratio contributes to the model. Its Odds Ratio is 1.01, indicating that an increase in the number of singular tokens in comparison to the number of plural tokens causes a very small increase in the probability of deletion. While the variable reflecting the deletion of preceding tokens in the NP is significant when taken by itself, the variable *local deletion rate*, which contains information about the deletion of the 5 preceding tokens of /s/ both within the NP and preceding the NP, is a much better predictor. This means that the 'perseverance' effect within an NP described by Poplack (1980a) may be simply the result of the local deletion rate rather than a plural-specific process.

	Base		With Usage Variables	
Predictor	**p**	**Odds Ratio**	**p**	**Odds Ratio**
Dialect (Ref: 'coastal')				
Interior	0.000	0.42	0.000	0.57
Male	0.000	1.25	0.000	1.17
Speech Rate	0.000	1.14	0.000	1.09
Preceding Vowel (Ref: 'a/e')				
i/w/y	0.755	1.18	0.860	0.90
o	0.000	1.18	0.000	1.19
u	0.000	1.61	0.000	1.88
Following Segment Type (Ref: 'Voiceless Stop')				
Glide	0.114	0.63	0.144	0.65
Liquid	0.000	2.89	0.000	3.03
Nasal	0.000	2.09	0.000	2.22
Pause	0.788	0.98	0.533	0.94
Voiced stop	0.000	3.26	0.000	3.85
Voiceless fricative	0.002	1.38	0.000	1.56
Vowel	0.000	0.15	0.000	0.12
Following [d]	0.000	1.81	0.000	1.66
Following labial	0.083	0.88	0.104	0.88
Following velar	0.000	0.71	0.000	0.64
Part of Speech (Ref: 'Noun')				
First Modifier	0.000	1.45	0.006	1.25
Adjective	0.023	0.91	0.559	0.97
Pronoun	0.000	2.05	0.000	1.64
Syllable Count	0.021	0.96	0.306	1.02
Monosyllabic	0.000	0.69	0.000	0.73
Following Stress (Ref: 'stressed/phrase final')				
Unstressed/ monosyllabic	0.000	1.32	0.000	1.38
Stressed syllable	0.571	1.48	0.498	1.63
Word Frequency – Radio			0.001	1.11
Bigram Frequency - Following Word			0.000	1.16
Conditional Trigram Probability - Two Preceding Words			0.000	0.93
Consonant Proportion			0.009	1.27
Repetition number			0.000	0.93
Local deletion rate			0.000	12.78
Singular-to-Plural ratio			0.006	1.01

Pearson Goodness of fit:

p = 0.000 p = 0.944

Table 74. Binary logistic regression for deletion data for plural –*s*.

6.4.2 /s/ duration of plural -s

The results of the multiple regression for /s/ duration for plural –s is shown in Table 75. Five of the usage-based variables that have been evaluated for all of the data sets contribute to the model: *word frequency, bigram probability of the word with the following word, bigram mutual information with the following word, consonant proportion,* and *local deletion rate*. All of these factors have the expected effect, with more predictable words and words with higher consonant proportions having a shorter /s/ duration. These results are in keeping with the results for the full data set in Table 60, although the magnitudes of the estimated coefficients for the plural data set are smaller than for the full data set. Of the additional factors for plural –s, *plural disambiguation* is significant. When there is additional disambiguating information in the word other than the final /s/, the duration of the word-final /s/ is predicted to be 2.45 msec longer than if there is no way of disambiguating the singular from the plural if the /s/ is deleted. This is in the opposite direction from what would be expected.

	Base		With Usage Variables	
Predictor	p	Coefficient	p	Coefficient
Dialect: interior	0.190	-0.74	0.000	-2.59
Male	0.012	-1.27	0.069	-0.92
Speech Rate	0.000	-7.69	0.000	-7.32
Preceding Vowel (Ref: all others)				
o	0.000	-3.72	0.000	-3.57
Following Segment (Ref: Stop or fricative)				
Glide	0.000	-18.19	0.000	-17.61
Liquid	0.000	12.58	0.000	12.93
Nasal	0.000	9.92	0.000	10.09
Pause	0.000	35.85	0.000	34.03
Vowel	0.000	29.00	0.000	29.54
Following velar	0.000	10.14	0.000	10.18
Part of Speech (Ref: 'Noun', 'Adjective')				
First Modifier	0.000	-5.91	0.000	-4.99
Pronoun	0.000	-5.21	0.000	-5.97
SyllableCount	0.002	0.76	0.947	-0.02
Following Stress	0.000	6.94	0.000	6.99
Word Frequency – Radio			0.010	-0.98
Bigram Frequency - Following Word			0.000	-1.60
Bigram Mutual Information			0.000	-1.45
Consonant Proportion			0.008	-3.11
Local deletion rate			0.000	-13.88
Plural Disambiguation			0.000	2.45
R^2	20.8%		22.0%	

Table 75. Multiple regression for /s/ duration for plural –*s*.

6.4.3 Spectral COG of plural -s

The results of the multiple regression for the spectral COG for plural –s are shown in Table 76. In the base model, the part-of-speech (noun/adjective/pronoun/first modifier), speech rate, and following stress are not significant, so they are not included in the model.

In the model with usage-based variables, the *word frequency*, the *bigram probability of the word with the following word*, the *conditional bigram probability of the word given the preceding word*, the *local deletion rate*, and *plural disambiguation* are significant. Similar to the effect on /s/ duration for plural –s, the factor ***plural disambiguation*** predicts more lenition (lower spectral COG) when the /s/ is the only indication of plurality in the word.

	Base		**With Usage Variables**	
Predictor	p	Coefficient	p	Coefficient
Dialect: interior	0.000	711.48	0.000	655.07
Male	0.000	-298.14	0.000	-282.39
Preceding Vowel (Ref: 'i/o/u')				
a/e	0.000	212.19	0.000	208.43
Following Segment (Ref: Voiceless stop or fricative)				
Glide	0.000	-697.79	0.000	-686.70
Liquid	0.000	-374.35	0.000	-371.18
Nasal	0.006	-68.43	0.009	-65.48
Pause	0.000	148.27	0.000	174.92
Voiced stop	0.000	-129.83	0.000	-132.08
Vowel	0.000	295.94	0.000	327.24
Following labial	0.005	62.50	0.002	68.31
Following velar	0.000	184.77	0.000	200.79
Monosyllabic	0.000	-72.44	0.065	-35.45
Stressed syllable	0.069	390.50	0.074	377.30
Word Frequency - CallHome			0.032	-14.22
Bigram Frequency - Following Word			0.002	-23.94
Conditional Bigram Probability - Preceding Word			0.063	8.67
Local deletion rate			0.000	-532.71
Plural Disambiguation			0.001	34.97
R^2	21.3%		23.8%	

Table 76. Multiple regression for spectral COG in plural *–s*.

6.5 Verbal ending for first person plural (-mos)

While the verbal ending for the first person plural (*-mos*) is much less frequent than the plural (only 1167 tokens in the Broadcast News corpus), it is an important data set to evaluate because the word-final /s/ is part of a morpheme, but it does not cause any ambiguity when it is deleted.

6.5.1 Deletion in -mos

In the case of the verbal ending *–mos,* many of the independent variables that have been explored in the full data set are not applicable: the preceding vowel is always /o/, the grammatical status is always the same, the word is never monosyllabic, and the syllable containing the /s/ is always unstressed. In addition, the following place of articulation (labial/coronal/velar) and following stress are not significant in the base model and therefore not included in the results shown in Table 77.

There are only 3 usage-based variables that contribute significantly to /s/ deletion in the morpheme *–mos*: *word frequency, conditional trigram probability based on the two preceding words,* and the *local deletion rate*. As with the other analyses, more frequent words are predicted to have higher deletion rates, and words that are more probable based on the preceding two words are predicted to have lower deletion rates. The local deletion rate also has a great effect, with a higher odds of deletion when previous tokens of /s/ have been deleted.

	Base		With Usage Variables	
Predictor	**p**	**Odds Ratio**	**p**	**Odds Ratio**
Dialect (Ref: 'coastal')				
interior	0.000	0.44	0.002	0.57
Male	0.017	1.54	0.051	1.47
Speech Rate	0.035	1.12	0.083	1.11
Following Segment Type (Ref: 'Voiceless Stop')				
Glide	0.579	1.52	0.501	1.71
Liquid	0.000	4.44	0.000	5.43
Nasal	0.000	3.91	0.000	3.83
Pause	0.025	1.74	0.012	1.98
Voiced stop	0.001	3.21	0.001	3.57
Voiceless fricative	0.473	1.42	0.364	1.61
Vowel	0.000	0.33	0.000	0.29
Following [d]	0.014	2.75	0.013	2.97
Word Frequency - CallHome			0.039	1.16
Conditional Trigram Probability - Two Preceding Words			0.050	0.90
Local deletion rate			0.000	9.42

Pearson Goodness of fit:
p = 0.216 p = 0.682

Table 77. Binary logistic regression for deletion data for the 1st person plural morpheme -*mos*.

6.5.2 /s/ duration in -mos

Table 78 shows the results from the multiple regression of /s/ duration for the morpheme –*mos*. In the model with usage-based variables, the *word frequency*, the *bigram probability of the word with the following word*, the *reverse conditional bigram probability of the word based on the following word*, the *conditional trigram probability of the word given the two preceding words*, and the *local deletion rate* are significant. In contrast with the model for the full data set, but similar to the model for function words, words with higher frequency are predicted to have longer /s/ duration than less frequent words. The effect of the other measures of word probability are similar to the effect in other data sets; higher bigram frequency with the following word and higher reverse conditional bigram probability predict shorter /s/ duration, while higher conditional trigram probability based on the two preceding words predicts longer /s/ duration.

	Base		**With Usage Variables**	
Predictor	p	Coefficient	p	Coefficient
Dialect: interior	0.001	-9.01	0.000	-10.17
Male	0.000	-9.63	0.006	-7.48
Speech Rate	0.000	-4.73	0.000	-5.12
Following Segment (Ref: Voiceless stop/fricative, voiced stop, glide)				
Liquid	0.068	12.70	0.163	9.47
Nasal	0.000	28.31	0.001	26.16
Pause	0.000	26.11	0.000	18.91
Vowel	0.000	17.56	0.000	19.26
SyllableCount	0.070	2.65	0.175	2.91
Word Frequency - Radio			0.005	4.85
Bigram Frequency - Following Word			0.000	-6.25
Reverse Conditional Bigram Probability - Following Word			0.046	-2.38
Conditional Trigram Probability - 2 Preceding Words			0.000	2.71
Local deletion rate			0.052	-8.44
R^2	14.1%		18.1%	

Table 78. Multiple regression for /s/ duration data for the 1st person plural morpheme -*mos*.

6.5.3 Spectral COG

Five of the usage-based variables contribute significantly to the model of the multiple regression for the spectral COG for the morpheme –*mos*, as shown in Table 79: the *word frequency*, the *bigram probability of the word with the following word*, the *forward conditional bigram probability of the following word*, the *conditional trigram probability of the word given the two preceding words*, and the *local deletion rate*. The estimated coefficients for word frequency and the bigram frequency with the following word are the opposite of the coefficients for /s/ duration for the tokens of the morpheme –

mos. Words that are more frequent are predicted to have lower spectral COG, while words in bigrams that are more frequent are predicted to have higher spectral COG. The morpheme *–mos* is the only data set having the forward conditional bigram probability significant, with a higher transitional probability predicting lower spectral COG.

	Base		**With Usage Variables**	
Predictor	p	Coefficient	p	Coefficient
Dialect: interior	0.000	749.80	0.000	717.65
Male	0.000	-396.21	0.000	-414.01
Following Segment (Ref: Voiceless stop/fricative, voiced stop, liquid, nasal)				
Glide	0.018	-708.70	0.025	-647.40
Pause	0.046	199.07	0.227	120.43
Vowel	0.000	333.58	0.000	304.12
Following velar	0.000	295.24	0.002	249.72
SyllableCount	0.001	97.96	0.102	68.82
Following Stress	0.041	127.80	0.026	136.79
Word Frequency - Average			0.009	-126.73
Bigram Frequency - Following Word			0.001	183.50
Forward Conditional Bigram Probability - Following Word			0.024	-141.43
Conditional Trigram Probability - 2 Preceding Words			0.029	32.98
Local deletion rate			0.000	-483.02
R^2	23.4%		28.8%	

Table 79. Multiple regression for spectral COG for the 1st person plural morpheme *-mos*.

6.6 *Word-internal syllable-final /s/*

Word-internal syllable-final /s/ is an important subset to evaluate separately because it may behave differently than word-final /s/, particularly with respect to word probability based on the following word. As described in the literature and confirmed in Table 59 above, word-internal syllable-final /s/ is subject to lower deletion rates than word-final /s/.

6.6.1 Deletion in word-internal /s/

Table 80 shows the results of the binary multiple logistic regression for word-internal /s/. The constraints on word-internal /s/ are mostly the same as those for word-final /s/. The main difference is the effect of the following segment; for word-internal /s/, only a following liquid or a following [d] is significantly different from all of the other segments. The effect of a following liquid is just the opposite for word-internal /s/ as for word-final /s/ (only about half as likely to delete an /s/ before a liquid as compared to other segments word-internally, but more likely to delete an /s/ before a liquid as compared to most of the other segment types word-finally).

In the model including usage-based variables, *word frequency, bigram mutual information with the following word, conditional trigram probability based on the 2 preceding words, repetition number,* and *local deletion rate* are significant. The effect of word frequency and bigram mutual information are in the expected direction, with higher frequency words and higher mutual information predicting higher deletion rates. The effect of the conditional trigram probability based on the two preceding words is similar to the effect in other data subsets, with higher probability predicting lower deletion rates. In this data set, higher repetition number leads to higher deletion rates, as would be expected (but in contrast to the other data subsets where the models actually predict lower deletion rates for higher repetition numbers).

	Base		With Usage Variables	
Predictor	**p**	**Odds Ratio**	**p**	**Odds Ratio**
Dialect (Ref: 'coastal')				
interior	0.000	0.20	0.000	0.26
Male	0.000	1.21	0.004	1.17
Speech Rate	0.000	1.25	0.000	1.19
Preceding Vowel (Ref: 'a/e')				
i/w/y	0.001	0.81	0.224	0.92
o	0.000	0.58	0.139	0.80
u	0.130	1.17	0.057	1.24
Following Segment (Ref: All others)				
Liquid	0.000	0.44	0.000	0.43
[d]	0.000	1.67	0.000	1.69
Following velar	0.000	0.55	0.000	0.69
SyllableCount	0.000	0.83	0.301	1.04
Following Stress (Ref: 'stressed/phrase final')				
unstressed/ monosyllabic	0.000	1.56	0.000	1.66
Stressed syllable	0.001	0.76	0.000	0.71
Word Frequency - Radio			0.000	1.41
Bigram Mutual Information - Following Word			0.000	1.08
Conditional Trigram Probability - 2 Preceding Words			0.000	0.89
Repetition number			0.032	1.10
Local deletion rate			0.000	15.26

Pearson Goodness of fit:

p = 0.000 p = 0.863

Table 80. Binary logistic regression for deletion data for word-internal /s/.

6.6.2 /s/ duration in word-internal -/s/

Table 78 shows the results from the multiple regression of /s/ duration for word-internal /s/. In the model with usage-based variables, *word frequency*, *bigram mutual information with the following word, conditional trigram probability of the word given the two preceding words, trigram frequency of the word with the two preceding words,* and *local deletion rate* are significant. The effect of word frequency, bigram mutual information, conditional trigram probability, and local deletion rate are all consistent with the effect of these independent variables in other data subsets. This is the only data set where the trigram frequency based on the two preceding words is also significant, with higher frequency trigrams having shorter durations.

	Base		**With Usage Variables**	
Predictor	p	Coefficient	p	Coefficient
Dialect: interior	0.000	8.86	0.000	7.77
Male	0.000	-4.82	0.000	-4.50
Speech Rate	0.000	-6.37	0.000	-5.95
Preceding Vowel (Ref: all others)				
a/e	0.000	-2.87	0.191	-0.65
Following Segment (Ref: Voiceless stop or fricative)				
Liquid	0.000	15.12	0.000	15.70
Nasal	0.000	6.06	0.000	6.42
Voiced stop	0.000	-4.36	0.001	-4.01
Following labial	0.000	5.97	0.000	3.91
Following velar	0.000	12.79	0.000	9.87
SyllableCount	0.000	1.37	0.096	-0.47
Following Stress	0.000	9.03	0.000	10.54
Stressed syllable	0.000	4.99	0.000	5.95
Word Frequency - Prior Probability			0.000	-4.14
Bigram Mutual Info - Following Word			0.000	-2.41
Conditional Trigram Probability - 2 Preceding Words			0.000	1.26
Trigram Frequency - Two Preceding Words			0.000	-1.71
Local deletion rate			0.000	-9.29
R^2	16.3%		21.3%	

Table 81. Multiple regression for /s/ duration data for word-internal /s/.

6.6.3 Spectral COG in word-internal -/s/

Three of the usage-based variables contribute significantly to the model of the multiple regression for the spectral COG for word-internal /s/, as shown in Table 82: the *word frequency*, the *conditional trigram probability of the word given the preceding and following words*, and the *local deletion rate*. Higher word frequencies are predicted to have lower spectral COG, but words that are more probable given the preceding and following words are predicted to have higher spectral COG.

	Base		**With Usage Variables**	
Predictor	p	Coefficient	p	Coefficient
Dialect: interior	0.000	1055.86	0.000	1005.18
Male	0.000	-245.32	0.000	-234.46
Speech Rate	0.001	-20.31	0.127	-9.66
Preceding Vowel (Ref: 'o/u')				
a/e	0.000	191.91	0.000	199.72
i/w/y	0.000	286.67	0.000	276.00
Following Segment (Ref: Voiceless stop or fricative)				
Liquid	0.000	-307.47	0.000	-301.55
Nasal	0.011	-91.65	0.004	-104.67
Voiced stop	0.000	-290.00	0.000	-260.81
Following velar	0.000	145.72	0.000	114.18
Following Stress	0.017	33.68	0.003	41.37
Word Frequency - Average			0.002	-22.50
Conditional Trigram Probability - Preceding & Following Words			0.037	27.86
Local deletion rate			0.000	-640.56
R^2	27.6%		30.6%	

Table 82. Multiple regression for spectral COG for the word-internal -/s/.

7 Discussion

7.1 Word frequency

Word frequency significantly contributes to the model of /s/ lenition in most of the analyses performed, as shown in Table 83. In this table, the blocks that are gray indicate no significant effect of word frequency. The symbol '+' indicates more advanced lenition (higher deletion rate, shorter duration, lower spectral COG) with higher word frequency. The symbol '-' indicates less advanced lenition. When the word frequency is significant, it almost always contributes to lenition in the expected way: higher word frequency implies higher lenition. This effect is strongest in the categorical deletion variable, where a word with ten times greater frequency is predicted to have a deletion rate 9-19% higher in word-final /s/ for most of the data subsets evaluated. The effect of word frequency is even stronger for word-internal /s/, with ten times greater frequency predicting 41% more deletion.

Word frequency effects are also gradient phonetically; higher word frequency causes shorter /s/ duration and lower spectral COG. For a ten-fold increase in word frequency, on average the /s/ is about 3 msec shorter. While statistically significant, this only represents a change of approximately 4% of the total /s/ duration. The impact of word frequency on spectral COG is slight (a predicted change of about –15 Hz for a ten-fold increase in word frequency), indicating that there is some pressure for higher frequency words to be pronounced with weaker articulation.

	All	Content	Function	Pre-Consonant	Pre-Vowel	Plural	-mos	Internal
Deletion	+	+		+		+	+	+
Duration	+	+	-	+		+	-	+
Spectral COG	+			+		+	+	+

Table 83. Direction of effect of higher word frequency on /s/ deletion, duration, and spectral COG in each of the subsets tested.

Higher word frequency predicts more advanced lenition in nearly all of the subsets tested, which is consistent with a usage-based model of phonology where more frequent tokens undergo lenition at a faster rate because they have more opportunities to undergo lenition and have a higher level of automation. There are two main categories where word frequency is not significant: word-final /s/ followed by a vowel and function words. In both of these categories, the result that word frequency has no effect is consistent with a usage-based theory using exemplars.

In the case of word-final /s/ followed by a vowel, the conditioning environment for lenition is not present; instead, lenition occurs in this context due to regularization. As Phillips (1984:336) describes, “Physiologically motivated sound changes affect the most frequent words first; other sound changes affect the least frequent words first.” While lenition of /s/ is physiologically motivated before a consonant, it is not physiologically motivated before a vowel. Word-final /s/ is being lost even when the word is followed by a vowel by analogy, in order to avoid allomorphy (Lipski 1986b, Terrell 1981). The higher frequency words may more robust against such changes due to regularization because they have greater familiarity. In addition, as Bybee (2001a:143) notes, words that are “of very high frequency” may have “lexicalized alternations for different contexts", meaning that some high frequency words actually have two pronunciations, one before a vowel (with little lenition) and a separate one before a consonant (with more lenition). It should be noted that this effect is not large in the current data. If the high frequency words had two clear alternations, one before a consonant and one before a vowel, then there would be a statistically significant effect of frequency for word-final /s/ before vowels, in the opposite direction from the effect of frequency before consonants (that is, higher frequency should predict less deletion when only looking at the pre-vowel subset of data). This is not the case; for the pre-vowel data, word frequency isn't significant at all. Moreover, while some of the very frequent words have extremely low deletion rates when followed by a vowel (ex. *les*, with less than 1% deletion before vowels), other words have deletion rates that are higher (*los* at 6% and *las* at 3%).

As for the lack of word frequency effects among function words, this category may behave differently because it includes many of the most frequent words. Because they occur so frequently, many of them may be stored as multi-word units in conjunction with other words that they co-occur with frequently. If this is the case, then we would expect the bigram frequency to be more important than individual word frequency, and there is in fact a significant effect of bigram frequency with the following word for function words. Interestingly, when considering only the subset of function word tokens with low bigram frequency (that is, bigram frequency lower than the median bigram frequency), the effect of word frequency is still not significant, and the effect of bigram frequency is significant.[3] In their study of vowel reduction in the ten most frequent function words in English, Jurafsky *et al.* didn't include word frequency, presumably because they felt that other measures of word probability would be better measures among words that all have high frequency.

7.2 Other measures of word probability

Table 84 shows the effect of word probability based on the following word. Depending on the data set and the dependent variable under investigation (deletion/duration/spectral COG), one or more of the independent variables in this category are significant: *bigram frequency with the following word, bigram mutual information with the following word* and/or the *reverse bigram probability of the word given the preceding word.*[4] When one or more of these variables is significant, it almost always contributes in the expected way: more frequent bigrams and more probable words are predicted to have more lenition. While there is an effect of word predictability in relation to the following word, it is interesting that in some cases, only the bigram frequency is significant (mostly for categorical deletion in the different subsets), and in

---

[3] In fact, it is not clear that the variable *word frequency* is meaningful for the most frequent words, if they are stored as multi-word units. As an example, there are a total of 3,947 tokens of the word *los* in the Broadcast News corpus. 343 of these are followed by the word *Estados*. If the sequence *los Estados* is stored separately as an unanalyzed whole, then the actual token frequency of the individual word *los* should not include these 343 tokens. Of course, the situation is considerably more complicated, because even if frequent multi-word sequences are stored in memory, they are not completely unanalyzed.

[4] The forward bigram probability of the following word (which Bush 2001 claims is important in /t,d/ palatalization across word boundaries) is only significant in the model of spectral COG for the verbal morpheme *–mos*.

others both the bigram frequency and the bigram mutual information or conditional bigram probability are significant. [5]

As described in the previous section, there is no significant effect of word frequency in the case of word-final /s/ followed by a vowel. This is consistent with a usage-based model of phonology because frequent words can have more than one alternate stored in memory, depending on the context. I expected that word-final /s/ followed by a vowel would also behave differently than word-final /s/ followed by a consonant with respect to the bigram frequency with the following word. In a usage-based model, where frequent word sequences are hypothesized to be stored in separate lexical entries, word-final /s/ in the most frequent bigrams should behave as though it is actually word-internal, so when it is followed by a vowel, there should be much less lenition. However, the data shows that higher bigram frequency favors lenition, even in this environment. Still, the magnitude of the estimated coefficients for bigram frequency in the subset of data followed by a vowel are lower than for the subset of data followed by a consonant, indicating that the effect of bigram frequency is not as strong when followed by a vowel. For example, for duration, word-final /s/ followed by a consonant is predicted to be 4.24 msec shorter when bigram frequency is increased ten-fold, while word-final /s/ followed by a vowel is predicted to be 1.9 msec shorter.

In the case of function words, the other category that does not have an effect of word frequency, there is a strong effect of word probability based on the following word. The strength of this effect, as measured by the odds ratio for deletion and the estimated

---

[5] Note that although these three measures in theory should be independent of each other, they are somewhat related due to the limited size of the corpora used to calculate them.

coefficients for /s/ duration and spectral COG is stronger in function words than in content words. These results are consistent with the hypothesis that the most frequent words are stored in the lexicon as multi-word units. If many of the function words are actually stored with the following word in the lexicon as a single unit, then we would expect a large effect for the bigram frequency with the following word rather than for individual word frequency. In addition, the bigram mutual information is important for function words but not for content words for /s/ duration and spectral COG, indicating that when a function word and the following word are more cohesive (i.e. are more like a single unit), a word-final /s/ is more likely to be deleted.

	All	Content	Function	Pre-Consonant	Pre-Vowel	Plural	-mos	Internal
Deletion	+	+	+	+	+	+		+
Duration	+	+	+	+	+	+	+	+
Spectral COG	+	+	+	+	+	+	+/-	

Table 84. Direction of effect of higher probability based on following word (bigram frequency, bigram mutual information, conditional bigram probability) in each of the subsets tested.

The effect of higher probability based on the preceding word or two preceding words is shown in Table 85. When the relative frequency or conditional probability based on the preceding word is significant, it usually contributes in the opposite way than the relative frequency or conditional probability based on the following word; higher probability words are less likely to undergo lenition. Note that this variable is significant in conjunction with other significant usage-based variables, and it does not imply that words that are more predictable overall undergo less lenition. The variables for word frequency and word probability based on the following word ensure that words that are absolutely more predictable do favor lenition. Furthermore, in general, the magnitude of the odds ratios and estimated coefficients are lower for measures of word predictability involving the preceding word than those involving the following word, meaning that the overall effect of the preceding word is lower than the effect of the following word.

	All	Content	Function	Pre-Consonant	Pre-Vowel	Plural	-mos	Internal
Deletion				-	-	-	-	-
Duration	+/-	-	-	+/-	-		-	+/-
Spectral COG			-	-	-	-	-	-

Table 85. Direction of effect of higher probability based on preceding word(s) (bigram/trigram frequency, conditional bigram/trigram probability) in each of the subsets tested.

There are two main ways that the syllable-final /s/ data in the present study differs from that of Jurafsky *et al*. (2001). First, their data supports their Probabilistic Reduction Hypothesis, which claims that all measures of word probability should contribute to lenition, with an increase in any measure of word probability causing a greater extent of reduction. This is not the case for syllable-final /s/. While word frequency and probability based on the following word do contribute in the expected direction, probability based on the preceding word contributes in just the opposite direction.

Second, Jurafsky *et al.* find an effect of conditional probability in function words (evaluating vowel reduction and shortening in function words), but the effect in content words (evaluating /t,d/ shortening and deletion) is much weaker. For the categorical variable of /t,d/ deletion, word frequency is significant but no measures of conditional probability are. In the present study, I find similar effects of conditional probability for both function words and content words.

7.3 *Repetition*

The effect of repetition number is shown in Table 86. In theory, "old" words should be more predictable and therefore phonetically shorter than "new" words, meaning that higher word repetition should mean more advanced lenition. However, for the models in which repetition is significant, most of them find just the opposite. In their analysis, Gregory *et al*. found an effect of word repetition for /t,d/ duration, with higher repetition predicting shorter duration. They did not find any effect of word repetition for categorical /t,d/ deletion. The results for syllable-final /s/ do not agree with their results. Repetition is significant in the case of categorical deletion, with higher repetition predicting lower deletion rates. However this variable is not significant in the case of /s/ duration, and it has mixed effects for spectral COG where it is significant.

It is not clear why the repetition number of the word has such an effect. It may be due to the large number of words that are only spoken once (81% of tokens of syllable-final /s/ are considered to be in the first use).

	All	Content	Function	Pre-Consonant	Pre-Vowel	Plural	-mos	Internal
Deletion	-	-	-	-		-		+
Duration			-					
Spectral COG		+	-	-	+			

Table 86. Direction of effect of higher word repetition number in each of the subsets tested.

7.4 *Phonetic context proportion*

As shown in Table 87, the effect of the consonant proportion for individual words is relatively weak. In many of the analyses performed, the phonetic context proportion is not significant. Where it is significant, with the exception of the function word data set, words that have a higher proportion of pre-consonant context have higher rates of lenition.

While I have evaluated this variable in all of the data sets tested, the result that it is significant in the full data set must be approached with caution due to the fact that there is a relationship between the actual phonetic environment of a particular token and the proportion of each phonetic context for the word in the corpora used to generate the statistics. The average consonant proportion for tokens in word-final, pre-consonant position is 0.65, and it is only 0.44 for tokens in word-final, pre-vowel position. Therefore, the most valid way to evaluate this variable is to look at the subsets of tokens that are followed by a consonant or followed by a vowel. For word-final /s/ followed by a consonant, there is an effect of consonant proportion on deletion and on spectral COG, with word-final /s/ in words that are more often followed by a consonant having more advanced lenition. For word-final /s/ followed by a vowel, there is an effect of consonant proportion only on /s/ duration, also with words that appear more often followed by a consonant having more advanced lenition. The fact that the models predict more advanced lenition for tokens of syllable-final /s/ in words with higher consonant proportion, even after taking into consideration the actual phonetic context of the token, suggests that this is part of the speaker's linguistic knowledge. These results give some support for Bybee's claim that the way that a particular word is used affects its lexical representation.

Given this evidence of more advanced lenition in words that appear more often before a consonant, it is somewhat surprising that word-internal /s/ followed by a consonant generally has less advanced lenition than word-final /s/ followed by a consonant. Word-internal /s/ always appears in the conditioning environment for the change while word-final /s/ appears in an alternating environment, depending on the following word. In an exemplar-based model, the lexical entry is updated each time a new token is encountered. Since a word-internal /s/ is always in the conditioning environment, each update to the lexical representation should favor lenition. Contrary to Bybee's (2001a: 148) claim that "sound changes taking place in uniform environments and within morphemes seem to occur at an accelerated pace of change," lenition of syllable-final /s/ is slower word-internally.

	All	Content	Function	Pre-Consonant	Pre-Vowel	Plural	-mos
Deletion	+		-	+		+	
Duration	+	+			+	+	
Spectral COG				+			

Table 87. Direction of effect of consonant proportion in each of the subsets tested.

7.5 Local deletion rate

The local deletion rate by far has the greatest impact of the usage-based factors evaluated. It is significant for each of the subsets tested, for all three measures of lenition: deletion, duration, and spectral COG, as shown in Table 88. Most of the binary logistic regression models predict about 10 times more deletion if the local deletion rate is 100% as compared to when the local deletion rate is 0%. While local deletion rate is significant in all of the subsets modeled, it has an especially large impact on deletion when the word is followed by a vowel (where /s/ is 36 times more likely to be deleted when the local deletion rate is 100% than when it is 0%). The local deletion rate is also significant on the gradient processes; the duration is predicted to be approximately 10 msec shorter, and the spectral COG is predicted to be about 500 Hz lower when the local deletion rate is 100% in comparison with 0%.

This type of perseverance effect is consistent with an exemplar model that includes an auditory buffer as described in section 5.2.6. Even when other factors related to the speaker (speaker dialect and sex) and how carefully the utterance is pronounced (the speech rate) are taken into effect, the local deletion rate adds significantly to the model. This implies that the speaker's phonetic target is slowly changing over time.

	All	Content	Function	Pre-Consonant	Pre-Vowel	Plural	-mos	Internal
Deletion	+	+	+	+	+	+	+	+
Duration	+	+	+	+	+	+	+	+
Spectral COG	+	+	+	+	+	+	+	+

Table 88. Direction of effect of higher local deletion rate on /s/ deletion, duration, and Spectral COG in each of the subsets tested. '+' indicates more advanced lenition.

7.6 Conclusion

As described above, the main premise of a usage-based model of language is that the structure of language is primarily shaped by how it is used, and that linguistic knowledge includes knowledge about variation. The data in the present study provide evidence that several of the variables associated with a usage-based model of variation contribute to lenition in syllable-final /s/. Word frequency, other measures of word probability, the phonetic context that words appear in most frequently, and the local deletion rate all influence the extent of lenition. While these usage-based variables are not typically considered to be part of a speaker's linguistic competence in a generative framework, the fact that they contribute significantly to the regression models indicates that they form part of the speaker's knowledge. Furthermore, these variables contribute both to categorical deletion of syllable-final /s/ and gradient lenition as measured by /s/ duration and spectral COG, and the effects of the variables are generally very similar for all three measures of lenition. This indicates that the variability in syllable-final /s/ lenition is not simply the result of phonetic implementation. Any linguistic theory must therefore be able to incorporate both the categorical and the gradient effects of these usage-based variables.

It should be noted that while they are statistically significant, the effect of most of the usage-based variables is small in comparison with many of the non-usage based variables that have been explored in other studies, particularly the following segment, the speaker's dialect and sex, and the speech rate.

The results of this study have several implications on lexical representations. In an exemplar model as advocated by Bybee (2001a) and Pierrehumbert (2001), the 'rule' of syllable-final /s/ lenition is accounted for within the lexicon. The effects of most of the variables are compatible with this type of model. The frequency effects for individual words found in this study are consistent with Pierrehumbert's model in which more frequent words undergo more advanced lenition because they have more opportunities to be affected. The relative frequency effects found for the plural marker as well as for bigrams with the following word both can be derived from the exemplar model, and in an exemplar framework, they provide support for Bybee's claim that the lexicon must contain units larger than morphemes. The strongest evidence that the lexical entry contains fine-grained phonetic detail comes from the result that the phonetic context in which particular words occur has an impact on the word's pronunciation. Words with higher consonant proportion have more advanced lenition, even after taking into consideration the actual phonetic context, an effect which cannot be easily derived if lenition is considered to be either a rule or simply the result of phonetic implementation.

While the effect of most of the usage-based variables that are significant in the present study is consistent with a theory of exemplars, conditional probabilities are somewhat problematic; whereas they have been shown to play a role in syllable-final /s/ lenition, they cannot be incorporated into the lexicon in a straightforward way. To account for effects such as these, Jurafsky *et al.* (2001) claim that speakers have knowledge of probabilistic relations between words, and that reduction occurs not within the lexicon but "at production processing levels" (p. 246). This type of model where the extent of lenition is calculated rather than stored in the lexicon is also compatible with most of the usage-based variables considered in the present study, and it also has the advantage of being able to handle other types of word probability that are out of the scope of the present study such as syntactic and semantic considerations. However, it is problematic to incorporate the effect of a word's most frequent phonetic context within this type of framework. The data from syllable-final /s/ lenition therefore suggest that the extent of lenition may be determined both within the lexical entry and by some kind of processing during production.

References

Alba, Orlando. 1982. Función del acento en el proceso de elisión de la /s/ en el español de República Dominica. In Orlando Alba (ed.), El español de Caribe: ponencias del VI Simposio de Diaectología. Santiago de los Caballeros, 15-26.

Amastae, Jon. 1989. The intersection of s-aspiration/deletion and spirantization in Honduran Spanish. *Language variation and change* 1, 169-183.

Agresti, Alan. 1990. Categorial Data Analysis. New York: John Wiley & Sons.

Barrutia, Richard and Tracy D. Terrell. 1982. *Fonetica y fonologia españolas*. New York: John Wiley & Sons.

Browman, Catherine P. and Louis Goldstein. 1992. Articulatory Phonology: An Overview. *Phonetica* 49, 155-180.

Brown, Esther L., and Rena Torres Cacoullos. 2003. Spanish /s/: A different story from beginning (initial) to end (final). In Rafael Núñez-Cedeño, Luis López, and Richard Cameron (eds.) *A Romance perspective on language knowledge and use*. Amsterdam: John Benjamins, 21-38.

Bybee, Joan L. 1999. Usage-based Phonology. In M. Darnell, E. Moravcsik, F. Newmeyer, M. Noonan and K. Wheatley (eds.) Functionalism and formalism in Linguistics, Volume I: General Papers. Amsterdam: John Benjamins, 211-242.

Bybee, Joan L. 2000. The phonology of the lexicon: evidence from lexical diffusion. In M. Barlow and S. Kemmer (eds.) *Usage-based models of language*. Stanford, CA: CSLI, 65-85.

Bybee, Joan. 2001a. *Phonology and Language Use*. Cambridge: Cambridge University Press.

Bybee, Joan. 2001b. Frequency effects on French liaison. In Joan Bybee & Paul Hopper (eds.) *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins, 337-359.

Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14:261-290.

Cameron, Richard. 1993. 'Ambiguous agreement, functional compensation and nonspecific tú in the Spanish of San Juan, Puerto Rico, and Madrid, Spain.' *Language Variation and Change* 5, 305-334.

Canfield, D. L. 1981. *Spanish pronunciation in the Americas.* Chicago: University of Chicago Press, Chicago. (Reprint of 1961 edition)

Cedergren, Henrietta J. 1978. En torno a la variación de la s final de sílaba en Panamá: Analisis cuantitativo. In Humberto López Morales, ed. *Corrientes actuales en la dialectología del Caribe hispánico*. Puerto Rico: Editorial Universitaria, Universidad de Puerto Rico, 37-49.

Cepeda, Gladys. 1995. Retention and Deletion of Word-Final /s/ in Valdivian Spanish (Chile). *Hispanic Linguistics* 6-7: 329-353.

Christensen, Ronald. 1990. Log-linear models and logistic regression. Second edition. New York: Springer-Verlag.

Cohen, Jacob. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20:37-46.

Cohen, Jacob. 1968. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70:213-220.

Cole, Ronald, Beatrice Oshika, Mike Noel, Terri Lander and Mark Fanty. 1994. Labeler Agreement in Phonetic Labeling of Continuous Speech. Proceedings of International Conference on Spoken Language Processing. Yokohama, Japan, 2131-2134.

Garrett, Susan, T. Morton, and C. McLemore. 1997. *LDC Spanish Lexicon*. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Fenk-Oczlon, Gertraud. 2001. Familiarity, Information Flow, and Linguistic Form. In Joan Bybee and Paul Hopper (eds*.) Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins Publishing Company, 431-448.

Fontanella de Weinberg, M. 1974. Aspectos sociolingüísticos del uso de -s en el español bonaerense. *Orbis* 23:85-98.

Fosler-Lussier, Eric and Nelson Morgan. 1999. Effects of speaking rate and word frequency on pronunciations in conversational speech. *Speech Communication* 29, 137-158.

Fosler-Lussier, J. Eric. 1999. Dynamic Pronunciation Models for Automatic Speech Recognition. PhD. Dissertation, University of California, Berkeley.

Godfrey, J., E. Holliman, and J. McDaniel. 1992. SWITCHBOARD: Telephone speech corpus for research and development. *Proceedings of ICASSP-92*. San Francisco, 517-520.

Gregory, Michelle, William D. Raymond, Alan Bell, Eric Fosler-Lussier, and Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. *Chicago Linguistic Society* 35, 151-66.

Harrington, Jonathan and Cassidy. Steve. 1999. Techniques in Speech Acoustics. Kluwer Academic Publishers.

Hochberg, Judith G. 1986. "/S/ deletion and pronoun usage in Puerto Rican Spanish". In David Sankoff (ed.), Diversity and Diachrony. Amsterdam: John Benjamins,. 199-210.

Honsa, Vladimir. 1965. The Phonemic Systems of Argentinean Spanish. *Hispania* 48, 275-83.

Hoole, Philip, Ziegler, Wolfram, Hartman, Erich, and Hardcastle, William. 1989. Parallel electropalatographic and acoustic measures of fricatives. *Clinical Linguistics & Phonetics* 3(1), 59-69.

Hooper, Joan. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In W. Christie, ed., *Current Pogress in Historical Linguistics*. North Holland, Amsterdam, 96-105.

Hundley, James E. 1987. Functional constraints on plural marker deletion in Peruvian Spanish. *Hispania* 70:891-895.

Hulst, Harry van der. 2000. "Modularity and Modality in Phonology." In Burton-Roberts et al., eds. *Phonological Knowledge. Conceptual and Empirical Issues.* Oxford: Oxford University Press, 207-243.

Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In Johnson & Mullennix (eds.) *Talker Variability in Speech Processing*. San Diego: Academic Press, 145-165.

Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William Raymond. 2001. Probabalistic relations between words: Evidence from reduction in lexical production. In Joan Bybee and Paul Hopper (eds.*) Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins Publishing Company, 229-253.

Kerswill, Paul and Susan Wright. 1990. The validity of phonetic transcription: Limitations of a socio-linguistic research tool. *Language variation and change*, 2:255-275.

Kingsbury, Paul. 1997. *Hub4 supplement to the Spanish Lexicon*. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Kirchner, Robert. 1998. *An Effort-Based Approach to Consonant Lenition*. UCLA Ph.D Dissertation.

Labov, William. 1994. *Principles of Linguistic Change: internal factors*. Blackwell, Oxford.

Lander, Terri, Beatrice Oshika, Ronald A. Cole, and Mark Fanty. 1995. Multi-language Speech Database: Creation and Phonetic Labeling Agreement. *Proceedings of the International Congress of Phonetic Science,* Stockholm, Sweden.

Langacker, R. W. 1987. *Foundations of cognitive grammar: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.

Langacker, Ronald. 2000. A Dynamic Usage-Based Model. In Michael Barlow and Suzanne Kemmer (eds.) *Usage-Based Models of Language*. Stanford: CSLI, 1-63.

Lapesa, Rafael. 1980. *Historia de la lengua española*. Madrid.

Lavoie, Lisa M. 2000. *Consonant Strength. Phonological Patterns and Phonetic Manifestations*. New York: Garland Publishing, 2001.

Lipski, John M. 1982. Central American Spanish in the United States: the case of El Salvador. Presented at the 3rd annual conference *El Español en los Estados Unidos*, Indiana Univerisity.

Lipski, John M. 1983a. La norma culta y la norma radiofonica: /s/ y /n/ en espanol. *Language problems and language planning* 7:239-262.

Lipski, John M. 1983b. Reducción de /s/ en el español de Honduras. *Nueva Revista de filología hispánica* XXXII:273-288.

Lipski, John M. 1984a. On the weakening of /s/ in Latin American Spanish. *Zeitschrift fur Dialektologie und Linguistik* 51, 31-43.

Lipski, John M. 1984b. /s/ in the Spanish of Nicaragua. *Orbis*, 33:171-181.

Lipski, John M. 1985. Spanish in United States broadcasting. In L. Elias-Olivares, E.A. Leone, R. Cisneros & J.R. Guttierez (eds.). *Spanish language use and public life in the United States.* Mouton: Berlin, 217-233.

Lipski, John M. 1986a. Instability and reduction of /s/ in the Spanish of Honduras. *Revista Canadiense de Estudios Hispánicos*. XI(1):27-47.

Lipski, John M. 1986b. Reduction of Spanish Word-Final /s/ and /n/. *Canadian Journal of Linguistics.* 31(2), 139-156.

Lipski, John M. 1994. Tracing Mexican Spanish /s/: A Cross-Section of History. Language Problems and Language Planning; 1994, 18(3):223-241.

Lipski, John M. 1999. The Many Faces of Spanish /s/-Weakening: (Re)Alignment and Ambisyllabicity. In Javier Gutierrez-Rexach and Fernando Martinez-Gil (eds.) *Advances in Hispanic Linguistics: Papers from the 2nd Hispanic Linguistics Symposium, I-II*. Somerville, MA: Cascadilla, 198-213.

Lunneborg, Clifford E. 1994. Modeling Experimental and Observational Data. Belmont, CA: Duxbury (Wadsworth) Press.

MacWinney, B. 1995. *The CHILDES Project: Tools for Analyzing Talk.* Hillsdale, NJ: Lawrence Erlbaum.

Manrique, Ana Maria Borzone de, and Maria Ignacia Massone. 1981. Acoustic analysis and perception of Spanish fricative consonants. *Journal of the Acoustical Society of America* 69:1145-1153.

Morris, Richard E. 2000. Constraint Interaction in Spanish /s/-Aspiration: Three Peninsular Varieties. In: Héctor Campos et al (eds.) *Hispanic Linguistics at the Turn of the Millennium: Papers from the 3rd Hispanic Linguistics Symposium*. Cascadilla Press, 14-30.

Navarro Tomás, Tomás. 1966. *El Español en Puerto Rico; contribución a la geografía lingüística Hispanoamericana*. Rio Piedras: Universidad de Puerto Rico.

Nosofsky, R. M. 1988. Similarity, frequency and category representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14:54-65.

Núnez Cedeño, Rafael. 1980. *La fonología moderna y el español de Santo Domingo.* Santo Domingo: Editorial Taller.

Phillips, Betty S. 1980. Old English *an ~ on*: A new appraisal. *Journal of English Linguistics* 14, 20-23.

Phillips, Betty S. 1984. Word frequency and the actuation of sound change. *Language* 60:320-42.

Phillips, Betty S. 2001. Lexical diffusion, lexical frequency, and lexical analysis. In Joan Bybee and Paul Hopper (eds.) *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins Publishing Company, 124-136.

Pierrehumbert, Janet. 1994. Knowledge of variation. In Beals *et al*. (eds), *CLS 30, Volume 2: Papers from the Parasession on Variation and Linguist Theory*. Chicago Linguistics Society, 232-56.

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (eds.) *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins Publishing Company, 137-157.

Pierrehumbert, Janet. 2003. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and speech.* 46(2-3), 115-154.

Poplack, Shana. 1980a. Deletion and Disambiguation in Puerto Rican Spanish. *Language*. 56(2):371-385.

Poplack, Shana. 1980b. The notion of the plural in Puerto Rican Spanish: Competing constraints on /s/ deletion. In *Locating Language in Time and Space*. W. Labov (ed.). Academic Press, New York, 55-68.

Poplack, Shana. 1981. Mortal phonemes as plural morphemes. In D. Sankoff & H. Cedergren (eds.). *Variation Omnibus*. Linguistic Research, Edmonton, Alberta, 59-72.

Resnick, Melvyn C. and Robert M. Hammond. 1975. The status of quality and length in Spanish vowels. *Linguistics* 156:79-88.

Reynolds, W. 1994. *Variation and phonological theory*. PhD. Dissertation, University of Pennsylvania.

Scherre, Maria Marta and Anthony J. Naro. 1992. The serial effect on internal and external variables. *Language Variation and Change*. 4, 1-13.

Shadle, C. H. & Mair, S. J. 1996. Quantifying spectral characteristics of fricatives. *Proceedings, 4th International Conference on Spoken Language Processing*. Philadelphia, 1521-1524.

Tabain, Marija. 2001. Variability in Fricative Production and Spectra: Implications for the Hyper- and Hypo- and Quantal Theories of Speech Production. *Language and Speech* 44(1), 57-94.

Terrell, Tracy. 1977. Constraints on the aspiration and deletion of final /s/ in Cuban and Puerto Rican Spanish. *The Bilingual Review,* 4, 35-51.

Terrell, Tracy. 1978a. La aspiracion y elision de /s/ en el español porteño. *Anuario de Letras* 16:41-66.

Terrell, Tracy. 1978b. Sobre la aspiración y elision de /s/ implosive y final en el español de Puerto Rico. *Nueva Revista de Filología Hispánica,* 27, 24-38.

Terrell, Tracy. 1979. Final /s/ in Cuban Spanish. *Hispania*, 62, 599-612.

Terrell, Tracy. 1981. Diachronic representation by dialect comparison of variable constraints: s-aspiration and deletion in Spanish. In David Sankoff & Henrietta Cedergren (eds.), *Variation Omnibus*. Carbondale: Linguistic Research, 115-124.

Valdivieso, Humberto and Juanita Magaña. 1991. Variacion fonetica de /s/ en el habla espontanea. *Revista de linguistica teorica y aplicad*a 29: 97-113.

Wester, Mirjam, Kessens, Judith M., Cucchiarini, Catia, and Strik, Helmer. 2001. "Obtaining Phonetic Transcriptions: A Comparison between Expert Listeners and a Continuous Speech Recognizer. *Language and Speech* 44(3), 377-403.

Widdison, Kirk A. 1995a. An acoustic and perceptual study of the Spanish sound change s>h. *Rivista di Linguistica* 7, 175-190.

Widdison, Kirk A. 1995b. Two models of Spanish s-aspiration. *Language Sciences*, 17(4), 329-343.

Widdison, Kirk A. 1997. Phonetic explanations for sibilant patterns in Spanish. *Lingua* 102, pp. 253-264.

Wright, Richard. 1997. *Lexical Competition and Reduction in Speech: A Preliminary Report.* Progress Report No. 21 (1996-1997), Indiana University. Available at www.indiana.edu/~srlweb/publication/short219.pdf

Young, Steve, Joop Jansen, Julian Odell, Dave Ollason, and Phil Woodland. 1995. *The HTK Book (for HTK Version 2.0)*. Cambridge University.