Variation and Gradience in a Noisy Harmonic Grammar with Lexically Indexed Constraints: The Case of Spanish –*s* Deletion and Aspiration

Jennifer Griffin

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Masters of Arts in the Department of Linguistics.

Chapel Hill
2011

Approved by:

Elliott Moreton

Jennifer L. Smith

David Mora-Marin

ABSTRACT

JENNIFER GRIFFIN: Variation and Gradience in a Noisy Harmonic Grammar with Lexically Indexed Constraints: The Case of Spanish –*s* Deletion and Aspiration
(Under the direction of Elliott Moreton)

This thesis presents a new way of modeling variation in production and perception within and across lexical categories. By adding lexical indexes to both the input and relevant faithfulness constraints in a Noisy Harmonic Grammar model, I will show that production frequencies be used to predict well-formedness judgments of variable forms. First I show that by using this model, an artificial learner in Praat (Version 5.1.43) can learn the appropriate production frequencies of variants showing –*s* deletion and aspiration in Spanish. In Experiment 1 I show that Puerto Rican Spanish speakers choose sentences with aspirated adjectives as more well-formed than sentences with aspirated nouns. In Experiment 2, participants' perception of ambiguous phonemes along a continuum from [h] to [s] is significantly influenced by the lexical category of the root to which the ambiguous fricative is attached. These results support the predictions made about perception judgments based on variable production frequencies.

Acknowledgements

First I would like to thank my infinitely patient thesis committee: Elliott Moreton, Jen Smith and David Mora-Marin. This project has benefitted significantly from your guidance and I greatly appreciate all of your advice. Thank you for making me a better linguist.

I would also like to thank the Spring 2011 P-Side group. Y'all provided great feedback on my experimental design and helped me explain concepts better (particularly the methods for Experiment 2).

Special thanks to Inmaculada Gomez-Soler and Danilo Loor for acting as my Spanish language consultants. I couldn't have created natural sounding token sentences without you (quite literally, since I spent numerous hours recording both of you reading off the token sentences until they were perfect).

I would also like to thank Christopher Wiesen at the Odum Institute, who performed all of the statistical analysis of the data.

Table of Contents

List of Abbreviations

NHG	Noisy Harmonic Grammar
OT	Optimality Theory

Chapter 1

Introduction

In this paper, I propose a Noisy Harmonic Grammar model (NHG) with weighted lexically indexed constraints that can account for phonological phenomena that apply variably across different lexical categories. In particular, I will examine the case of –*s* aspiration and deletion in Spanish. In Spanish, word-final -*s* is reduced to -*h* or deleted at different frequencies depending on the lexical category of the word (Evanini 2007; Poplack 1980; Cedergren 1973). Under my NHG model, inputs as well as relevant faithfulness constraints will be indexed for lexical categories such as noun and adjective. These constraints will be assigned different weights depending on the lexical category to which they are indexed. I test this model by having a learning software in Praat (Version 5.1.43) learn the grammar based on frequencies of variant forms that exhibit word-final –*s* aspiration or deletion observed in the language. Once the learner has generated the correct frequencies for –*s* aspiration and deletion in the lexical categories of noun and adjective, I will show that by using the Harmonic Values for each variable candidate and Coetzee & Pater's Acceptability Metric (2008), I can account for gradience in grammaticality judgments by native speakers. This allows the same phonological model to account for both production and perception data. Well-formedness judgments were gathered by running two speech-perception experiments on seven native Puerto Rican Spanish speakers whose dialects display –*s* aspiration and deletion. The Puerto Rican dialect shows the same pattern of –*s* aspiration and deletion affecting adjectives more often than nouns (Poplack 1980). This means the relationships in terms of relative weight between the constraints should be the same in Puerto Rican Spanish as in Panama City Spanish, although the value of the weights themselves may change slightly between dialects. Since both dialects show the same general pattern in production, variants of nouns and adjectives should hold the same well-formedness status in both dialects, allowing me to gather perception data from Puerto Rican Spanish speakers that will reflect the well-formedness pattern predicted for both dialects by the Harmonic Values of each variant. Finally, I will discuss linguistic implications of this model and further research opportunities for this model.

Chapter 2

Literature Review

In historical linguistics, it has often been claimed that variation is a stepping stone on the way to sound change (Janson 1983; Labov 1981). Variation occurs as a sound change is transmitted throughout the language through lexical diffusion. In between the initial state and the final state of sound change, variation occurs as the sound change is slowly transmitted throughout the language (Janson 1983; Labov 1981). Janson (1983) states that “synchronic variation is intimately connected with diachronic change,” (p. 18). Janson (1983) goes further to argue that changes in production of variant forms invokes changes in perception of those variants. This connection between variation in production and perception is precisely what this paper aims to provide evidence for. Frequency has been shown to play a crucial role in whether certain words are more likely to show a particular phonological variant (Bybee 2002). For example, English final /t/ or /d/ deletion affects high-frequency words much more than low-frequency words (Bybee 2002). The aspiration and deletion of syllable-final /s/ in many dialects of Spanish shows this same pattern. Therefore, in order to rest solely the effect of a variable such as lexical category on the production and perception of variant forms, it is imperative that the frequency of the tokens is controlled for.

The question of how to model variation in production and perception has been at the core of variation studies for many years. In much of the phonological literature, there persists a claim that speech production and perception are linked in a significant way

(Coetzee & Pater 2008; Boersma & Hayes 2001; Pierrehumbert 2001; Janson 1983; Labov 1981). Many models have been proposed to account for and test this theoretical link between production and perception. According to Coetzee & Pater (2008), a Noisy Harmonic Grammar model is the best model for capturing free variation as well as accounting for well-formedness judgments. Under this model, each constraint is weighted rather than ranked. Each candidate is assigned a Harmonic Value based on the number of violations it has incurred. The Harmonic Value of a candidate is calculated by multiplying candidate's score, which is a negative number based on the number of violations a given input has incurred for a given constraint, for a each constraint by the weight of that constraint. These products are then added together to yield the Harmonic Value for that candidate (Coetzee & Pater 208). This is best illustrated using the example tableau below in (1), in which the weight of each constraint is given below the constraint name.

1)

/books/	**MAX** 2	**DEP** 4	***S-WordFinal** 1	**Harmonic Value**
→books			*	-(1x1) = -1
book	*			-(1x2) = -2
bookes		*	*	-((1x4) + (1x1)) = -5

In the example above, the most harmonic candidate is the one with the greatest Harmonic Value, which in this case is 'books'.

Coetzee & Pater (2008) also provide a way in which Harmonic Values can be used to determine well-formedness. Using their Acceptability Metric, the Harmonic Values of two variable candidates can be used to determine which variant is more well-formed. The equation for the Acceptability Metric is given below.

2)

Accepability(x) = H(x) – H(y)

Where y is the most harmonic candidate for the same input as x.

A candidate which receives a positive Acceptability Value is well-formed while a candidate which receives a negative Acceptability Value is ill-formed (Coetzee & Pater 2008). This NHG model is an improvement over other methods which aim to account for variability and gradience, such as a variable rules model or Stochastic OT, because NHG is able to account for a broader range of gradient phenomena by allowing constraint violations to interact with each other. Variable rules models, which assign probabilities to the likelihood of a phonological rule being implemented based on extra-phonological contextual factors, dominated the literature on variation between the 1960s and 1990s (Pierrehumbert 2001). Boersma & Hayes (2001) show that Stochastic OT can indeed account for gradient well-formedness data regarding ambiguous phoneme perception, as illustrated in their study of English light [l] and dark [ɫ]. Although Stochastic OT can account for production variation as well as some kinds of gradient perception data, the constraints are still ranked in such a way that cannot account for all gradient phenomena such as gang effects (Coetzee & Pater 2008). NHG is also preferred over Stochastic OT in regards to learning algorithms, where Stochastic OT has been shown to be non-convergent (Pater 2008). Given these theoretical comparisons, a NHG was employed in this study.

It has often been seen in the study of phonology that phonological processes do not always apply equally across the lexicon. Specifically, the lexical category of an input can have a large effect over whether or not a phonological phenomenon will occur (Smith 2001; Evanini 2007; Cedergren 1973; Poplack 1980). Jennifer Smith (2001) shows that lexically indexed constraints, particularly noun faithfulness constraints, can help account for observed variation across lexical categories.

But what happens when variation occurs both across and within lexical categories? We must then employ a phonological model that can account for variable output candidates for a single input while accessing the lexical category of the input. Spanish –*s* aspiration and deletion is a perfect case study to test the NHG model with lexically indexed constraints because it shows the variability across lexical categories. As seen in Cedergren's (1973) study of Spanish s-deletion in Panama City Spanish, whether or not a word-final -*s* is deleted or reduced to an -*h* will be made more or less probable by a variety of factors. The style of the speaker, the lexical category of the input, the grammatical status of -*s* and the segment following the -*s* all play a role in determining whether or not an *s* will be deleted or reduced. This is illustrated in the following table from Evanini (2007):

3)

Factor Group	strongly promotes /s/ deletion	mildly promotes /s/ deletion	inhibits /s/ deletion
POS	adjective	noun	determiner
grammatical status of /s/	monomorphemic	plural -/s/ suffix	2nd Sg. -/s/ suffix
following segment	consonant	vowel	pause

Table 7. *Internal grammatical factor groups affecting /s/ deletion.*

(Evanini 2007, page 152)

The general pattern of –*s* aspiration and deletion occurring with different frequencies across lexical categories is seen in many dialects of Spanish, including Puerto Rican, Panamanian and Cuban (Poplack 1980; Morris 2000). Since Cedergren (1973) has provided actual production frequencies for each variant, I will use her data on Panama City Spanish as the basis for my artificial learner in Section 4.

Evanini (2007) proposes a Stochastic OT account for this variation in Spanish –s aspiration and deletion. According to Evanini (2007), effects on variation can be accounted for by simply adding variables to the selection point equation. Evanini accounts for effects from the grammatical category of –s, the following segment and the lexical category of the input in the following equation:

4)

$$\mathit{selectionPoint}_i = \mathit{rankingValue}_i + \mathit{styleSensitivity}_i \times \mathit{Style} + \boldsymbol{posSensitivity}_i \times \mathbf{POS} + \boldsymbol{gramSensitivity}_i \times \boldsymbol{Gram} + \boldsymbol{follsegSensitivity}_i \times \boldsymbol{FollSeg} + \mathit{noise}$$

(Evanini 2007, page 153)

Under this model, the selection point of a particular constraint (indexed here with i) is equal to the sum of its ranking value and its particular style sensitivity (or likelihood to be influenced by style) multiplied by an language-dependent value for style itself. Evanini also adds variables for position, grammatical category and following segment, each with their own values and each constraint with its own sensitivity to each variable. While this model may accurately account for variation across lexical categories, Evanini never addresses how the model might be learned or how it might account for gradience in grammaticality judgments. Evanini also never gives actual examples for possible values for variables such as style or following segment, leaving this selection point equation completely untested and uninformative.

For this study, I propose a hybrid model in which Noisy Harmonic Grammar, like that used by Coetzee & Pater (2008), is combined with constraints which are indexed to lexical category, much like those in Smith (2001). By combining these two models, this new model will have the benefits of being able to account for variation within and across lexical categories as well as perception data. It also has the advantage of being testable, both by running the model through an artificial language learner in Praat (Version 5.1.43) and by testing it against perception data from native Spanish speakers.

Chapter 3

The Model

For this study, I am proposing a Noisy Harmonic Grammar with lexically indexed constraints. Each relevant faithfulness constraint will be co-indexed to a lexical category such as noun or adjective. Each input will also be co-indexed to a lexical category. When an indexed input is given to a grammar, only the non-indexed constraints and the constraints that are indexed for that specific lexical category will be relevant. That is, an indexed constraint only affects which variant is chosen as the output if the input also shares the same index. While learning, the learner must acquire the weights of the constraints.

In a Noisy Harmonic Grammar, the weight of a given constraint changes from evaluation to evaluation (Pater 2008). This allows for variation between outputs because each time the speaker gets an input, the constraint weights will change, thus changing the harmonic values for the variable outputs. Thus, at each evaluation, a different candidate will have the best harmonic value based on the weights of the constraints and the number of violations that candidate has incurred.

For this study, I am choosing a noisy version of Harmonic Grammar so that I can model how a learner would learn a variable grammar. If I were to try to account for variability in a standard Harmonic Grammar, then the variant with the best harmonic value would always be the same because the weights for the constraints would never change. Adding noise to the model ensures that each optimal output (one with –s, one with –h, and one with -∅) will be chosen some of the time.

For this study, I will adopt my constraints from Cedergren's (2007) account of variation in Spanish –*s* aspiration and deletion. Crucially, however, I will add lexical indexes to relevant Faithfulness constraints. In doing so, I get the following constraint set:

5)

1. ***S-SyllFinal –** Assign one * for every syllable-final *s*.
2. **ID-PLACE$_N$** – Assign one * for every segment in an output that has a different place of articulation than its correspondent in an input noun.
3. **ID-PLACE$_A$ –** Assign one * for every segment in the output that has a different place of articulation than its correspondent in an input adjective.
4. **MAX$_N$** – Assign one * for every segment in an input noun that does not have a corresponding segment in the output.
5. **MAX$_A$** – Assign one * for every segment in an input adjective that does not have a corresponding segment in the output.
6. **ID-MANNER –** Assign one * for every segment in the output that has a different manner of articulation than its correspondent in the input.

(Evanini 2007)

The **ID-MANNER** constraints interact with ***S-SyllFinal** in order to create variation between –*s* and –*h*, while the **MAX** constraints interact with ***S-SyllFinal** in order to create variation between –*s* and –Ø. **ID-MANNER** will have a much higher weight than the other constraints in order to rule out any ungrammatical forms such as –*n* as a variant of –*s*. Since –*s* is the least frequent variant of Spanish –*s* aspiration and deletion for both nouns and adjectives, ***S-SyllFinal** does not need to be indexed for any particular lexical category. Likewise, -*n* is never a variant of –*s* in either nouns or adjectives, so **ID-MANNER** does not need to indexed for lexical category. Since **ID-PLACE** and **MAX** produce the variant forms –*h* and –Ø by interacting with ***S-SyllFinal,** and the frequencies for each variant differs depending on the lexical category of the input, these two Faithfulness constraints will be indexed for lexical category.

Perception data from native Spanish speakers will also be accounted for under this model. Using Coetzee & Pater's Acceptability Metric (2008), I will show that the Acceptability Value for each candidate in relation to another will predict how well-formed one variant will seem to native speakers. For Example, let's say ***S-SyllFinal** has a weight of 16 while **ID-PLACE $_N$** has a weight of 15 and **ID-PLACE $_A$** has a weight of 14. We can calculate the Acceptability Values of candidates such as /libros/ and /buenos/ using the tableaus in (6) and (7), in which the weights of each constraint is located underneath the constraint:

6)

/libros$_N$/	***S-SyllFinal** 16	**ID-PLACE$_N$** 15	**ID-PLACE$_A$** 14	**Harmonic Value**
libros	*			-16
libroh		*		-15

7)

/buenos$_A$/	***S-SyllFinal** 16	**ID-PLACE$_N$** 15	**ID-PLACE$_A$** 14	**Harmonic Value**
buenos	*			-16
buenoh			*	-14

Using the Acceptability Metric equation, we can use the Harmonic Values to calculate the Acceptability Values of each candidate.

8)

Acceptability(x) = H(x) – H(y)
Where y is the most harmonic candidate for the same input as x

Acceptability (libros) = -16- (-15) = -1
Acceptability(buenos) = -16 – (-14) = -2
Acceptability (libroh) = -15 – (-16) = 1
Acceptability (buenoh) = -14 – (-16) = 2

Since a violation of a constraint will yield a negative Harmonic Value of a candidate, all Harmonic Values will be negative. Keeping this in mind, Coetzee & Pater (2008) stipulate that all candidates that receive negative Acceptability Values are ill-formed while candidates that receive positive Acceptability Values are well-formed. Thus, the greater the difference in Harmonic Values, the higher the Acceptability Value of a given candidate will be, and the more well-formed that variant will seem to native speakers.

Given the equation for Coetzee & Pater's Acceptability Metric (2008), the candidate with the highest Acceptability Value will be the most well-formed. Thus, we can arrange the four candidates in question in accordance to their well-formedness:

9)

less well-formed -------------------------------------more well-formed
buenos--------- libros --------- libroh ---------- buenoh

Chapter 4

Artificial Learner Data

The first test to see whether a NHG model with lexically indexed constraints can successfully account for the pattern of production frequencies for each variant of Spanish –*s* aspiration and deletion is to determine whether a language learner can learn the correct production frequencies based on the grammar. To test this, relevant constraints and production frequencies will be run through an artificial language learner in Praat (Version 5.1.43). If an artificial language learner can learn the correct pattern of production frequencies for each variant under this model, then it supports NHG with lexically indexed constraints as a plausible model for variation across lexical categories.

For the purposes of this paper, I will analyze only the effect of the lexical category on Spanish s-deletion. If successful, indexes other than lexical categories may be added to relevant constraints in order to account for all effects on variation.

In order to obtain the frequencies for *s*-deletion within lexical categories, we must turn to Cedergren's (1973) original data from her study of Panama City Spanish. The results of her study are summarized in the (10).

10)

Lexical Category	Frequency of -s	Frequency of s-deletion (h or ∅)
Noun	14.4%	85.6%
Adjective	10.7%	89.3%

As seen above, -*s* resists deletion and reduction 4% more in nouns than adjectives. Based on the frequencies given originally by Cedergren (1973) on overall –*s* aspiration and deletion, it seems that- *s* is fully deleted 48% of the time while it is reduced to -*h* 41% of the time (with *s* showing up faithfully 11% of the time). Although Cedergren (1973) does not give the frequency of –*h* vs. -∅ for nouns and adjectives, we can estimate based on this data that the distribution would be about the same as it would for the general data. We can postulate the following frequencies for –*h* and -∅:

11)

Frequency of –s	Frequency of –h	Frequency of -∅
11%	41%	48%

12)

Lexical Category	Frequency of -*s*	Frequency of –*h*	Frequency of -∅
Noun	14%	39%	47%
Adjective	11%	41%	48%

These are the production frequencies I will use to test my model in an artificial language learner.

4.1 Methods

In order to test this model in Praat (Version 5.1.43), I chose a Linear OT decision strategy. Unlike other decision strategies in Praat, Linear OT limits the constraint weights to positive values, thus eliminating any negative disharmonies (Praat Version 5.1.43; Keller 2000, 2006). This means that each violation of a constraint will negatively impact the Harmonic Value of the candidate in violation, thus decreasing its harmony according to the number of its violations.

In addition to the Linear OT decision strategy, I employed the following parameters when testing my model:

13)

Evaluation noise: 2.0
Initial Plasticity: 1.0
Replications per plasticity: 25,000
Plasticity Decrement: 0.1
Number of Plasticities: 4
Rel. plasticity spreading: 0.1
Number of chews: 1

Because I wanted the learner to undergo 100,000 learning trials at 4 decreasing plasticities, the number of replications per plasticity was set at 25,000. This would ensure that the learner will begin with a plasticity of 1.0 and continue to change the constraint weights according to this plasticity for the first 25,000 learning trials. For the next 25,000 learning trials, the plasticity will reduce to 0.1, then to 0.01 for the next 25,000 learning trials, and finally to 0.001 for the final 25,000 learning trials. This strategy is to mimic the decreasing plasticity of a real language learner (Boersma & Hayes 2001. All other values for the above parameters were according to the default settings in Praat (Version 5.1.43).

The final components input into the Praat learner are the grammar and the distribution frequency text files. The grammar includes the constraints as well as the initial value, disharmony and plasticity for each constraint. The grammar also includes the necessary tableaus to show which kind of candidate is supposed to violate each constraint. For this grammar, ***S-SyllFinal** was given an initial weight of 100 while all of the faithfulness constraints were given an initial weight of 0 (Pater 2008). The learner was also provided with the following tableaus in (14) and (15).

14)

/buenos$_A$/	**ID-MANNER**	***S-SyllFinal**	**ID-PLACE$_N$**	**ID-PLACE$_A$**	**MAX$_N$**	**MAX$_A$**
buenos		*				
buenoh				*		
bueno						*
buenon	*					

15)

/libros$_N$/	**ID-MANNER**	***S-SyllFinal**	**ID-PLACE$_N$**	**ID-PLACE$_A$**	**MAX$_N$**	**MAX$_A$**
libros		*				
libroh			*			
libro					*	
libron	*					

For the production frequencies observed in the language, the learner was given the following output distributions, based on Cedergren's (1973) data on Panama City Spanish given in Table 4.

4.2 Results

As predicted, the artificial learner in Praat (Version 5.1.43) generated output frequencies for each variant form that were quite close to the distribution frequencies with which the learner was provided.

16)

Input	Candidate	Predicted Frequency	Frequency Produced by Learner
/buenos$_A$/	buenos	11.00%	9.69%
	buenoh	41.00%	42.05%
	bueno	48.00%	48.26%
	buenon	0.00%	0.00%
/libros$_N$/	libros	14.00%	13.16%
	libroh	39.00%	39.95%
	libro	47.00%	46.89%
	libron	0.00%	0.00%

As shown in (16), the frequencies produced by the artificial learner in Praat (Version 5.1.43) after 100,000 learning trials are incredibly similar to those it was fed at the beginning of the learning cycle. The model seems to have successfully produced the adult grammar characteristic of variable Spanish –*s* aspiration and deletion.

But how do we know if the learner has finished learning? Typically, the next step would be to test the adult grammar to see if the constraint weights have remained constant. However, due to the variable nature of the target grammar, the learner must constantly tweak the constraint weights with each new observed output (Boersma & Hayes 2001). Under the Gradual Learning Algorithm, the learner changes the weights of constraints if the observed form in the language differs from the predicted form by the data (Boersma & Hayes 2001). When a learner encounters a grammar that exhibits free variation, as in the case of Spanish –*s* aspiration and deletion, the weights of the learner's grammar at any given time may not produce the observed variant form it encounters in the language. Thus, the learner will constantly readjust the constraint weights to a certain degree based on what he/she observes at any given time in the language, giving rise to a grammar that can never truly converge in the traditional sense (Boersma & Hayes 2001).

Because no truly variable grammar can ever converge in the traditional sense (in which the constraint weights remain constant once the adult grammar has been learned), the closest test to see whether a NHG with lexically indexed constraints has successfully acquired the grammar is to run the entire learning cycle numerous times. If the grammar successfully produces very similar constraint weights with each learning cycle, then the grammar has come as close to convergence as any variable grammar can, thus proving successful in learning.

17)

Constraint	**Initial Weight**	**Weight after Learning Cycle 1**	**Weight after Learning Cycle 2**	**Weight after Learning Cycle 3**	**Weight after Learning Cycle 4**
***S-SyllFinal**	100.000	16.767	16.571	16.881	16.799
ID-PLACE$_N$	0.000	15.136	15.043	15.382	15.236
ID-PLACE$_A$	0.000	14.683	14.492	14.770	14.662
MAX$_N$	0.000	14.855	14.641	14.980	14.836
MAX$_A$	0.000	14.450	14.266	14.479	14.485
ID-MANNER	0.000	24.109	24.986	23.508	23.982

18)

Input	**Candidate**	**Selection Frequency (Cycle 1)**	**Selection Frequency (Cycle 2)**	**Selection Frequency (Cycle 3)**	**Selection Frequency (Cycle 4)**
/buenos-$_A$/	buenos	9.69%	9.86%	9.44%	9.83%
	buenoh	42.05%	42.14%	41.42%	42.42%
	bueno	48.26%	48.00%	49.14%	47.74%
	buenon	0.00%	0.00%	0.20%	0.10%
/libros$_N$/	libros	13.16%	13.31%	13.78%	13.14%
	libroh	39.95%	38.30%	38.01%	38.35%
	libro	46.89%	48.39%	48.21%	48.51%
	libron	0.0%	0.00%	0.40%	0.30%

As (18) shows, the selection frequencies for each variant are remarkably similar across all four learning cycles. Since no variable grammar can completely converge due to the relatively unpredictable nature of the variation, successful replications of the same learning cycles such as this are as close to convergence as one can possibly hope for (Boersma & Hayes 2001). The stability of the selection frequencies for each variant across four separate learning cycles shows that this grammar is able to learn the correct output frequencies to a spectacular degree of accuracy. This is because it is able to learn the correct weights for each constraint. Stable constraint weights make for stable production frequencies, thus successfully acquiring the target grammar. The constraint weights for each learning cycle are given below (see the Appendix for tableaus).

As (17) shows, the weights for each constraint has remained relatively stable over time. This produces a stable proportion of variant percentages in accordance with the distribution frequencies with which the learner was provided at the beginning of each learning trial. The model has successfully learned constraint weights which always produce the output distribution frequencies of variants of the adult grammar. Therefore, we can conclude that a NHG with lexically indexed constraints is capable of learning and producing the variation that is observed in Spanish –*s* aspiration and deletion.

In order to test whether or not production frequencies can be tied to perception judgments made by native speakers, I must first plug in the Harmonic Values for each variant form into the Acceptability Metric in order to ascertain their Acceptability Values. Since in this case each candidate only violates one constraint one time (as illustrated in (14) and (15)), then we can use the weights of each candidate as mock Harmony Scores for each variant by simply multiplying them by -1. I chose the constraint weights from Learning Cycle 1 in (17) in order to ascertain the Harmonic Values for each candidate and input them into Coetzee & Pater's (2008) Acceptability Metric equation to yield the Acceptability Values in (19).

19)

Acceptability (libros) = $H_{libros} – H_{libroh}$ = -16.767 - (-15.136) = -1.631
Acceptability(buenos) = $H_{buenos} – H_{buenoh}$ = -16.767 – (-14.683) = -2.084
Acceptability (libroh) = $H_{libroh} – H_{libros}$ = -15.136 – (-16.767) = 1.631
Acceptability (buenoh) = $H_{buenoh} – H_{buenos}$ = -14.683 – (-16.767) = 2.084

Keeping in mind that the higher the Acceptability Value, the more acceptable the variant, we can arrange the four variants in the following scale of well-formedness:

20)

less well-formed ---more well-formed
buenos---------- libros ------------ libroh -------------- buenoh

Since Puerto Rican Spanish and Panama City Spanish show the same pattern in distribution of these four variants, we can expect the two dialects to share the same general hierarchy of acceptability based on the relationship between relative constraint weights. Keeping this in mind, I will use perception judgments from Puerto Rican Spanish speakers to test whether perception data mimics the acceptability scale outlined above.

Chapter 5

Gathering and Testing Well-Formedness Judgments in NHG

In order to test whether NHG with lexically indexed constraints can account for well-formedness judgments based on lexical category, it was first necessary to collect such judgments. While perception studies have been done before on Spanish *–s* aspiration and deletion, most are sociolinguistic in nature and none have focused on the role that lexical category plays on shaping these well-formedness judgments. In order to do this, I conducted two speech perception experiments in which native Puerto Rican Spanish speakers were asked to make judgments about nouns and adjectives with word-final *–s*, *-h*, or an artificial fricative that lay somewhere in between the two variants along a continuum.

There are three main reasons I chose Puerto Rican Spanish speakers as the subjects of my studies. First, Puerto Rican Spanish is one of the dialects which show the same general pattern of *–s* aspiration and deletion as Panamanian Spanish. As in Panamanian Spanish, adjectives are affected more than nouns (Cedergren 1973; Poplack 1980). This means that the constraints governing the production of each variant should hold the same relationships as those that govern variants in Panama City Spanish. If Harmonic Values for candidates are to be used to predict grammaticality judgments, then the same variants should be more or less well-formed in each dialect. Secondly, the Puerto Rican community in North Carolina has the second highest population of any Hispanic community in the state at 8.2%. This community is outnumbered only by those of Mexican descent, who account for 65% of the North Carolina Hispanic population (Census 2000). This means that Puerto Rican Spanish is the second most widely spoken dialect of Spanish in North Carolina, preceded only by Mexican Spanish, which varies widely across social classes as well as region (Morris 2000). To facilitate the recruitment of an adequate number of subjects for these studies, Puerto Rican Spanish seemed like the obvious dialect of choice. Finally, there have been a few perception experiments done on Puerto Rican Spanish in regards to word-final –*s* aspiration and deletion. Shana Poplack (1980) conducted a study which tested Puerto Rican Spanish speaker's ability to identify plural forms in fully deleted variants once the variants were taken out of the greater context of the sentence. Sara Mack (2011) conducted a study to see if there was a link between perceived sexual orientation and deleted -s variants. However, no study have focused on the role of lexical category in perception judgments of variant forms.

For the following two experiments, seven native Puerto Rican Spanish speakers were recruited using mass emails to all UNC students and personnel as well as recruitment flyers posted around UNC's campus. All seven speakers were of adult age and grew up in Puerto Rico with Puerto Rican Spanish speaking parents. All subjects were also fluent in English and of relatively high social class, either having earned college degrees and working white-collar jobs or in the process of completing college degrees.

5.1 Experiment 1 – Perception of Sentences

For Experiment 1, I presented native Puerto Rican Spanish speakers with two lexically identical sentences at a time and asked them to choose which sentence sounds better, or more natural, to them. In each sentence one word-final –*s* was aspirated to an –*h*. The participants had to choose between a sentence in which the –*h* occurred at the end of a noun and a sentence in which the –*h* occurred at the end of an adjective. In other words, they were forced to choose whether it sounded better to them hear a word-final –*h* on the end of an adjective or a noun. If my hypothesis was correct and grammaticality judgments can be predicted by the harmonic values of variant forms, then the participants would choose the sentences with an –*h* on the adjective more often than the sentences with the –*h* on the noun. In other words, if the well-formedness scale for each variant in (20) is correct, then participants should choose sentences with aspirated adjectives as more well-formed more than sentences with aspirated nouns since the grammar predicts that aspirated adjectives are more well-formed than aspirated nouns. As seen in Section 5.1.2, this is precisely what happened for all seven participants.

5.1.1 Methods

For Experiment 1, 16 lexically unique sentences, each of which included an adjectival phrase, were recorded by a native Spanish speaker. The speaker is a 24-year-old male from Ecuador. Each lexically unique sentence was recorded twice, once in which the speaker produced a clear –*h* at the end of the noun, and once in which the speaker produced the –*h* at the end of the adjective. All other –*s* sounds in the sentences were clearly annunciated as such in order to avoid any additional aspiration. The sentences were recorded in the soundproof booth in the Phonetics Lab on UNC's campus using a RadioShack Hands-Free Headset Microphone.

In order to produce the stimuli, each pair of lexically identical sentences were presented in both possible orders. In the first order, the aspirated adjective was presented in the first sentence, and in the second order the aspirated adjective was presented in the second sentence. This was to ensure that the participants were actually choosing which sentence sounded best to them rather than continuously choosing Sentence 1 or Sentence 2. Each pair of sentences were presented in both orders and repeated twice for a total of 64 stimuli. The order in which the sentences were presented were randomized, excluding doublets, in the Praat experiment script. The experiment was run in Praat in the soundproof booth in the Phonetics Lab on UNC's campus. The Praat script for Experiment 1 can be found in the Appendix.

In order to ensure that the sentences were testing solely the role of lexical category on the perception of –*s* aspiration, several controls were employed in selecting the nouns and adjectives used. In her study of Panama City Spanish, Cedergren (1973) identified several factors that affected the frequencies of the variant forms of Spanish –*s* aspiration and deletion, as illustrated in Table 1. In order to control for the grammatical status's role in the perception of –*s* aspiration, only plural –*s* suffixes were used in the sentences, while monomorphemic –*s* and second person singular suffix –*s* were excluded. The effect of following segment was also controlled for by creating an equal number of vowels and consonants following the fricatives in question. The control of following segment was incorporated both within the adjectival phrase and within the larger sentence, the former being accomplished by the nouns or adjectives following the fricatives in question beginning with an equal number of vowels and consonants and the latter being accomplished by putting either *ayer*, meaning 'yesterday', or *mañana*, meaning 'tomorrow', at the end of the sentence. This was to ensure that no matter whether the fricative in question was phrase-final or not, the following segment was controlled for.

A few additional controls were incorporated into the study for good measure. While Cedergren (1973) mentions nothing about the position within an adjectival phrase having an effect on the frequency of *–s* aspiration or deletion, an equal number of pre-nominal and post-nominal adjectives were used. This was to ensure that no phrase-final weakening effects were playing a role in the results of the study. Lexical frequency was also controlled for using a Spanish Frequency Dictionary (Juilland & Chang-Rodriguez 1964). The influence of lexical frequency on phonological processes has long been noted in the literature, particularly in cases of weakening or deletion (Diaz-Campos & Ruiz-Sanchez 2008; Bybee 2002). In regards to lexical frequency as a motivator for language change and the rise of phonological variation. Studies of variation in terms of deletion or weakening of various languages, including English and Spanish, show that high lexical frequency correlates to high rates of deletion or weakening (Bybee 2002). In other words, the more frequent a lexical item is, the more likely it is to undergo weakening or deletion. Given the cross-linguistic nature of the influence of lexical frequency on phonological processes, the lexical frequency of the nouns and adjectives examined in this study must surely be controlled. In order to ensure that lexical frequency of the nouns and adjectives was not influencing the results, all nouns and adjectives were carefully chosen based on their lexical frequency. In order to select the adjectives and nouns used in the stimuli, I used the *Frequency Dictionary of Spanish Words* by Alphonse Juilland & E. Chang-Rodriguez (1964). I used this particular lexical frequency dictionary because Juilland & Chang-Rodriguez (1964) provide lexical frequency for each morphological variant of a lexical entry. For example, they do not only provide the lexical frequency for *muchacho*, but also *muchacha, muchachos,* and *muchachas*. This ensured that the plural form I was choosing did indeed have the lexical frequency I was looking for, as well as accounting for feminine and masculine forms. The frequency of each lexical item was based on its number of occurrences within a corpora of approximately 500,000 Spanish words that included five lexical worlds: dramatic literature, fictional literature, essayistic literature, technical literature, and periodical literature. Of a range of observed frequencies between 1 and 29,000, all nouns and adjectives chosen for this study had frequencies between 11 and 19.

All controls considered, the following set of unique adjectival phrases were chosen (lexical frequency values from Juilland & Chang-Rodriguez (1964) are given in parenthesis to the right of their corresponding lexical item):

21)

Second Word	**pre-nominal adjective**	**post-nominal adjective**
vowel-initial	difíciles (11) obstáculos (11)	organismos (17) invisibles (13)
	malas (11) impresiones (19)	oficinas (10) oficiales (14)
consonant-initial	viejos (16) profesores (13)	discursos (16) profundos (16)
	simples (13) descripciones (17)	muchachas (14) tristes (18)

With all the controls properly in place, the stimuli chosen ensured that only lexical category would influence the participants' perception of the sentences. The independent variables in this experiment are word frequency, following segment, adjectival position, and gender, and lexical category. The dependent variable is whether the participant chooses to aspirate the noun or the adjective. A full list of the sentences used as well as the Praat script for Experiment 1 is located in the Appendix.

5.1.2 Results

As anticipated, all seven participants chose the sentences in which the –*h* was on the adjective more often than they chose the sentences in which the –*h* was on the noun. This pattern reflects the production frequencies quite well. Since –*s* aspiration occurs more commonly on adjectives than on nouns (Cedergren 1973; Evanini 2007; Poplack 1980), it makes perfect phonological sense that an –*h* plural marker would sound more natural on an adjective than on a noun. After all, the more rare an item is, the more it should stick out like a sore thumb, thus making the participants more likely to reject it as an unnatural or awkward sounding utterance.

22)

Participant #	**- h on noun**	**-h on adjective**
1	47%	53%
2	47%	53%
3	41%	59%
4	44%	56%
5	42%	58%
6	45%	55%
7	42%	58%
Average	**44%**	**56%**

To test whether these results were statistically significant enough to refute the null hypothesis, a logistic regression was run using a mixed model to determine whether or not each fixed effect influenced the participants' responses. The random effect was participant. The p-values, parameter estimates and standard errors for each fixed effect is given below.

23)

Fixed Effects	**Parameter Estimates**	**P-Value**	**Standard Error**
Intercept	0.2206	0.5483	0.3675
Lexical Category	-0.2422	0.0109	0.0952
Adjectival Position (pre-nominal vs. post-nominal)	-1.0489	0.0349	0.4973
Following Segment (within noun phrase)	0.8099	0.0128	0.3253
Following Segment (outside of noun phrase)	0.1004	0.7075	0.2675

Position of the adjective was shown to be a significant variable, with a p-value of 0.0349. Following segment showed some interesting results. The following segment within the adjectival phrase (ie, the beginning segment of the second word in the adjective-noun or noun-adjective sequence) was significant, with a p-value of 0.0349. However, outside of the adjectival phrase, following segment was insignificant, with a p-value of 0.7075. This is unexpected since no studies have indicated why following segment should only matter in one syntactic context as opposed to another. Most importantly for this study, the lexical category of the aspirated word was shown to be significant, receiving a p-value of 0.0109. This means that the lexical category of the aspirated word had a major impact on which sentence sounded best to our participants.

These results fit nicely with our NHG model with lexically indexed constraints. Since Puerto Rican Spanish has been shown by other investigators to show the same general pattern of –*s* aspiration and deletion occurring more often on adjectives than on nouns, we can assume that their constraint weights would look quite similar to those we saw in Panama City Spanish in Section 4 (although the specific weights of the constraints may vary between dialects). Since more variation would exist between *–h* and *–s* forms in nouns than in adjectives, we would expect the constraint weights to reflect that pattern just as they did in Panama City Spanish. In other words, the constraint weights for ***S-SyllFinal** and **ID-PLACE**$_{\mathbf{N}}$ would have to be closer together than the weights for ***S-SyllFinal** and **ID-PLACE**$_{\mathbf{A}}$. With these assumptions in mind, we can use Coetzee & Pater's Acceptability Metric to estimate the approximate the proportions of acceptability scores for the variants. Given the closeness of the weight of the markedness constraint to those of the lexically indexed faithfulness constraints, we should expect to see the same outcome of the Acceptability Metric equation in (2). Using this metric, a fully faithful candidate would have a higher Acceptability value as a noun than it would if it were an adjective. Likewise, an aspirated *–h* variant would have a higher Acceptability value as an adjective than as a noun (See Section 4 for details). The higher the Acceptability value, the more well-formed the variant. By using the Harmonic Values of each variable candidate to determine its Acceptability value in reference to another candidate, this NHG model with lexically indexed constraints can accurately predict the well-formedness judgments of variable forms by native speakers of Puerto Rican Spanish, showing that it can indeed account for perception data.

5.2 Experiment 2 – Perception of Artificial Fricatives

The purpose of Experiment 2 is to test whether the lexical category of a variant affects how native Puerto Rican Spanish speakers perceive the actual fricative in question. In other words, if it is more rare for an aspirated *–s* to occur on a noun than an adjective, is it more obvious if a native Spanish speaker hears an –*h* on a noun as opposed to an adjective?

For this experiment, participants were presented with nouns and adjectives with word-final fricatives that have been altered to be more h-like or s-like and were asked to choose whether they heard an –*h* or an –*s*. If my hypothesis is correct and lexical category does affect how speakers of language that varies across lexical category perceive the variable sounds in question, then the participants would answer that they heard –*h* more on nouns than on adjectives. For the exact same fricative, the participants would answer that they heard an –*s* more on adjectives than on nouns. In other words, the point on the continuum between –*h* and –*s* at which the fricatives start to sound s-like will occur earlier in adjectives than in nouns. This pattern of responses would show that the –*h* stands out more on nouns than adjectives, which is precisely the result that we find with 6 out of the 7 participants.

5.2.1 Methods

In order to create the stimuli, I first had to create the fricatives that would make up the continuum from –*h* to –*s*. The fricatives were made by taking a naturally occurring –*h* and a naturally occurring –*s* from the recordings for Experiment 1 and modifying them. I spliced the –*h* and –*s* off of the end of a word ending in –*e* at a point where the amplitude was at 0. This was to eliminate any potential unnatural sounding "clicks" that I might get from a sudden change in amplitude when I added the fricatives to a lexical base later on. In order to make the continuum of fricatives, I started with a full –*h* and made fricatives that were slowly becoming more s-like. Following Alan C. L. Yu's (2010) methodology, I did this by taking the full –*h* and –*s* and multiplying them by the percentages that I wanted in Praat, and then superimposing them upon each other to create a new hybrid fricative. For example, if I wanted to create a fricative that was 95%h and 5%s (thus slightly more s-like than a regular –*h*), then I multiplied the amplitude of the –*h* sound file by 0.95, multiplied the amplitude of the –*s* sound file by 0.05, and then used Praat to create a new fricative from the two modified fricatives. To create the new fricative in Praat, I highlighted the modified –*h* and –*s* sound files, selected New, then Sound, and then Create Sound from Formula. I used the following specifications for the formula window that popped up:

24)

Name: 95h05s (the numeric values would change depending on percentages of each sound)
Channels: Mono
Start time(s): Sound_h.xmin
End time(s) Sound_h.xmax
Sampling frequency (Hz): 44100
Formula: Sound_h(x) + Sound_s(x)

The Channels and Sampling frequency were kept at the default settings while the formula superimposes the sound named 'h.wav' and the sound named 's.wav' on top of each other to make one *h/s* hybrid sound. I increased the percentage of –*s* and decreased the percentage of –*h* by 5% for each new fricative until I reached the full –*s* sound to get 21 total fricatives. Each fricative was then tapered for the first and last 10 ms of the fricative to make it sounds more natural. In order to taper each fricative, I selected it in the Praat objects window and then selected New, then Sound, then Create Sound from Formula. In the Formula box, I then input the following formulas:

25)

```
if (x>0.010) then self else self * (x/0.010) endif
if (x < (xmax-0.010)) then self else self*(100 * (xmax-x)) endif
```

These formulas successfully tapered the ends of each fricative, thus avoiding a sudden burst of fricative energy at the end of each word which would sound unnatural (Moreton 2010).

The next step in creating the stimuli were to attach the fricatives to roots of nouns and adjectives. From the recorded nouns and adjectives in the sentences for Experiment 1, I randomly chose the noun *profesores,* meaning 'professors', and *invisibles,* meaning 'invisible' from among the nouns and vowels that had an *–e* as the final vowel. It was important that the fricatives be spliced onto a root ending in *–e* because they were originally taken from fricatives that followed in *–e,* thus reducing the risk of the formant transitions in the preceding vowel and fricatives not matching up and creating an unnatural sound. Crucially, I used two versions of both the noun and the adjective: one version that was originally recorded with a word-final *–h*, and one that was recorded with a word-final *–s*. This was to ensure that any formant transitions in the vowel preceding the word-final fricatives (which differ for vowels preceding an *–h* and those preceding an *–s*) wouldn't be to blame for any discrepancies in perception between the noun and adjective. In other words, if I had just used an original *–h* ending noun and an original *–s* ending adjective, then the different formant transitions of the vowel preceding the original fricative could have tipped the participant off as to what sound they were about to hear. Each original fricative was spliced off of the end of the lexical root at a zero-crossing so that the amplitude of the root and the amplitude of the fricative that would be concatenated onto the end would match up (both being zero), thus eliminating any risk of clicking and making the new word sound most natural. I used Praat (Version 5.1.43) to concatenate each fricative onto all four lexical bases to create 84 individual words.

Once all the potential stimuli had been created, I listened to them to decide whether to use all of the individual words as stimuli or not. I decided on using only words with fricatives between 95-40%h along with full *–h* and *–s* sounds as endpoints. This is because to my ear, anything more s-like than a fricative that was 60%h and 40%s sounded entirely s-like. Anything farther towards the *–s* end of the continuum than 40%h 60%s would surely be judged as *–s,* so I decided to reduce the number of stimuli to natural fricative endpoints as well as all fricatives between 95-40%h. This generated 56 unique stimuli. Each stimulus was repeated twice in Experiment 2, thus generating 112 total tokens. The independent variables in this experiment were step on the continuum, the original fricative, and the lexical category of the stimulus. The dependent variable is whether the participant chooses *–h* or *–s.* The Praat script for Experiment 2 can be found in the Appendix. If the well-formedness scale in (20) is correct, then an aspirated noun should be viewed as more marked, or more obvious, than an aspirated adjective since aspirated nouns are less well-formed.

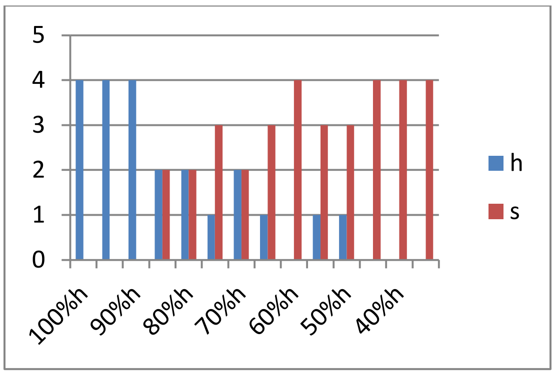
5.2.2 <u>Predicted Results</u>

Experiment 2 yielded some very interesting results. 6 of the 7 participants showed the same pattern of answers, judging the same artificial fricatives much more often as an *–s* when the stimulus was an adjective as opposed to when it was a noun. Participant #5, however, answered that she heard *–s* for almost every stimulus, which I will address later on in this section.

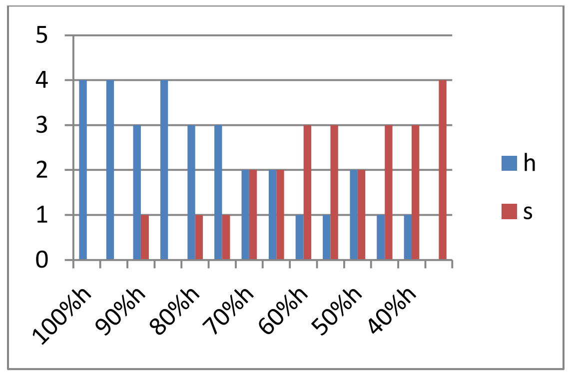
Below, (26)-(31) display the results for Participants 1, 2, 3, 4, 6 and 7, respectively, for the adjective and noun stimuli.

26)

Results for Adjectives

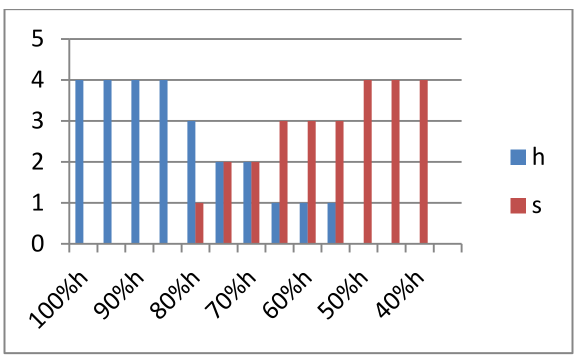


Results for Nouns

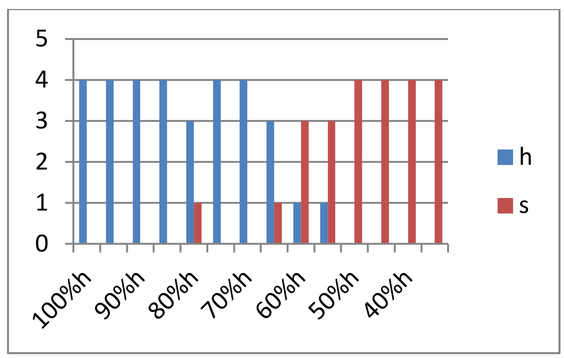


27)

Results for Adjectives

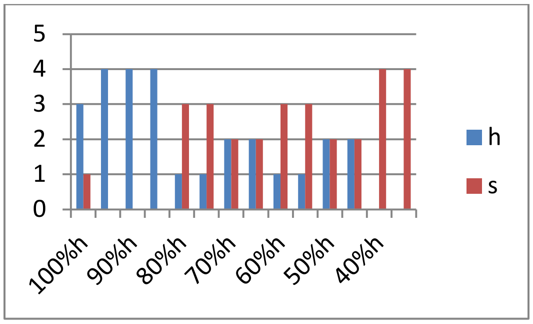


Results for Nouns

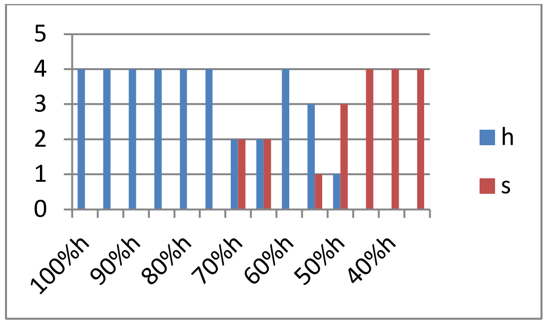


28)

Results for Adjectives

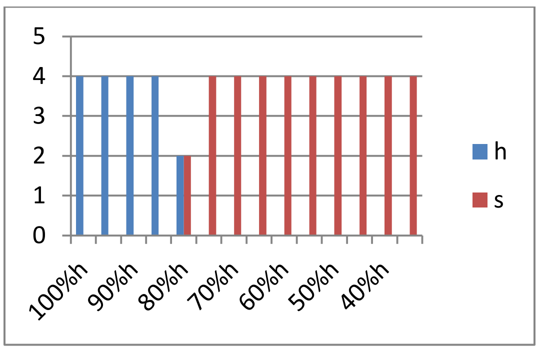


Results for Nouns

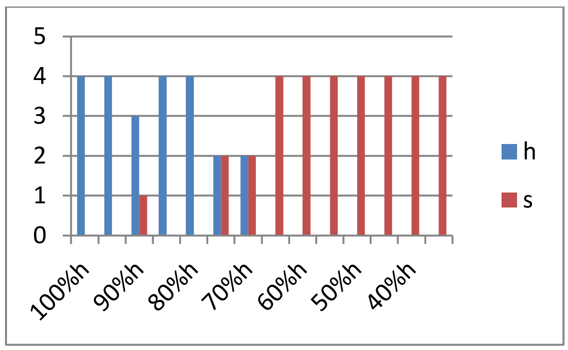


29)

Results for Adjectives

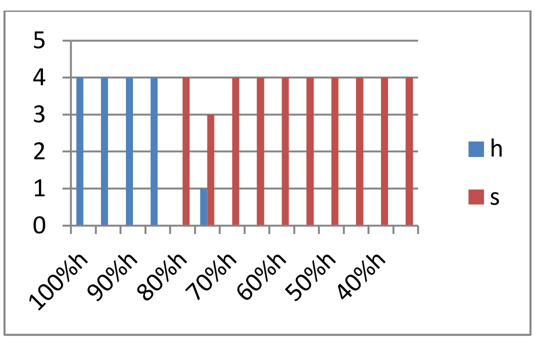


Results for Nouns

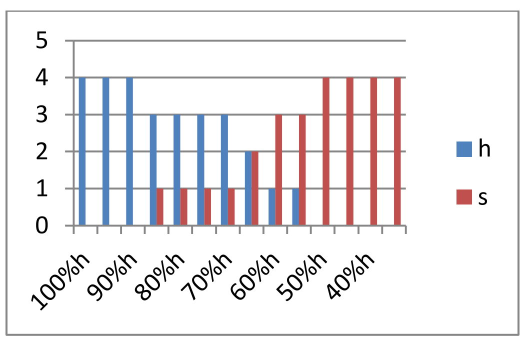


30)

Results for Adjectives

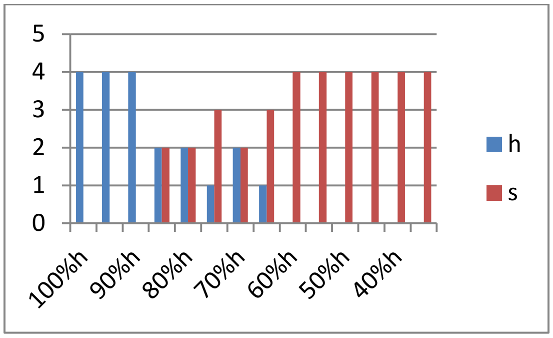


Results for Nouns

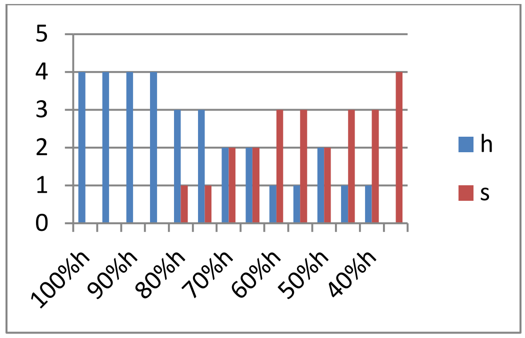


31)

Results for Adjectives



Results for Nouns

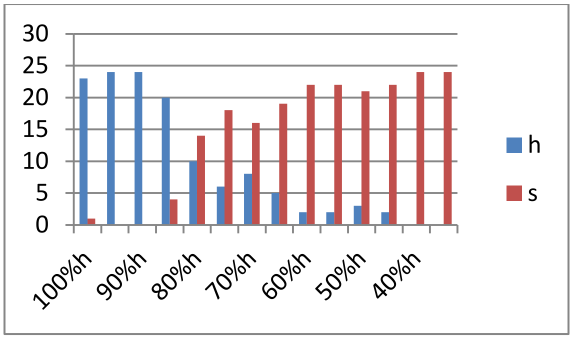


As the graphs above show, every participant started to choose *–s* more than *–h* at a later point on the continuum for nouns than for adjectives. This suggests that the *–h* stands out more on nouns than adjectives. Given that *–h* variants are more rare on nouns than they are on adjectives, this matches up well with the observed output frequencies.

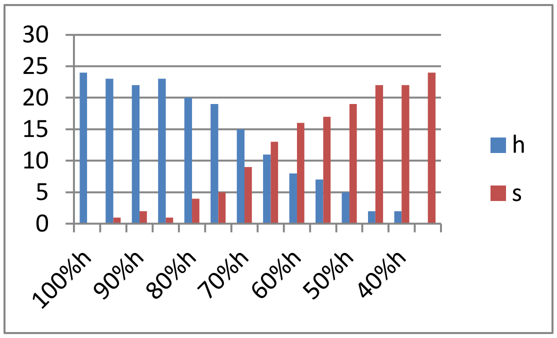
For a snapshot of the affect of lexical category on the perception of word-final fricatives on the entire Puerto Rican Spanish dialect, the following graphs display the combined results for the 6 participants who performed consistently with respect to one another. Since each fricative was spliced onto four separate lexical bases and these graphs represent the responses from 6 participants, each fricative had 24 tallied responses.

32)

Total Results for Adjectives



Total Results for Nouns



(32) shows a very interesting pattern. For adjectives, the point along the continuum at which most speakers begin to hear an *–s* more often than they hear an *–h* is right around the fricative which is 80%h and 20%s. For nouns, however, the point on the continuum at which most speakers begin to hear an *–s* comes much later, with the number of *–s* responses outnumbering the number of *–h* responses only beginning at the fricative which is 65%h and 35%s.

After the results were tallied, a logistical regression was run using a mixed model on the fixed effects of original fricative (h or s), the step on the continuum (ie, percentage of -h and -s in each artificial fricative), and lexical category of the stimulus. The random effect was participant. Since Participant 5 responded significantly beyond the scope of the other participants' responses, the data from her responses was not analyzed in these regressions. The p-values, parameter estimates and standard error for each fixed effect is given below.

33)

Fixed Effects	**Parameter Estimates**	**P-Value**	**Standard Error**
Intercept	-6.9124	<0.0001	0.5555
Step on Continuum	0.1037	<0.0001	0.0078
Original Fricative (-*h* or –*s*)	0.3258	0.1325	0.2166
Lexical Category	-0.8812	<0.0001	0.2217

Predictably, the step on the continuum did significantly affect the participants' responses with a p value of <0.0001, with a higher percentage of *–h* increasing the likelihood of the participant hearing an –*h*. Whether the stimulus originally had an –*h* or an –*s* did not significantly change the participants' responses, yielding a p value of 0.1325. Most importantly for this study, the lexical category of the stimulus yielded a p-value of <0.0001, showing that lexical category did significantly change the way that the participants' heard each fricative (with a higher chance of hearing –*h* if the stimulus was a noun).

These results are surprising when compared to other ambiguous phoneme studies. Numerous experiments have shown that extra-acoustic factors such as phonotactics have a significant effect on phoneme perception (Ganong 1980). For instance, a participant is much more likely to perceive an ambiguous speech sound as a phoneme that is allowed in a given phonotactic context in their native language (Moreton 2002). For example, English speakers are much more likely to perceive an ambiguous liquid as [l] when it is preceded by a [s] rather than a [t] because the constraint set on English onset clusters disfavors [coronal-coronal] onset clusters such as [tl] (Moreton 2002; Pitt 1998; Massaro & Cohen 1983). Studies of ambiguous phoneme perception in Stochastic OT show that whether a liquid was interpreted as light [l] or dark [ɫ] depended heavily on the morphophonemic context of the liquid, where constraints restricted the likelihood of either variant occurring in a specific context (Boersma & Hayes 2001). Such phonotactic effects are not present in our results of phoneme perception in Spanish word-final fricatives. In standard Spanish, the grammar prohibits [h] from occurring in the coda position, favoring instead coronal consonants such as [s] (Harris 1983). The grammar of Puerto Rican Spanish strays from this standard grammar, allowing significant variation between aspirated and faithful fricative codas. If phonotactics were solely responsible for the perception of phonemes, then we would expect the perception of word-final Spanish fricatives to follow the same pattern as phonotactic effects in other languages. Since the grammar of Puerto Rican Spanish allows more –*h* variants of adjectives than nouns, then the perception mechanism of the grammar should trend towards perceiving more –*h* on adjectives than nouns, which is the opposite of what we see in (32). The relative frequency of each variant does not directly mirror how the variable fricative is perceived. In other words, participants do not hear more –*h* on adjectives even though the aspirated variant is more frequent in adjectives than in nouns. One explanation may be that an aspirated noun is more marked than an aspirated adjective, making it more obvious to the participant that a non-standard form is being uttered. According to the scale of acceptability for aspirated and unaspirated nouns and adjectives in (20), aspirated nouns are indeed less well-formed than aspirated adjectives, making aspirated nouns more marked. An interaction of lexical-category-based markedness and traditional disfavoring of an aspirated coda seem to be behind the participants performance in Experiment 2. It is more obvious when a noun violates the standard grammar, so it is viewed as marked and less well-formed than an aspirated adjective.

5.2.3 <u>Participant 5 Results</u>

As stated in Section 5.2.2, Participant 5 responded to the stimuli in Experiment 2 in a very different way than the other six participants. This participant answered that she heard an –*s* for 106 out of the 112 stimuli, including 7 of the 8 stimuli which included a full, unmodified –*h*. Unlike the other six participants, it seems that Participant 5 is indeed normalizing her perception for context-induced variation, hearing mostly the faithful fricative –*s*. The stimuli to which Participant 5 responded that she heard an –*h* are listed in (34).

34)

Lexical Category	**Original Fricative** (-*h* or –*s)*	**Modified Fricative** (Percentages of –*h* and –*s*)
adjective	-*h*	80% h, 20% s
adjective	-*h*	85% h, 15% s
adjective	-*h*	50% h, 50% s
adjective	-*h*	85% h, 15% s
adjective	-*h*	100% h (unmodified)
noun	-*h*	85% h, 15% s

What is striking about these results is not only the fact that the participant reported that she heard –*s* almost every time, but that almost all of the stimuli which sounded like they contained an –*h* to her are adjectives. If the lexical categories of the stimuli to which the participant responded with an answer of –*h* were more mixed, it would be easy to write this data off as the results of a participant who constantly just clicked the "s" box on the screen. However, she reports that she hears an –*h* almost exclusively on adjectives. In other words, she rarely hears an –*h* at the end of words, but when she does, she hears it on adjectives. If the participant is indeed choosing the variant which is more "correct" or "standard", the fact that she chooses –*h* almost exclusively for adjectives indicates that the aspirated variant is more acceptable on adjectives than on nouns. This would line up with her performance in Experiment 1, where, like the other six participants, she chose the sentences with aspirated adjectives as sounding better than the ones with aspirated nouns.

The only significant difference between Participant 5 and the other six participants is her age. While the other six participants were aged 18-35, Participant 5 was 63. According to Janson (1983), changes in perception among speakers of a dialect happen after changes in production in that dialect. Thus, the same speaker who regularly produces variant forms may not perceive them as different. This was the case for Swedish participants who were given artificial vowels on a continuum between /o:/ and /a:/ and asked to choose which vowel they heard. There was a significant difference in the phoneme boundary between older participants and younger participants, showing that different generations perceived the same artificial stimuli differently (Janson 1983). The age discrepancy between Participant 5 and the other six participants lends support to this theory. In response to the variation found in the production of word-final –*s*, the younger speakers perceived the fricatives differently than the older speaker.

Chapter 6

Conclusions

This study of Spanish –*s* deletion and aspiration has shown that the lexical category of a given input not only has a direct effect on the production frequencies of phonological variants, but also has an effect on the perception of those variants. By adding lexical indexes to the input and relevant Faithfulness constraints in a NHG model, both variation within and across lexical categories in production can be accounted for. By adding indexes to inputs and constraints, extra-phonological influences on variation can be accounted for. This model may be expanded to account for other extra-phonological factors such as formality register or lexical frequency which have been shown to influence variation. This study has shown that lexical category can indeed be successfully incorporated into the phonological model, although further investigation needs to be done in order to ascertain whether this is true of all extra-phonological factors or whether there is some restriction on which kinds of linguistic and social factors may be incorporated into a phonological model of variation.

This study has also shown that in the NHG framework, production frequencies of variable candidates can be used to predict well-formedness judgments of those candidates. By using the constraint weights generated by the artificial learner in Praat (Version 5.1.43) in Chapter 3, I was able to successfully predict well-formedness judgments by Puerto Rican Spanish speakers in Chapters 4 and 5. The model predicted that aspirated adjectives are more well-formed than aspirated nouns. This was reflected in the results for Experiment 1, in which participants regularly chose the sentence with an aspirated adjective as more well-formed than a sentence with an aspirated noun. It was also reflected in the results for Experiment 2, in which 6 out of 7 participants classified the same artificial fricative more as –*h* on nouns than on adjectives due to the less well-formed status of aspirated nouns (thus classifying them as more marked, or bad, than aspirated adjectives). Further investigation needs to be done to ascertain whether these perception judgments are uniform across all speakers of dialects of Spanish in which aspirated adjectives are more common than aspirated nouns. Studies may also be done to test whether production frequencies of other phenomena of phonological variation predict well-formedness judgments as successfully as this study of Spanish s-deletion predicted well-formedness judgments of Puerto Rican Spanish speakers.

Appendix A

Artificial Learner Scripts

1. Grammar Script

```
File type = "ooTextFile"
Object class = "OTGrammar 2"

<LinearOT>
0 ! leak
6 constraints
constraint [1]: "*S\s{syllfinal}" 100 100 1
constraint [2]: "ID-PLACE\s{N}" 0 0 1
constraint [3]: "ID-PLACE\s{A}" 0 0 1
constraint [4]: "MAX\s{N}" 0 0 1
constraint [5]: "MAX\s{A}" 0 0 1
constraint [6]: "ID-MANNER" 0 0 1

0 fixed rankings

2 tableaus
input [1] : "buenos" 4
  candidate [1]: "buenos" 1 0 0 0 0 0
  candidate [2]: "buenoh" 0 0 1 0 0 0
  candidate [3]: "bueno" 0 0 0 0 1 0
  candidate [4]: "buenon" 0 0 0 0 0 1

input [2] : "libros" 4
  candidate [1]: "libros" 1 0 0 0 0 0
  candidate [2]: "libroh" 0 1 0 0 0 0
  candidate [3]: "libro" 0 0 0 1 0 0
  candidate [4]: "libron" 0 0 0 0 0 1
```

2. Output Distributions Script

```
File type "ooTextFile"
"PairDistribution"
6 pairs
"buenos" "buenos" 10
```

"buenos" "buenoh" 42
"buenos" "bueno" 48

"libros" "libros" 14
"libros" "libroh" 39
"libros" "libro" 47

Appendix B

Artificial Learner Tableaus

Cycle 1

	ranking value	*disharmony*	*plasticity*
ID-MANNER	24.109	21.099	1.000000
ID-PLACEN	15.136	17.126	1.000000
ID-PLACEA	14.683	15.324	1.000000
***Ssyllfinal**	16.767	14.188	1.000000
MAXA	14.450	13.645	1.000000
MAXN	14.855	12.812	1.000000

buenos	ID-MANNER	ID-PLACEN	ID-PLACEA	*Ssyllfinal	MAXA	MAXN	
buenos				*			-14.188
buenoh			*				-15.324
☞ bueno					*		-13.645
buenon	*						-21.099

libros	ID-MANNER	ID-PLACEN	ID-PLACEA	*Ssyllfinal	MAXA	MAXN	
libros				*			-14.188
libroh		*					-17.126
☞ libro						*	-12.812
libron	*						-21.099

Cycle 2

	ranking value	*disharmony*	*plasticity*
ID-MANNER	24.986	23.353	1.000000
ID-PLACEN	15.043	18.985	1.000000
***Ssyllfinal**	16.571	17.264	1.000000
ID-PLACEA	14.492	15.595	1.000000
MAXN	14.641	14.035	1.000000
MAXA	14.266	12.591	1.000000

buenos	ID-MANNER	ID-PLACEN	*Ssyllfinal	ID-PLACEA	MAXN	MAXA	
buenos			*				-17.264
buenoh				*			-15.595
☞ bueno						*	-12.591
buenon	*						-23.353

libros	ID-MANNER	ID-PLACEN	*Ssyllfinal	ID-PLACEA	MAXN	MAXA	
libros			*				-17.264
libroh		*					-18.985
☞ libro					*		-14.035
libron	*						-23.353

Cycle 3

	ranking value	*disharmony*	*plasticity*
ID-MANNER	23.508	22.440	1.000000
MAXA	14.479	18.398	1.000000
***Ssyllfinal**	16.881	15.811	1.000000
MAXN	14.980	14.275	1.000000
ID-PLACEA	14.770	14.266	1.000000
ID-PLACEN	15.382	13.433	1.000000

buenos	ID-MANNER	MAXA	*Ssyllfinal	MAXN	ID-PLACEA	ID-PLACEN	
buenos			*				-15.811
☞ buenoh					*		-14.266
bueno		*					-18.398
buenon	*						-22.440

libros	ID-MANNER	MAXA	*Ssyllfinal	MAXN	ID-PLACEA	ID-PLACEN	
libros			*				-15.811
☞ libroh						*	-13.433
libro				*			-14.275
libron	*						-22.440

Cycle 4

	ranking value	*disharmony*	*plasticity*
ID-MANNER	23.982	21.868	1.000000
***Ssyllfinal**	16.799	18.608	1.000000
MAXN	14.836	17.777	1.000000
ID-PLACEA	14.662	17.483	1.000000
MAXA	14.485	14.302	1.000000
ID-PLACEN	15.236	11.961	1.000000

buenos	ID-MANNER	*Ssyllfinal	MAXN	ID-PLACEA	MAXA	ID-PLACEN	
buenos		*					-18.608
buenoh				*			-17.483
☞ bueno					*		-14.302
buenon	*						-21.868

libros	ID-MANNER	*Ssyllfinal	MAXN	ID-PLACEA	MAXA	ID-PLACEN	
libros		*					-18.608
☞ libroh						*	-11.961
libro			*				-17.777
libron	*						-21.868

Appendix C

Sentences Recorded for Experiment 1

1. Había difícile**h** obstáculo**s** ayer.

2. Había difícile**s** obstáculo**h** ayer.

3. Habrá difícile**h** obstáculo**s** mañana.

4. Habrá difícile**s** obstáculo**h** mañana.

5. Dimos mala**h** impresione**s** ayer.

6. Dimos mala**s** impresione**h** ayer.

7. Daremos mala**h** impresione**s** mañana.

8. Daremos mala**s** impresione**h** mañana.

9. Hablamos con viejo**h** profesore**s** ayer.

10. Hablamos con viejo**s** profesore**h** ayer.

11. Hablaremos con viejo**h** profesore**s** mañana.

12. Hablaremos con viejo**s** profesore**h** mañana.

13. Oímos simple**h** descripcione**s** ayer.

14. Oímos simple**s** descripcione**h** ayer.

15. Oiremos simple**h** descripcione**s** mañana.

16. Oiremos simple**s** descripcione**h** mañana.

17. Aprendimos acerca de organismo**h** invisible**s** ayer.

18. Aprendimos acerca de organismo**s** invisible**h** ayer.

19. Aprenderemos acerca de organismo**h** invisible**s** mañana.

20. Aprenderemos acerca de organismo**s** invisible**h** mañana.

21. Visitamos oficina**h** oficiale**s** ayer.

22. Visitamos oficina**s** oficiale**h** ayer.

23. Visitaremos oficina**h** oficiale**s** mañana.

24. Visitaremos oficina**s** oficiale**h** mañana.

25. Oímos discurso**h** profundo**s** ayer.

26. Oímos discurso**s** profundo**h** ayer.

27. Oiremos discurso**h** profundo**s** mañana.

28. Oiremos discurso**s** profundo**h** mañana.

29. Hablamos con muchacha**h** triste**s** ayer.

30. Hablamos con muchacha**s** triste**h** ayer.

31. Hablaremos con muchacha**h** triste**s** mañana.

32. Hablaremos con muchacha**s** triste**h** mañana.

Appendix D

Experiment 1 Text File

"ooTextFile"

"ExperimentMFC 5"

stimuliAreSounds? <yes>

stimulusFileNameHead = "Experiment_1_Tokens/"

stimulusFileNameTail = ".wav"

stimulusCarrierBefore = ""

stimulusCarrierAfter = ""

stimulusInitialSilenceDuration = 0.5 seconds

stimulusMedialSilenceDuration = 0

numberOfDifferentStimuli = 32

"1_2" ""

"2_1" ""

"3_4" ""

"4_3" ""

"7_8" ""

"8_7" ""

"9_10" ""

"10_9" ""

"13_14" ""

"14_13" ""

"15_16" ""

"16_15" ""

"19_20" ""

"20_19" ""

"21_22" ""

"22_21" ""

"25_26" ""

"26_25" ""

"27_28" ""

"28_27" ""

"31_32" ""

"32_31" ""

"33_34" ""

"34_33" ""

"37_38" ""

"38_37" ""

"39_40" ""

"40_39" ""

"43_44" ""

"44_43" ""

"45_46" ""

"46_45" ""

numberOfReplicationsPerStimulus = 2

breakAfterEvery = 0

randomize = <PermuteBalancedNoDoublets>

startText = "You will hear a series of sentences.

For each pair of sentences, choose the sentence that sounds best to you.

Click to start."

runText = "Choose the sentence that sounds best to you."

pauseText = ""

endText = "This part of the experiment has finished.

Click to continue."

maximumNumberOfReplays = 3

replayButton = 0.3 0.7 0.2 0.4 "Click here to hear the sentences again." ""

okButton = 0 0 0 0 "" ""

oopsButton = 0 0 0 0 "" ""

responsesAreSounds? <no> "" "" "" "" 0 0

numberOfDifferentResponses = 2

0.2 0.4 0.5 0.7 "Sentence 1" 40 "" "1"

0.6 0.8 0.5 0.7 "Sentence 2" 40 "" "2"

numberofGoodnessCategories = 0

Appendix E

Experiment 2 Text File

"ooTextFile"

"ExperimentMFC 5"

stimuliAreSounds? <yes>

stimulusFileNameHead = "Experiment 2 Tokens REVISED TAPERED/"

stimulusFileNameTail = ".wav"

stimulusCarrierBefore = ""

stimulusCarrierAfter = ""

stimulusInitialSilenceDuration = 0.5 seconds

stimulusMedialSilenceDuration = 0

numberOfDifferentStimuli = 56

"invisibleh100h" ""

"invisibleh95h" ""

"invisibleh90h" ""

"invisibleh85h" ""

"invisibleh80h" ""

"invisibleh75h" ""

"invisibleh70h" ""

"invisibleh65h" ""

"invisibleh60h" ""

"invisibleh55h" ""

"invisibleh50h" ""

"invisibleh45h" ""

"invisibleh40h" ""

"invisibleh0h" ""

"invisibles0h" ""

"invisibles40h" ""

"invisibles45h" ""

"invisibles50h" ""

"invisibles55h" ""

"invisibles60h" ""

"invisibles65h" ""

"invisibles70h" ""

"invisibles75h" ""

"invisibles80h" ""

"invisibles85h" ""

"invisibles90h" ""

"invisibles95h" ""

"invisibles100h" ""

"profesoreh100h" ""

"profesoreh95h" ""

"profesoreh90h" ""

"profesoreh85h" ""

"profesoreh80h" ""

"profesoreh75h" ""

"profesoreh70h" ""

"profesoreh65h" ""

"profesoreh60h" ""

"profesoreh55h" ""

"profesoreh50h" ""

"profesoreh45h" ""

"profesoreh40h" ""

"profesoreh0h" ""

"profesores0h" ""

"profesores40h" ""

"profesores45h" ""

"profesores50h" ""

"profesores55h" ""

"profesores60h" ""

"profesores65h" ""

"profesores70h" ""

"profesores75h" ""

"profesores80h" ""

"profesores85h" ""

"profesores90h" ""

"profesores95h" ""

"profesores100h" ""

numberOfReplicationsPerStimulus = 2

breakAfterEvery = 0

randomize = <PermuteBalancedNoDoublets>

startText = "For each word you hear,

choose whether you hear an 'h' or an 's' at the end.

Click to start."

runText = "Choose whether you hear an 'h' or an 's' at the end of each word."

pauseText = ""

endText = "The experiment has finished. Thank you."

maximumNumberOfReplays = 0

replayButton = 0 0 0 0 "" ""

okButton = 0 0 0 0 "" ""

oopsButton = 0 0 0 0 "" ""

responsesAreSounds? <no> "" "" "" "" 0 0

numberOfDifferentResponses = 2

0.2 0.4 0.5 0.7 "h" 40 "" "h"

0.6 0.8 0.5 0.7 "s" 40 "" "s"

numberofGoodnessCategories = 0

References

1. Boersma, Paul & Weenik, David (2010). Praat: doing phonetics by computer [Computer program]. Version 5.1.43, retrieved 1 August 2010 from http://www.praat.org

2. Boersma, Paul & Hayes, Bruce (2001). Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry*, 32, 45-86.

3. 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.

4. Cedergren, Henrietta (1973). On the nature of variable constraints. *New Ways of analyzing variation in English*, ed. by C.J. Baile & Roger Shuy, 13-22. Georgetown: Georgetown University Press.

5. Coetzee, Andries W. (2006). Variation as accessing 'non-optimal' candidates. *Phonology*, 23, 337-385.

6. Coetzee, Andries W. & Pater, Joe (2008). Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *National Language Linguistic Inquiry*, 26, 289-337.

7. Evanini, Keelan (2007). How to formalize variation: Stochastic OT models and /s/ deletion in Spanish. *LACUS Forum*, 33, 147-158.

8. Gaskell, M. Gareth, Hare, Mary & Marslen-Wilson, William D. (1998). A Connectionist Model of Phonological Representation in Speech Perception. *Cognitive Science*.

9. Harris, J.W. (1983). *Syllable Structure and Stress in Spanish. A Nonlinear Analysis.* Cambridge, Mass.: MIT Press.

10. Hayes, Bruce & Wilson, Colin (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39, 379-440.

11. Ganong, William F. (1980). Phonetic Categorization in Auditory Word Perception. *Journal of Experimental Psycholoy: Human Perception and Performance*, 6, 110-125.

12. Janson, Tore (1983). Sound Change in Perception and Production. *Language*, 59, 18-34.

13. Juilland, A. & Chang-Rodriguez, E. (1964). *Frequency Dictionary of Spanish Words*. La Haya: Mouton.

14. Labov, William (1981). Resolving the Neogrammarian Controversy. *Language*, 57, 267-308.

15. Mack, Sara (2011). A Sociophonetic Analysis of /s/ Variation in Puerto Rican Spanish. *Selected Proceedings of the 13th Hispanic Linguistics Symposium*.

16. Moreton, Elliott (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition*, 84, 55-71.

17. Morris, Richard E. (2000). Constraint interaction in Spanish /s/-aspiration: three peninsular varieties. *Hispanic linguistics at the turn of the millennium: selected proceedings from the Third Hispanic Linguistics Symposium.* Somerville, MA: Cascadilla.

18. Pater, Joe (2008). Gradual Learning and Convergence. *Linguistic Inquiry,* 30, 334-345.

19. Pierrehumbert, Joan (2001). Stochastic Phonology. *Glot International*, 5, 195-207.

20. Poplack, Shana (1980). *The Notion of Plural in Puerto Rican Spanish: Competing Constraints on (s) Deletion.* Academic Press, Inc.

21. Smith, Jennifer L. (2011). Category-specific effects. *Companion to Phonology*. Oxford: Wiley-Blackwell.

22. Smith, Jennifer L. (2001). Lexical category and phonological contrast. *Papers in Experimental and Theoretical Linguistics 6: Workshop on the Lexicon in Phonetics and Phonology,* 61-72. Edmonton: University of Alberta.

23. U.S. Census Bureau. 2000. Online <http://www.census.gov/>.

24. Yu, Alan C. L. (2010). Perceptual Compensation is Correlated with Individuals' "Austistic" Traits: Implications for Models of Sound Change. PLoS ONE, 5.