Sociophonetic Variation in the Production and Perception of Obstruent Voicing in Buenos Aires Spanish

by

Marcos Rohena-Madrazo

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Linguistics

New York University

September, 2011
Dedication

lunes, martes, miércoles: tres
jueves, viernes, sábado: seis
¡y domingo siete!

Para ti, Abuela.
Acknowledgements

First of all, I would like thank each and every one of the participants in this investigation, who selflessly donated their time, energy, and speech. Without their generosity, this dissertation would not have been possible.

I would like to express my deepest gratitude to my advisor, Lisa Davidson. Thank you for being such a patient guide, for always having your door open to discuss new ideas, for encouraging me to step back and question my assumptions, and for instilling in me a passion for experimental research. I am truly grateful for all you have taught me and am honored to have been your student.

I would also like to thank the other members of my committee: Gregory Guy, for always bringing it back to the big picture; John Singler, for delivering the harsh (and well-deserved) critiques in a way that always made me smile; Susannah Levi, for her detailed comments that provided a new perspective; and Laura Colantoni, for believing in my work and offering words of encouragement at just the right time. Your input throughout this process has undoubtedly made me a better scholar.

The inspiration for this project, I owe to Mercedes Paz. Thank you for your invaluable support and for all the wonderful moments we shared. Likewise, I would like to express my heartfelt appreciation to the Paz and Fink families for their hospitality.
and boundless generosity. Thank you also to Santiago Kalinowski, Helgalís Ramos, Giselle Román, Lourdes Lávaque, Pedro Rodríguez, Daniela Lauría and all my friends in Buenos Aires, native and transplanted, who made me feel right at home.

I am grateful to Sean Martin, Jason Shaw, Tuuli Morrill, Kevin Roon, Jocelyn Doxsey, Andrea Cattaneo, Stacy Dickerman, Carina Bauman, Vincent Chanethom, Luciana Muniz, Kyle Major, Libby Coggshall, Sang Im Lee, Jon Brennan, Suzanne Dikker, Luiza Newlin-Lukowicz, Emily Nguyen, Cara Shousterman, Eytan Zweig, Lisa Levinson, Tom Leu, Jen Nycz, Sonya Fix, Dan Lassiter, Txuss Martín, Philipp Angermeyer, Laziz Nchare, Oana Savescu, Mike Taylor, Silvia Perpiñán, Meghan Armstrong, Hilary Barnes, Eva Suárez, Ryan LaBrozzi, Ana de Prada, and the other fellow graduate students at NYU Lingustics and other departments for the hearty meals, numerous drinks, and fruitful discussions that we have shared throughout these years. Also, many thanks to the professors at NYU Linguistics for promoting such a collegial environment, in particular, to Renée Blake and Louise Vasvári for helping me uncover academic interests I never knew I had. I am also grateful to María Vaquero, who taught at the University of Puerto Rico, for being the first to usher me into the endlessly fascinating field of Hispanic Linguistics.

I owe many thanks to my research assistants, Linden Bairey, Autumn Gerami, Scott Lawrence, Sophia LeFraga, Joe Raite, Matt Sauter, Jess Tauber, and Molly Trisler, who helped me through the less glamorous aspects of phonetics research.

Thanks to David Pe and my fellow Grad Halls Staff, Bryan Brazeau, Elaine Tang, Jeff Steen, Lauren Aument, Monica Putt, Olga Sánchez-Eltell, Shavanna Calder, Theo Mitchell, and Thomas Fai, for cheering me on and making my last year at NYU so memorable in such an unexpectedly delightful way.
I am especially grateful to Jeff Crumbliss, who was there every week to help me surpass the hurdles and celebrate the successes of the dissertation process (and of life in general!). I cannot imagine what it would have been like without your help. Thank you for making me learn so much about myself.

I am very fortunate to have such a fantastically varied group of friends keeping me afloat. Many thanks to Juan Carlos Guzmán, Omar Acevedo, Joel Maysonet, Marla Torrado, Yari Otero, Carmen Otero, Chetzil Peñalverty, and to all my friends from Puerto Rico, who have supported me on and off the island. In particular, a tidal wave of gratitude goes out to Lorna Torrado for being the best pep-talker in the universe, from the first chapter to the defense. And Sara Ortiz, I will be thanking you many more times than just once a month. Many thanks to my NYC friends, Fu Chung, Andrew Solomon, Chris Mark, Emily McNeil, Jill Emmanuele, Dave Bowles, Josh Siegel, Vicky Cattani, and Pablo Ariel for keeping me from falling prey to the grad school vortex.

Thanks to Tricia Irwin for showing me the \LaTeX{} ropes, without her help this dissertation would have never been written, literally. Thanks to Kara Becker for blazing trails and showing me that it can be done. Thanks to Maryam Bakht, whose quinces and company got me through my segmentation and much more. Amanda Dye’s G-chat banter, dependably full of football and Spanishness, made the late night work sessions much more enjoyable. Gracias y ¡hala!

Special thanks to Javier Uriarte for being such a faithful friend, from across the block or across globe, through the NYU experience and beyond. We have overcome!

If I am the onset, Danny Erker is the coda. Our fricative dork-out sessions have been, without a doubt, the most intellectually stimulating (and fun!) part of the dissertation process. I look forward to many more, my friend.
I am particularly indebted to Miquel Simonet for teaching me so much about linguistics research at every level, from the most abstract theoretical questions to the most concrete scripting code. You seamlessly embody the role of friend, colleague, and mentor; for this I am deeply grateful.

One person has been a constant not just in my academic career but in my life ever since I started this roller coaster called graduate school: Amy Wong. There is no one in the world whose interests/obsessions overlap so perfectly with mine. Truly, a match made in foodie-linguistics heaven. Sharing with you offices, holidays, thoughts, theories, mooncakes, guavacakes, and countless hours of conversation has been an immense pleasure. Thank you so very much for being my best friend throughout.

Many thanks to my family: to Abuela Alicia for showering all of us with unconditional love, to Abuelo Santos for always telling me ¡usted es de los buenos!, to Giovanna, Marfa, and Cuqui for showing me that I can do it too, to Marfa del Mar for always checking in, to titi Sunchi for her illuminating cheer, to Anibal for encouraging me by means of ¡termina!, to Ana Cecilia for always wanting to talk shop, and to Larisa for being the sister I never had. I am very fortunate to have all of you in my life.

Finally, I would like to thank my parents Gloria Madrazo and Ricardo Rohena for inspiring and motivating me in the most of diverse ways. No one has been more influential in my life than you. I am eternally grateful for your unwavering love and support. ¡Los quiero infinitamente!
Abstract

This dissertation presents an instrumental study of variation in fricative voicing in Buenos Aires Spanish (BAS), particularly with respect to the devoicing change of the postalveolar fricative: /ʃ/>/ʃ/. It proposes a novel way of determining the completion of this change by comparing the percentage voicing of the postalveolar fricative to that of /s/, thus providing a system-internal benchmark for voicelessness in a language without a fricative voicing contrast. The findings from the production study show that there is still much variation in the voicing of /ʃ/ that is both socially and phonetically conditioned. Although the change has reached completion only in the younger middle class group, all other social groups show evidence that the devoicing change is still progressing.

Cross language discrimination experiments using Portuguese stimuli were conducted to explore the effects that socially conditioned variation and positionally defined allophonic variation in obstruent voicing may have on the delimitation of perceptual categories used to discriminate non-native obstruent voicing contrasts. The results from the ABX task show that, instead of improving the discrimination of Portuguese /ʃ//ʒ/, having socially differentiated variation in the native speech community actually hinders the discrimination of this contrast, regardless of the listener’s social background. Listeners, regardless of age and social class, do not seem to tap into the socially conditioned
differential distribution of [f] and [z] in BAS in order to discriminate the non-native contrast. The goodness rating task results show that listeners who maintain the voicing of /ʒ/ in their production give Portuguese /ʃ/ significantly lower ratings than Portuguese /ʒ/. However, the converse is not true: listeners for whom the /ʒ/>/ʃ/ change has been completed give similarly high ratings to both /ʒ/ and /ʃ/. Rather than being determined by social differences, the evaluation of the variation (/ʃ~ʒ/) seems to be affected by the listener’s own production. The more innovative speakers in the devoicing change seem to have a more expansive postalveolar fricative perceptual category, whereas the conservative speakers’ category has a smaller range of acceptable realizations.
Contents

Dedication iii

Acknowledgements iv

Abstract viii

List of Figures xv

List of Tables xx

List of Appendices xxiv

1 Introduction 1

1.0 Overview ................................................. 1

1.1 Background on yeismo in Buenos Aires .................. 3

1.1.1 Historical development: j > z > f .................. 3

1.1.2 The devoicing of /x/ in Buenos Aires: previous studies .. 7

1.2 Issues in fricative voicing ................................. 10

1.2.1 Aerodynamic factors in fricative devoicing .......... 10
3.1 Introduction ................................................................. 68
3.2 Background on cross-language speech perception .................. 69
3.3 Models of cross-language speech perception ......................... 70
  3.3.1 Perceptual Assimilation Model ................................. 70
  3.3.2 Speech Learning Model ............................................ 73
  3.3.3 PAM vs. SLM ....................................................... 74
  3.3.4 Sociolinguistic variation and cross-language speech perception 77
  3.3.5 Test case: voicing contrasts and voicing variation in Spanish .. 78
    3.3.5.1 Stop voicing contrast ..................................... 78
    3.3.5.2 Variable fricative voicing in Spanish ..................... 82
3.4 Goals of the study ..................................................... 83
3.5 Perception experiment 1: AX discrimination ......................... 84
  3.5.1 Materials ............................................................. 84
  3.5.2 Participants .......................................................... 85
  3.5.3 Experimental task ................................................ 87
  3.5.4 Variables and predictions ....................................... 89
    3.5.4.1 Position ...................................................... 89
    3.5.4.2 Voicing pair ................................................ 91
    3.5.4.3 Social class ............................................... 97
    3.5.4.4 Age ......................................................... 98
  3.5.5 Results and discussion ........................................... 98
    3.5.5.1 Position and voicing pair ................................. 99
    3.5.5.2 Social class and age ..................................... 103
  3.5.6 Interim conclusion ............................................... 104
3.6 Perception experiment 2: ABX discrimination ....................... 107
3.6.1 Stimulus materials ........................................... 108
3.6.2 Participants ................................................. 108
3.6.3 Experimental task ......................................... 109
3.6.4 Variables and predictions ................................. 111
  3.6.4.1 Position ............................................. 111
  3.6.4.2 Voicing pair ....................................... 112
  3.6.4.3 Social class ....................................... 113
  3.6.4.4 Age ............................................... 113
  3.6.4.5 Voicers vs. devoicers ............................. 114
3.6.5 Results and discussion ................................ 114
  3.6.5.1 Position and voicing pair ........................ 115
  3.6.5.2 Social class and age .............................. 119
  3.6.5.3 Voicers vs. devoicers ............................. 120
3.6.6 Reaction times ............................................ 121
3.7 Conclusions .................................................. 123

4 Perceptual assimilation of non-native obstruent voicing contrasts 125
  4.1 Introduction .............................................. 125
  4.2 Background ............................................... 126
  4.3 Methodology .............................................. 127
    4.3.1 Participants ....................................... 127
    4.3.2 Stimuli ............................................. 128
    4.3.3 Procedure ......................................... 128
  4.4 Predictions ............................................... 131
  4.5 Results and discussion ................................ 133
    4.5.1 Stop stimuli ..................................... 134
List of Figures

1.1 Word classes and range of realizations in Castilian Spanish (adapted from Hualde 2004:13) .............................................. 4
1.2 Palatograms of /j/ in Peninsular Spanish (Fernández Planas 2007) and of /z/ and /s/ in Argentinean Spanish (Kochetov and Colantoni 2009) . . 5
1.3 A possible gestural score for the sequence /asa/, exhibiting alignment of the tongue tip and glottal gestures . . . . . . . . . . . . . . . . 13
1.4 A possible gestural score of the sequence /asa/, exhibiting misalignment of the tongue tip and glottal gestures . . . . . . . . . . . . . . . . 14
1.5 Waveform, spectrogram, and broad transcription of the phrase a desayunar “to have breakfast” produced by a young, upper class male (from Rohena-Madrazo’s 2008 corpus) . . . . . . . . . . . . . . . . . . . . . . . . 16
1.6 Map of the City of Buenos Aires (Ciudad Autónoma de Buenos Aires), showing the neighborhood subdivisions as well as the partidos in Buenos Aires Province that are adjacent to the city limits . . . . . . . . . . . . 22
1.7 Map of the Buenos Aires metropolitan area with the administrative subdivisions or partidos . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23
2.1 Example of /ʒ/ in the word payaso, “clown,” spoken by subject 21 (younger, upper class female) in the word list task. 32
2.2 Examples of /s/ in the word acerqué, “I brought close”, spoken by subject 21 (younger, upper class female) in the word list task. 33
2.3 Waveforms and spectrograms for the word allá, “there”. Note the completely voiced realization [aʒa] to the left by subject 21 (younger, upper class, female) and the mostly voiceless realization [aʃa] to the right by subject 5 (younger, middle class, female). 37
2.4 Waveforms and spectrograms for the sequence /asa/ in the word pasajeros, “passengers”. Note the completely voiced realization [aza] to the left by subject 5 (younger, middle class, female) and the mostly voiceless realization [asa] to the right by subject 21 (younger, upper class, female). 37
2.5 Examples of both /ʒ/ and /s/ in the word embellecer, “to beautify”, spoken by subject 31 (younger, upper class male) in the sentence reading task. 38
2.6 Examples of both /ʒ/ and /s/ in the word embellecer, “to beautify”, spoken by subject 6 (younger, middle class, male) in the word list task. 39
2.7 Example of an interdental fricative realization of /s/ in the word cazaba, “s/he hunted”, spoken by subject 34 (older, upper class, male) in the word list task. 41
2.8 Example of a palatal lateral realization of /ʒ/ in the word allá, “there,” spoken by subject 34 (older, upper class, male) in the word list task. 42
2.9 Example of a palatal glide realization of the /ʒ/ in the sequence in the word embellecer, “to beautify”, spoken by subject 26 (older, middle class, male) in the word list task. 42
2.10 Example of a voiced affricate realization of /ɾ/ in the name Yolanda spoken by subject 30 (older, upper class, female) in the sentence reading task. ............................................. 43

2.11 Example of a mostly voiceless affricate realization of /ɾ/ in the word llevale, “take to him/her”, spoken by subject 26 (older, middle class, male) in the sentence reading task. ................................. 44

2.12 Example of a preaspirated realization of /s/ in the word cazaba, “s/he hunted”, spoken by subject 5 (younger, middle class female) in the word list task. ..................................................... 44

2.13 Example of a preaspirated realization of /ɾ/ in the word allá, “there”, spoken by subject 14 (younger, middle class female) in the word list task. ........................................................................ 45

2.14 Boxplots showing the percentage voicing rates of the two stimulus fricatives in three positions ................................................................. 48

2.15 Boxplots showing the percentage voicing by fricative. ‘y’ represents /ɾ/. ................................................................. 51

2.16 Average percentage voicing rates for /s/ and /ɾ/ in each phrase initial and intervocalic position. ................................................................. 51

2.17 Average percentage voicing rates for /s/ and /ɾ/ in each style. ........... 52

2.18 Average duration (in seconds) of /s/ and /ɾ/ for each style. ................ 55

2.19 Average duration (in seconds) of /s/ and /ɾ/ by position and style. ....... 55

2.20 Average percentage voicing rates for /s/ and /ɾ/ for the subjects stratified by age and social class, pooled across tasks. ......................... 58

2.21 Boxplots showing the percentage voicing for /s/ and for /ɾ/ produced by each subject, stratified by age and social class. ......................... 61
3.1 Example from Torreira and Ernestus (to appear:9) showing the measures of the syllable /to/ in the word *autopista* “highway”. IV: intervocalic voicing; CD: closure duration; VD: vowel duration. 81

3.2 Boxplots showing the sensitivity rates (d’) for the voicing pairs in initial and intervocalic positions. 100

3.3 Accuracy rates for voicing pairs in initial and intervocalic positions. 100

3.4 Boxplots showing the sensitivity rates (d’) for the voicing pairs in initial and intervocalic positions. 102

3.5 Accuracy of discrimination of non-native voicing contrasts, collapsed across position. 102

3.6 Accuracy rates for obstruent voicing discrimination by voicing pair and position in the ABX experiment. 115

3.7 Accuracy rates for obstruent voicing discrimination by voicing pair in the ABX experiment. 117

3.8 Boxplots showing the distribution of the results of the ABX experiment. 118

3.9 Reaction time results for each voicing pair in initial and intervocalic position in the ABX experiment. 121

4.1 Beanplots showing the rating distribution for each stimulus obstruent for initial and intervocalic position. 134

4.2 Distribution of the transcription responses for each of the stop stimuli collapsed across position and vowel contexts. 135

4.3 Distribution of the transcription responses for each of the fricative stimuli collapsed across position and vowel contexts. 140

xviii
4.4 Beanplots showing the rating distribution for each stimulus obstruent for initial and intervocalic position after the elimination of infrequent responses. .......................... 148

4.5 Beanplots showing the goodness rating distribution for postalveolar stimuli divided by social class. .......................... 152

4.6 Correlations of rating and age for the postalveolar fricative BP stimuli. 153

4.7 Beanplots showing the goodness rating distribution for postalveolar stimuli divided by voicers and devoicers. .......................... 154

4.8 Correlations of the listeners’ mean % voicing of /\l/ in production and their ratings of the two postalveolar fricative BP stimuli. 156
List of Tables

2.1 Demographic information of the participants in the production experiment. 35
2.2 Realizations for the stimulus obstruent /s/. 40
2.3 Realizations for the stimulus obstruent /z/. 40
2.4 Distribution of preaspirated tokens by subject. 46
2.5 Distribution of tokens by obstruent and position. The symbol “#” indicates phrase initial position and “/” indicates word initial, phrase medial position. 46
2.6 Results of a multiple linear regression presenting the factors that significantly affect the voicing of fricatives. 47
2.7 Summary of the linear mixed effect regression model adjusted to the voicing levels of the fricative, with subject as a random factor, for the all the data combined. (Intercept = /z/, sentences, VCV, older, upper class). 49
2.8 Summary of the linear mixed-effect regression model adjusted to the duration of the fricative, with subject as a random factor, for the all the data combined. 54
2.9 Summary of the linear mixed effect regression model adjusted to the voicing levels of the fricative, with subject as a random factor, for the fours subgroups divided by class and age. ..................... 58

2.10 Summary of the linear mixed effect regression model adjusted to the voicing levels of the fricative, with subject as a random factor, for the subset of devoicers. ................................. 63

2.11 Summary of the linear mixed effect regression model adjusted to the duration of the fricative, with subject as a random factor, for the subset of devoicers. ................................. 63

3.1 Schema of the Social Variant Assimilation pattern of a non-native contrast to two non-contrastive but socially distinct sociophonetic variants. 77

3.2 Demographic information of the participants in the AX experiment. . 86

3.3 Discrimination accuracy scale (with f/v as CG). ......................... 94

3.4 Discrimination accuracy scale (with f/v as TC). ......................... 95

3.5 Discrimination accuracy scale (with s/z as CG). ......................... 95

3.6 Discrimination accuracy scale (with s/z as SC). ......................... 95

3.7 Schema of the Social Variant Assimilation pattern of a non-native BP contrasts by BAS listeners. ................................. 96

3.8 A discrimination accuracy scale under Perceptual Assimilation Model. 97

3.9 A discrimination accuracy scale incorporating Social Variant Assimilation Hypothesis. ................................. 97

3.10 Discrimination accuracy scale for voicing pairs in the AX experiment. 103

3.11 Demographic information of the participants in the ABX discrimination experiment. ................................. 110

3.12 Discrimination accuracy scale (with s/z as SC). ......................... 112
3.13 Discrimination accuracy scale (with s/z as CG) ........................................ 112
3.14 Discrimination accuracy scale incorporating Social Variant Assimilation Hypothesis ................................................................. 112
3.15 Discrimination accuracy scale from the ABX experiment ....................... 119

4.1 Demographic information of the participants in the perceptual assimilation task (same as in the ABX task) ......................................................... 129
4.2 Orthographic responses for BP /p/ stimuli .............................................. 134
4.3 Orthographic responses for BP /t/ stimuli ................................................ 135
4.4 Orthographic responses for BP /k/ stimuli .............................................. 136
4.5 Orthographic responses for BP /b/ stimuli .............................................. 136
4.6 Distribution by position of <b> and <v> responses for BP /b/ ................. 137
4.7 Results of the linear mixed effect regression model adjusted to the ratings of BP /b/, with response and position as fixed factors and subject as a random factor ................................................................. 138
4.8 Orthographic responses for BP /d/ stimuli .............................................. 138
4.9 Orthographic responses for BP /g/ stimuli .............................................. 138
4.10 Orthographic responses for BP /f/ stimuli ............................................. 141
4.11 Orthographic responses for BP /v/ stimuli ............................................. 141
4.12 Distribution by position of <b> and <v> responses for BP /v/ ................. 142
4.13 Results of the linear mixed effect regression model adjusted to the ratings of BP /v/, with response and position as fixed factors and subject as a random factor ................................................................. 143
4.14 Orthographic responses for BP /s/ stimuli ............................................. 143
4.15 Distribution by position of <s> and <z> responses for BP /s/ .................. 143
4.16 Results of the linear mixed effect regression model adjusted to the ratings of BP /s/, with response and position as fixed factors and subject as a random factor. .................................................. 144
4.17 Orthographic responses for BP /z/ stimuli. ................................. 144
4.18 Distribution by position of <s> and <z> responses for BP /z/. ......... 145
4.19 Results of the linear mixed effect regression model adjusted to the ratings of BP /z/, with response and position as fixed factors and subject as a random factor. .................................................. 145
4.20 Orthographic responses for BP /f/ stimuli. ................................. 145
4.21 Distribution by position of <sh> and <y> responses for BP /f/ ........ 146
4.22 Results of the linear mixed effect regression model adjusted to the ratings of BP /f/, with response and position as fixed factors and subject as a random factor. .................................................. 146
4.23 Orthographic responses for BP /ʒ/ stimuli. ................................. 147
4.24 Distribution of <s> and <z> responses for BP stimuli /s/ and /z/. . . . 149
4.25 Results of the linear mixed effect regression model adjusted to the goodness ratings, with response and position as fixed factors and subject as a random factor. .................................................. 151
4.26 Results of the linear mixed effect regression model adjusted to the ratings of BP /f/ and /ʒ/, with subject as a random factor. ................. 155

xxiii
List of Appendices

Appendix A  Word List Reading Task  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 173
Appendix B  Sentence Reading Task  . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 177
Appendix C  AX Discrimination Task Instructions  . . . . . . . . . . . . . . . . . . . . . . . 187
Appendix D  ABX Discrimination Task Instructions  . . . . . . . . . . . . . . . . . . . . . . 194
Appendix E  Perceptual Assimilation and Rating Task Instructions and Worksheet 203
Chapter 1

Introduction

1.0 Overview

Phonetic variation is an undeniably pervasive fact of speech. The chances are infinitesimi-

mally small that two utterances of the same phrase, word, or phoneme will ever be pro-
duced in exactly the same manner. In spite of this constant variation speakers are able to
genralize from the varied input and associate the realizations into what is perceived as
the same phrase, or word, or phoneme. The nature of subphonemic variation is one of
the core questions in understanding how listeners learn and represent the sounds of their
language. Sometimes subphonemic variation is phonologically conditioned, socially or
stylistically conditioned, or phonetically conditioned. One case of subphonemic vari-
ation that has been of interest in Hispanic linguistics is that of Buenos Aires Spanish
/ʒ/, which has been described as undergoing a devoicing change to /ʃ/ since the mid-
20th century. This dissertation will present an instrumental approach to the study of
variation in fricative voicing in Buenos Aires Spanish (BAS)—particularly with respect
to the devoicing of /ʒ/—by examining the aerodynamic, articulatory and social factors
influencing the completion of this sound change. Furthermore, it will examine how the
phonological, phonetic and social variation that BAS speakers exhibit in obstruent pro-
duction might affect how they perceive obstruent voicing contrasts in another language,
thereby shedding light on how this variation influences the delimitation of phonological
categories.

The structure of this dissertation is as follows. The remainder of this chapter will
provide background on the development of /ʒ/ in BAS as well as a review of the past
studies that have examined the variation in voicing. Furthermore, the general aerody-
namic and articulatory issues underlying variation in fricative voicing will be discussed
and contextualized with respect to BAS. Lastly, a brief description will be provided of
the main questions addressed in the dissertation as well as a discussion of the fieldwork
procedures. Chapter 2 will present data from two production tasks in order to provide
a profile of the phonetic and social factors influencing sibilant variation in BAS. It will
propose a novel method of comparing the voicing rates of /ʒ/ to those of /s/ in order to
determine whether the devoicing change has reached completion. Chapter 3 will present
the results from two perception experiments where BAS listeners discriminate Brazil-
ian Portuguese obstruent voicing contrasts in order to determine how phonologically
conditioned variation (spirantization of voiced stops) and socially conditioned variation
(devoicing of /ʒ/) in the native speech community might affect cross-language percep-
tion. It will evaluate the different predictions made by two models of cross-language
speech perception (the Perceptual Assimilation Model and the Speech Learning Model)
in the light of the experimental results. Chapter 4 will present the results from a per-
ceptual assimilation and rating task, where the BAS listeners transcribe and rate the
Brazilian Portuguese stimuli. The aim of this chapter is to corroborate the assimilation
patterns inferred by the discrimination results as well as to determine whether sociolin-
guistic variation affects how speakers from different social groups assimilate and rate
the non-native phonemes. Finally, Chapter 5 will present a general discussion the results of all the experiments presented in the previous chapters, with aims to ascertain the effects that sociophonetic variation in speech production might have on variation in speech perception. The chapter will conclude with proposals for future research.

1.1 Background on yeísmo in Buenos Aires

The Spanish spoken in Buenos Aires, Argentina is one of the most recognizable varieties in the Spanish-speaking world. The principal characteristics that distinguish this variety from the rest are the exclusive use of vos as the second person singular informal pronoun, its intonation contours, and particularly the presence of a sibilant postalveolar fricative: /\!/ . This fricative corresponds to the phoneme /j/ of other varieties of Spanish in words such as llave “key” and playa “beach”. The variable voicing of this postalveolar fricative has been the focus of many investigations throughout the 20th century, which have described the process of devoicing of /\!/ to [ʃ].

1.1.1 Historical development: j > ʒ > f

Historically in Spanish there existed a phonemic distinction between the central palatal consonant /j/ and the lateral palatal consonant /ʎ/, represented orthographically as <y> and <ll>, respectively. Given this distinction, words like cayó vs. calló (“s/he fell” vs. “s/he kept quiet”) and haya vs. halla (“that there be” vs. “s/he finds”) were minimal pairs. This distinction is still maintained currently in some varieties of Spanish spoken in areas of Northern Spain, of the Andean Highlands, and of the Guaraní-influenced areas of South America. However, in the vast majority of the other varieties of Spanish

\[^1\] However, even in these zones, the /l/ vs. /ʎ/ distinction is being lost in the urban areas among the speakers of younger generations (Lipski 1994:139-140, Hualde 2005:44).
Figure 1.1: Word classes and range of realizations in Castilian Spanish (adapted from Hualde 2004:13)

<table>
<thead>
<tr>
<th>Word</th>
<th>Realizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>hiato</td>
<td>iá ~ já</td>
</tr>
<tr>
<td>hiena</td>
<td>jé ~ jé</td>
</tr>
<tr>
<td>yema</td>
<td>jé ~ jé ~ jé</td>
</tr>
</tbody>
</table>

the /j/ and /ʝ/ have undergone a merger, whereby /ʝ/ > /j/. In Hispanic linguistics this phenomenon is known as yeísmo. The resulting phoneme, /ʝ/, is highly variable with respect to its realizations both within and across different dialects of Spanish. It exhibits realizations ranging from a glide [j], to a fricative [ʝ], to a stop [ʝ] and even to an affricate [ʝʝ], depending on factors such as dialect zone, phonological position, or speech style (Hualde 2005:165). In fact, Hualde (2004) classifies the /ʝ/ as a quasi-phoneme, since it could be analyzed as an allophone of unstressed /i/ in syllable initial position, evidenced by the variation in the production of words such as hielo “ice” ([iélo]∼[jélo]). This means that that in spontaneous speech there can be a merger between the supposed minimal pairs of hiena /ʝéna/ “hyena” and llena /ʝéna/ “full”. The variable range of overlapping realizations of the quasi-contrastng /i/ and /ʝ/ are shown in Figure 1.1. The transcriptions of the first sounds in each of the three words hiato, hiena, and yema show that each has a different set of possible alternating realizations within the same variety (or speaker); however, all three sets share [j] as a variant.

In Buenos Aires Spanish, the /ʝ/ underwent a process of articulatory fortition, which resulted in a voiced postalveolar sibilant realization: /ʝ/. In Hispanic linguistics this phenomenon is known as žeísmo or yeísmo rehilado and is one of the characteristic

---

2It is not necessarily the case that all varieties of Spanish that exhibit yeísmo have /ʝ/ as the underlying phoneme. Notable examples are žeísmo, as will be discussed here, as well as Northern Mexican Spanish where the etymological /ʝ/ is realized consistently as a palatal glide [j] (Hualde 2005:166).

3Navarro Tomás (1934:274) defines rehilamiento as “relatively intense and resonating vibration with which some sounds are articulated” (“vibración relativamente intensa y resonante con que se producen ciertas articulaciones”); in other words, it can be defined as turbulent airflow that is produced at the point of constriction of some fricatives (Martínez Celdrán and Fernández Planas 2007:53-55).
Figure 1.2: From left to right, palatograms of /ʝ/ in Peninsular Spanish (Fernández Planas 2007) and of /ʒ/ and /s/ in Argentinean Spanish (Kochetov and Colantonii 2009). The darker squares indicate higher degree of tongue-palate contact.

aspects of the Spanish spoken in the Río de la Plata area. In this case the quasi-phonemic status of /ʝ/ is no longer applicable, since the possible mergers in other varieties are clearly minimal pairs in BAS: *llena* [zena] vs. *hiena* [jéna]. Figure 1.2 shows the point of constriction of the BAS sibilant /ʒ/ as being anterior to the palatal /ʝ/ of peninsular Spanish and posterior to the /s/ of BAS, demonstrating the postalveolar, or prepalatal, realization in Argentinean Spanish. The development of this change will be discussed in more detail below.

In Argentina, as in most of the Spanish-speaking world, there was a merger of /ʝ/ and /ʎ/. The first documents from Buenos Aires that present evidence of this merger, by orthographic confusion between <y> and <ll>, i.e. *yeismo*, date back to 1720 (Fontanella de Weinberg 1987:163). In the mid-18th century the phoneme /ʝ/ undergoes a process of fortition and begins to be realized as [ʃ]. According to Fontanella de Weinberg (1973:342), the first documentation of this *yeismo rehilado* is found in “El amor de la estanciera”, an anonymous *sainete gauchesco* or popular comedic play depicting gaucho life, dating from the last decade of the 18th century. In this *sainete* there is a character who speaks Spanish with a Portuguese accent and whose speech is transcribed as accented using Spanish orthography. The Portuguese word *gente* /ʒêtel/ “people” is transcribed in Spanish orthography as <yente>. Fontanella de Weinberg (1973:342) claims
that the notation of <y> for the Portuguese phoneme /ʒ/ shows, “without a doubt,” the existence of rehilamiento in the pronunciation of Buenos Aires Spanish of the period (“muestra, sin duda, la existencia del rehilamiento en la pronunciación bonaerense de la época”). Upon more careful examination, this is not necessarily unequivocal evidence, since speakers of other non-żeista varieties of Spanish would probably assimilate Portuguese /ʒ/ to Spanish /j/, which could only be rendered <yente> in Spanish orthography. For this reason, Fontanella de Weinberg’s evidence as to the earliest attestation remains inconclusive.

The realization of /j/ as [ʒ] spread through Buenos Aires and by the late 19th century had become generalized. In fact, Maspero (1875:55) remarks that in the Spanish of the Río de la Plata basin the <y> preceding a vowel is usually pronounced like the <j> in French (“Y devant une voyelle se prononce d’ordinaire comme notre J français”). This fact unmistakably confirms that, at least by the end of the 19th century, the phoneme /j/ was pronounced [ʒ], since in French orthography the letter <j> can only represent the phoneme /ʒ/.

By mid-20th century, when ʒeismo had already been incorporated into the Spanish of Buenos Aires, several studies were published describing the Spanish spoken there (Zamora Vicente 1949, Malmberg 1950, Guitarte 1955, among others). These earlier studies make mention of the fact that this relatively new phoneme /ʒ/ is sometimes devoiced and realized as [ʃ]. Several studies conducted towards the latter half of the 20th century (Wolf and Jiménez 1979, Fontanella de Weinberg 1978) found that this innovation of devoicing of /ʒ/ exhibited characteristics of a change in progress.

Labov (1972, 2001) describes two types of change in progress: change “from above” and change “from below”. The key difference between them, at least for the purposes of this investigation, is that in changes “from above” the innovation tends to be imported
from outside the speech community and the speakers are conscious of the innovation. The innovation is associated with prestige; therefore, speakers present higher rates of use of the innovation in more careful styles. On the other hand, in changes “from below” the innovation arises from spontaneous changes in the linguistic system and the speakers of the community tend not to be conscious of the innovation, for this reason they do not usually show stylistic shift. Changes “from below” tend to show a curvilinear pattern in their distribution by social class, where the central social strata show the highest use of the innovation. Both of these changes are usually led by younger, middle class females.

The voicing variation in /s/ and its dispersion through the Buenos Aires speech community has been shown to be subject to social factors such as age, sex, and socio-economic class. For example, words like ayer “yesterday” vary between [aʃeɾ] and [aʃer] and llama “flame” can be realized as either [ʃama] or [ʃama], depending on the speaker’s social background. Differing accounts on this phenomenon as a change in progress will be discussed below.

1.1.2 The devoicing of /s/ in Buenos Aires: previous studies

The devoicing of /s/ in Argentinean Spanish was first mentioned in Alonso and Rosenblat’s notes in Espinosa’s study (1930:130), where they mention that the phoneme /s/ has as emphatic variants [dʒ] and [ʃ]. Later on, however, Zamora Vicente (1949) documents the devoicing change and contradicts the claims of [ʃ] being an emphatic variant of the phoneme /s/, claiming instead that [ʃ] is a characteristic realization among speakers of “middle cultural level” and that it is quite widespread in Buenos Aires. Guitarte (1955) adds that the voiceless variant is more prevalent in the speech of women and he agrees with Zamora Vicente in that it is common among the middle social strata. (Malmberg 1950:107), in his meticulous study of the phonetics of Argentinean Spanish, mentions
that his informants often produced “imperfect” voicing when pronouncing /ʒ/; however, contrary to Zamora Vicente’s findings, in his entire study Malmberg only noted three cases of the voiceless variant [ʃ].

Towards the end of the 20th century, several studies document the advancement of the devoicing of /ʒ/ not only in the capital (Wolf and Jiménez 1979, Wolf 1984), but also in the provinces of Buenos Aires (Fontanella de Weinberg 1978) and Santa Fe (Donni de Mirande 1992). Fontanella de Weinberg (1978) studies the devoicing of /ʒ/ in Bahía Blanca in the province of Buenos Aires, focusing on the phonological restructuring of the palatal system. The author finds that the phonological distinction between /ʒ/ and /ʃ/ is being lost (/ʃ/ being the phoneme for foreign loanwords such as show and shorts) for speakers who devoice /ʒ/ as well as for those speakers who voice /ʃ/. Fontanella de Weinberg (1978, 1983) finds stylistic effects differentiated by education level, where the speakers with a primary or university level education decrease their use of [ʃ] in more careful styles, while the speakers with secondary education increased their use of voiceless variant in the more careful styles, regardless of the fact that, according to the author, [ʃ] was not a prestigious variant. In addition, she also finds that the devoicing change is led by younger speakers, particularly females, which is a characteristic trait of changes in progress.

Wolf and Jiménez also study the devoicing of /ʒ/, but in the city of Buenos Aires, where they find results similar to those found by Fontanella de Weinberg: the factors of age and sex are more important in determining the use of either the voiced or voiceless realization of the phoneme /ʒ/. They find that the women younger than 30 are leading the change, followed by the men, who lag behind by a generation. Wolf and Jiménez (1979) report that the effect of social class is minimal and that speaker who devoice /ʒ/ are not conscious of the devoicing phenomenon. Furthermore, they conclude that the
devoicing seems to be a change in progress, and for some groups it might be nearing completion.

Wolf (1984) conducts a real time study with some of the speakers who participated in the Wolf and Jiménez (1979) study and she finds that for some younger speakers, the devoicing change seems to have been completed. However, she also finds that upper class speakers appear to be re-voicing their /ʃ/. In other words, the younger male speakers, who had shown some moderate usage of [ʃ] in the 1979 study, exhibited the voiced variant [ʒ] almost exclusively in the 1984 study. Apparently, these “re-voicers” were the only ones that were consciously aware of the change from /ʃ/ > [ʃ] and avoid it to maintain a social distinction between the upper class and the other social strata. For this reason, Wolf (1984) proposes that the devoicing of /ʃ/ has already been completed in the youngest generation and the re-voicing, as a type of age-grading, is evidence of stable variation in the speech community.

In the 21st century two new studies (Chang 2008, Rohena-Madrazo 2008) have been conducted using instrumental acoustic analysis in order to investigate the progress of the devoicing of /ʃ/ in BAS. Chang (2008) studies the devoicing levels of /ʃ/ and observes that the change is already quite generalized in the younger generation. Chang uses categorical methods (auditory categorization) as well as gradient methods (percentage of voicing and relative intensity) in order to measure the voicing of the fricative in a reading task, where participants read a comic strip. He finds that the speakers who were born before 1945 produce voiced realizations predominantly, whereas the speakers who were born after 1975 produce voiceless variants almost exclusively. The speakers who were born during the intervening period comprised between these years show variability in their devoicing of /ʃ/. Chang also investigates the effect of neighborhood of origin, but he does not find a significant correlation with the devoicing. Chang (2008:26) does not
find an effect of sex: young speakers, male or female, produce [j] consistently: “speakers in the youngest age bracket are all approaching a ceiling of 100% devoicing,” which seems to suggest that for them the change from /ʒ/ to /ʃ/ has already been completed.

Rohena-Madrazo (2008) also used instrumental measures in order to determine the devoicing of /ʒ/ in BAS. The data were collected from sociolinguistic interviews and the devoicing of /ʒ/ was determined by examining the voicing bar in the spectrogram as well as the periodicity in the waveform during the duration of the fricative segment /ʒ/. This study confirmed that the men in the younger generation have already caught up with the women in terms of the use of the innovative variant; therefore, for the majority of the younger speakers the postalveolar fricative had already been devoiced. Furthermore, contrary to the findings of Chang (2008), Rohena-Madrazo found that area of residence as an index of social class was a significant factor in the advancement of the change. The neighborhood divisions in this study, as compared to Chang’s (2008), perhaps reflected more clearly the socio-economic divisions in the city (see Section 1.3.1.2 below). Rohena-Madrazo (2008) found that among the older speakers, the middle class speakers were in the lead of the innovation, ahead of both the upper and working class, which is a characteristic of changes “from below”. However, the difference in voicing between the working and middle class is lost among the speakers of the younger generation, which is further evidence that the devoicing of /ʒ/ could be nearing completion.

1.2 Issues in fricative voicing

1.2.1 Aerodynamic factors in fricative devoicing

If one considers the articulatory and phonological processes involved in the production of a voiced fricative, then the change of /ʒ/ > [ʃ] in BAS is phonetically motivated.
According to van den Berg’s (1958) Myoelastic Theory of Speech Production, phonation is a result of differential air pressure between the subglottal and supraglottal cavities, whereby high pressure air in the subglottal cavity passes through the vocal folds into the lower pressure supraglottal cavity to equalize the pressure, thus making the vocal folds vibrate. Therefore, the oral air pressure needs to be as low as possible in order to ensure optimal voicing. Fricatives, on the other hand, require turbulence, which is attained by maintaining an air pressure differential of low air pressure at the constriction point and high air pressure in the oral cavity behind the constriction point in order to ensure maximal velocity of air passing through the constriction point, thus producing turbulence (Ohala 1997). Therefore, the oral air pressure needs to be as high as possible in order to ensure optimal frication. Consequently, the requirements for frication are at odds with the requirements for voicing. If impedance at one site reduces airflow substantially at the other, one or the other component of the speech signal is likely to be lost: too much glottal impedance impairs frication, too much oral constriction impairs voicing (see Hayes and Steriade 2004, Stevens et al. 1992).

According to Ohala (1983), these conflicting articulatory requirements are what make voiced fricatives relatively rare segments cross-linguistically. Given these reasons, the devoicing of the voiced postalveolar fricative in BAS is an expected change, particularly in a language like Spanish, where there is no phonemic voicing opposition in fricatives. In fact, a similar type of devoicing change already occurred in the history of Spanish.

Medieval Spanish possessed a full inventory of sibilants distinguished by voicing: /ʃ, s, š/ vs. /ʒ, z, ʒ/. These differences were all phonemic in Medieval Spanish; however, during the late sixteenth century this distinction was lost in favor of the voiceless variants, resulting in only three distinct phonemes: /ʃ, s, š/ (Penny 2000:43-45). These
three fricatives underwent yet another change, resulting in the modern Castilian lingual fricatives: /ʃ, s, ʂ/ > /x, s, θ/. Of these fricatives, the only ones present in Latin American Spanish are /x, s/. The voicing contrast in fricatives is preserved, to a certain extent, in most other modern Romance languages (Penny 2000). Presently, perhaps with the exception of Judeo-Spanish, there are no varieties of Spanish that present a phonemic voicing distinction like the one that existed in Medieval Spanish. Ohala (1993) argues that the fact that synchronic changes resemble diachronic changes is not due to teleological phonological pressures, but rather that the causes for phonetic change depend on the interaction of physiological and aerodynamic factors in speech production, which are essentially the same now as they were centuries ago.

However, it is not the case that the aerodynamics of the vocal tract is the only factor that could influence the voicing variation in fricatives. There are articulatory factors that could result in fricative voicing as well.

1.2.2 Articulatory factors in fricative voicing

One of the major drawbacks of traditional phonological notation is that it represents speech in discrete units that correspond to certain combinations of articulatory parameters: place of articulation, manner of articulation, voicing, etc. In a phonemic transcription, the boundaries between one segment and the next are clearly delimited. However, this representation completely obscures the continuous and gradient nature of the speech signal and of the articulatory gestures that produce this signal. A great deal of phonetic variability is to a large extent due to coarticulation, that is, “the pervasive, systematic,

---

4 Some varieties of Spanish do have voiced allophones of coda /s/, mainly non-aspirating varieties, so that whenever a coda /s/ is followed by a voiced consonant it is realized as [z], e.g. desde [dezhe] “from”, mismo [mizmo] “same” (Hualde 2005). However, this phenomenon is not productive in Buenos Aires Spanish since /s/ is usually aspirated (/s/ > [h]) in preconsonantal coda position (Lipski 1994). Also, see Bradley and Delforge (2006) for evidence from Highland Ecuadorian Spanish of an apparent voicing distinction in word boundary position: has ido [az iðo] “you have gone” vs. ha sido [a sīðo] “it has been”.

12
reciprocal influences among contiguous and often non-contiguous speech segments” (Farnetani and Recasens 1999:31).

If, rather than a segmental approach, one assumes an approach along the lines of Articulatory Phonology (Browman and Goldstein 1989, 1990, 1992), whereby neither segments nor features but a constellation of continuous articulatory gestures are encoded in the phonological representation of speech sounds, then coarticulation is the natural result of the offset of one consonantal or vocalic gesture overlapping with the onset of the next. Figure 1.3 shows an example of the gestural score of the sequence /asa/, with four supralaryngeal articulators (velum, tongue body, tongue tip, and lips) and one laryngeal articulator (glottis). The colored bars represent the timing and duration of the gestures relative to each other.

(Browman and Goldstein 1992:157) assume that “the larynx is positioned appropriately for voicing unless otherwise instructed,” therefore, they only posit a devoicing gesture for the laryngeal articulator. In Figure 1.3, this devoicing gesture is represented by “wide” on the glottal tier. In this gestural score for the sequence /asa/ there is one devoicing gesture, resulting in three distinct periods of laryngeal activity: a period of
voicing, followed by a period of voicelessness, which is then followed by another period of voicing. The devoicing gesture in the glottal tier is aligned with the tongue tip gesture both at the onset and offset, thus resulting in a voiceless fricative. In a prototypical situation, the laryngeal voicing gestures are produced in synchrony with the other supraglottal gestures corresponding to the vowel or the fricative. However, there is experimental data showing that speakers actually show a significant amount of gestural misalignment (e.g. Browman and Goldstein 1990, Fowler and Salzman 1993, Zsiga 1994; see Farnetani 1997 for an overview of connected speech processes). The gestural misalignment could be the result of overlap between two or more gestures, spatio-temporal reduction of one gesture, or a combination of these two factors (Jun 1996). If the timing for the laryngeal devoicing gesture is misaligned with the onset or offset of the consonantal gestures, then the voicing might carry over into the articulation of the fricative or begin before the fricative articulation is completed. This would result acoustically in a voiced or partially voiced fricative, as seen in Figure 1.4.

Figure 1.4 shows that the cause of the misalignment of the devoicing gesture with the tongue tip gesture is due to a reduction in time of the laryngeal gesture both at the

Figure 1.4: A possible gestural score of the sequence /asa/, exhibiting misalignment of the tongue tip and glottal gestures
onset and at the offset. This temporal reduction of the laryngeal gesture, resulting in the voicing of the intervocalic consonant, could be included under the typology of lenition. Kirchner (1998, 2004) describes lenition as a process where a certain consonant becomes “weaker,” i.e. it undergoes some sort of reduction in constriction degree or duration. This description of lenition is tightly linked with the concept of articulatory effort, since the resulting sound should exhibit greater ease of articulation with respect to the original sound. For example, Westbury and Keating (1986) argue that it is more difficult to stop vocal fold vibration during a singleton stop closure than it is to continue throughout the obstruent segment. Kirchner (1998:55) describes the same tendencies for singleton voiceless fricatives as well. In these cases, ease of articulation is characterized by the minimization of articulatory change during an utterance. Since voicing is the default speech-ready state, then to realize a devoicing gesture necessarily involves an articulatory change, thus making a voiceless obstruent harder to produce in intervocalic position. Given these tendencies, one possibility could be that the devoicing gesture in /asa/ is reduced to such an extent that it is not realized at all. In fact, elision of the devoicing gesture would be the preferred option under this articulatory account. Furthermore, the rates of reduction, gestural misalignment, and elision can also be affected by changes in speech style or rate of speech (Browman and Goldstein 1990, Byrd 1996) or automatization of gestural production in high frequency lexical items (Bybee 2002).

From a historical linguistics standpoint, intervocalic voicing of obstruents is a very common phenomenon (Campbell 1998:41), as in the transition from Latin to several Western Romance languages, where etymologically voiceless obstruents became voiced in intervocalic position, e.g. *vita* > [vida] “life,” *casa* > [kaza] “house/cottage” in Portuguese. However, there are synchronic examples of spontaneous voicing of voiceless obstruents as well. In the data from the corpus collected in Rohena-Madrazo (2008),
there was certainly much voicing variation in /\z/; however, there were instances where phonologically voiceless fricatives presented variation in voicing as well.

The phrase *a desayunar* “to have breakfast”, contains two fricatives, identified in broad phonetic transcription as /s/ and /\z/. The alveolar fricative /s/ in Spanish is a phonologically voiceless fricative; however, the pitch tracker, based on measurements of F0, indicates that there is voicing throughout the duration of the fricative. In fact, in this example /s/ seems to exhibit more voicing than /\z/. Similar processes of intervocalic fricative voicing occur variably in southern dialects of Italian, where /f, s/ > [v, z] (Nocchi and Schmid 2007). Also, for English, Stevens et al. (1992:2984) report that some phonologically voiceless fricatives sometimes exhibited residual glottal vibration during the fricative portion adjacent to the vowel. Cases of voicing of voiceless obstruents have also been found for stops in several varieties of Spanish, especially in casual speech (Machuca Ayuso 1997, Lewis 2001, Parrell 2010, Torreira and Ernestus
to appear). These lenition processes of voicing (and even spirantization) of voiceless stops occur despite the fact that stops in Spanish, unlike fricatives, do present a voicing contrast.\(^5\)

There are contradictory expectations between the aerodynamic account, which would predict voiceless fricatives, and the articulatory account, which would predict voiced fricatives, at least intervocally. It is clear from the examples discussed above, that the voicing of obstruents can be highly variable. With respect to the variation in BAS /\(z\)/, when previous studies talk about “devoicing” or completion of the change, it is not clear what these terms refer to exactly. Once speakers have reached ceiling levels of devoicing, as Chang (2008:62) mentions, there is no doubt that the change has been completed. However, it should be noted that the fact that /s/ is a phonologically voiceless fricative does not necessarily imply that it exhibits 0% voicing. As was shown in, voiceless obstruents in Spanish can be realized occasionally with partial or complete voicing (see also Martínez Celdrán and Fernández Planas 2007:69-84, Martínez Celdrán 2009, Parrell 2010, Torreira and Ernestus to appear). By extension, it should not be taken for granted that a speaker, for whom the devoicing of /\(z\)/ has reached completion, will exhibit 0% voicing in all cases of /\(z\)/. For this reason, examining the voicing of /\(z\)/ against a case of unmarked voicing like /s/ would be more informative than measuring the absolute levels of voicing for /\(z\)/ alone. Out of the voiceless fricative inventory of BAS /f, s, x/, using /s/ as a control against which to compare /\(z\)/ seems most appropriate, since /s/ is most similar to /\(z\)/ both articulatorily and acoustically; furthermore, /s/ is a phonologically voiceless fricative, whose voicing levels have not been associated with other social factors.

---

\(^5\)For a more comprehensive description of the allophonic variation in the stop voicing contrast in Spanish, see Section 3.3.5
1.3 The present study

These facts about obstruent voicing raise questions about the nature of the voicing variation in Buenos Aires Spanish sibilants. There have been numerous studies of the socially conditioned voicing variation in /ʒ/ in Buenos Aires Spanish, indicating that it is undergoing a devoicing change. However, from the examples discussed in the previous section, that there are articulatory and aerodynamic reasons why a fricative might be realized voiced or voiceless, which are independent from its underlying voicing specifications. Therefore, even the phonologically voiceless fricatives show some variation in voicing as well. How might the socially conditioned variation interact with this phonetically conditioned variation? If phonetically voiceless fricatives do exhibit phonetic voicing variation, then this variation needs to be taken into account in order to determine whether the devoicing of /ʒ/ has reached completion. Rather than assuming 0% voicing as the endpoint of the devoicing change, the voicing levels of a phonologically voiceless fricative, for example /s/, can serve as a criterion for determining the completion of the devoicing of /ʒ/. In Chapter 2 a production experiment will be presented which will address these issues, by examining the voicing variation in /ʒ/ and /s/. It will determine whether the voicing variation of /ʒ/ is affected in the same manner as /s/ by contextual and stylistic factors. Furthermore, in addition to presenting a sociophonetic profile of the distribution of the voicing variation of /ʒ/ in the BAS speech community, it will also use the voicing levels of /s/ in order to attempt to determine the groups in the community for whom the devoicing of /ʒ/ has reached completion.

Assuming that there still is socially conditioned voicing variation of /ʃ~ʒ/ in the BAS speech community, how might the speakers categorize or represent this variation? In BAS it has been reported (see Wolf 1984, Colantoni 2008, King 2009) that some speakers are aware of this variation and differentiate between [ʒ] and [ʃ] along social
lines. If so, is the representation of these socially distinct categories similar to that of phonological or phonetic categories? One way to attempt to answer this question is by means of cross-language speech perception. Current theories of cross-language speech perception (the Perceptual Assimilation Model and the Speech Learning Model) propose that listeners do not perceive non-native contrasts in a native-like fashion because of the influence of their own sound systems. However, how might sociolinguistic variation shape the influence of their native sound system? In BAS, will the variation in /ʃ-ʒ/ affect the perception of an analogous non-native voicing contrast /ʃ/-/ʒ/? Will it be different from the discrimination of a similar non-native voicing contrast like /s/-/z/, where the two contrasting non-native phonemes are not associated with social variation? Will differences in listeners’ production of /ʒ/ affect how they rate a /ʃ/-/ʒ/ contrast in a non-native language? Furthermore, how might the effects of social variation on category delimitation differ from the effect that native allophonic variation might have on the delimitation of categories? Will listeners have difficulty distinguishing a native voicing contrast in a non-native position?

In order to answer these questions, several perception experiments were conducted. In Chapter 3 the results of two cross-language experiments will address these questions of how BAS listeners might represent phonological, phonetic and social variation and how this representation affects the discrimination of non-native voicing contrasts. The experiment in Chapter 4 will corroborate the assimilation patterns of the non-native contrasts to the native system and how the goodness of fit ratings vary depending on phonemic, phonetic and social factors of BAS. The results from a perceptual assimilation and rating task will be presented to test the predictions made by the different models of cross-language speech perception (PAM and SLM) as well as to elucidate the degree of awareness that BAS listeners might have on the devoicing of /ʒ/.
1.3.1 **Fieldwork procedures**

The experiments that will be presented in the following chapters were conducted in two different field trips to Buenos Aires, Argentina: one in 2008 and the other in 2009. In 2008 the only experiment conducted was the AX discrimination task, which will be discussed in Chapter 4. In 2009 the administered experiment included the production tasks discussed in Chapter 2 (word list and sentence list), as well as the ABX discrimination task and the perceptual assimilation/rating task discussed in Chapter 3 and Chapter 4. It should be noted that the participants whose production data will be analyzed in Chapter 2 were a subset of those whose data will be analyzed in Chapter 3 and Chapter 4. Aside from the word list task and the sentence reading task, other production tasks were administered as well; however, they will not be reported on in this dissertation and will be reserved for future investigation. All the production tasks preceded the ABX discrimination task, which was then followed by perceptual assimilation and rating tasks.

1.3.1.1 **Participant recruiting and social classification**

The participants in the production and perception experiments were all 18 years of age or older and were native speakers of Spanish. None of them reported hearing or speech problems. They were all born and raised in the Buenos Aires metropolitan area and lived there at the time of the data collection. The participants were recruited using the “friend of a friend” method (Milroy 1980), starting from the investigator’s social network in Buenos Aires. All of the speakers and listeners in the experiments participated voluntarily without monetary compensation. The investigator is very grateful to them for having donated their time and energy.

The main language-independent social factors that will be considered in the experiments presented in this dissertation will be age, sex, and social class. While the cate-
gorization of the speakers by age and sex is relatively straightforward, the classification of speakers into socioeconomic categories is more problematic. There is a multitude of factors that could determine someone’s socioeconomic position such as education, income, place of residence, occupation, race, ethnicity, etc. (see Guy 1988). Since place of residence has been an important component in the determination of social class in past investigations, as well as the present one, it is necessary to examine some demographic facts of Buenos Aires.

1.3.1.2 Area of residence as a proxy for social class

Buenos Aires is the capital of Argentina and its largest city. The Buenos Aires metropolitan area is known as the Conurbano Bonaerense, which is comprised of the city of Buenos Aires proper (Ciudad Autónoma de Buenos Aires) and Greater Buenos Aires (Gran Buenos Aires). According to the provisional results of the 2010 census, the population of the city of Buenos Aires is 2,891,082 and the population of the 24 partidos\(^6\) of Greater Buenos Aires is 9,910,282, for a total metropolitan population of just over 12.8 million. (Instituto Nacional de Estadística y Censos: http://www.censo2010.indec.gov.ar/preliminares/cuadro_totalpais.asp). Figure 1.6 shows a map of the City of Buenos Aires and Figure 1.7 shows a map of the Buenos Aires metropolitan Area, with the City of Buenos Aires and the Greater Buenos Aires areas indicated. As mentioned before, one of the requisites for participation in the experiments in this investigation was to have been born and raised in the Buenos Aires metropolitan area.

During the 20th century there has been differential development of the northern zone of the Buenos Aires metropolitan area, particularly compared to areas of the south and

\(^6\)The partidos are intra-provincial administrative subdivisions of the Province of Buenos Aires.
Figure 1.6: Map of the City of Buenos Aires (*Ciudad Autóctona de Buenos Aires*), showing the neighborhood subdivisions as well as the partidos in Buenos Aires Province that are adjacent to the city limits (Source: Mapa de la Ciudad Autónoma de Buenos Aires - http://mapa.buenosaires.gov.ar)
Figure 1.7: Map of the Buenos Aires metropolitan area with the administrative subdivisions or partidos (Source: http://commons.wikimedia.org/wiki/File:Great_buenos_aires.png)
This development has resulted in increased property values and has attracted more affluent, higher status residents (Ciccolella 1999, Torres 2001). Furthermore, the north-south social distinction is quite salient in the Buenos Aires community. Several participants from Rohena-Madrazo (2008) brought up this topic in their sociolinguistic interviews. One participant, a 60-year-old male from Southern Buenos Aires, provided a very relevant example regarding his and his wife’s place of origin: *Los dos somos del sur, eso tiene una cierta connotación...sociocultural. La diferencia norte-sur existe.* ("We’re both from the south; that carries a certain sociocultural connotation. The north-south difference exists.").

Rohena-Madrazo (2008) devised a socioeconomic index taking into consideration factors such as education, occupation, and place of residence. The speakers in that study patterned into three groups: (I) professionals with post-secondary education (or post-secondary students) living in Northern Buenos Aires, (II) middle professionals with post-secondary education (or post-secondary students) living in Southern Buenos Aires, and (III) manual laborers with elementary-school education living in Southern Buenos Aires. These three groups were identified as upper, middle, and working class, respectively. Rohena-Madrazo (2008) found that the distribution of the devoicing of /s/ in these three groups showed a curvilinear pattern, characteristic of changes “from below”, where the middle social stratum shows higher rates of the innovation. However, in the younger generation, the difference in devoicing rates between the middle class and the working class was eroded; in other words, the differences in education and occupation between these two social classes did not seem to be significant factors affecting the devoicing of /s/. For the younger generation, there resulted two social class groups that differed in their devoicing rates: the upper class and the non-upper class. The component of the socioeconomic index that distinguished these two groups was area of residence.
(north vs. south), which suggests that area of residence, rather than education or occupation, seems to be emerging as the relevant social class indicator differentiating the group that disfavored devoicing of /s/ (northerners/upper class) and the group that favored the devoicing innovation (southerners/middle+working class).

The inclusion of three socioeconomic strata (upper, middle, and working class) in Rohena-Madrazo’s (2008) investigation allowed for the investigation of changes in the curvilinear pattern of the devoicing of /s/, which would definitely be fruitful to expand upon in the current investigation. However, because of differences in the methodology between the two studies, it was not feasible to include the working class group. During the data collection for Rohena-Madrazo (2008), as part of the sociolinguistic interviews, participants were asked to read a passage and a word list. Reading any text aloud is already something that many adults are unaccustomed to; however, during these reading tasks, the working class participants had considerably more difficulties reading the text and words, and many were visibly uncomfortable during the task. The current investigation does not include an interview, and the reading tasks are much longer and involved than the ones used in Rohena-Madrazo (2008). This would have increased the level of discomfort in the working class participants.

Taking into consideration these methodological differences, only upper and middle class speakers were recruited in for the experiments in this investigation. Based on the saliency of the differential socio-geographic profile discussed above and on the findings from Rohena-Madrazo (2008), for the purposes of this study, area of residence was used as a proxy for the speakers’ social class: speakers from northern Buenos Aires were considered to be upper class and the speakers from southern Buenos Aires were considered to be middle class.
The northerners/upper class participants were from the City of Buenos Aires barrios of Recoleta, Retiro, Palermo, Villa Crespo, and Belgrano; and from the Greater Buenos Aires partidos of Vicente López (Olivos, Munro, Florida), San Isidro (San Isidro, Acasuso, Martínez, Béccar), and Tigre (Nordelta, Don Torcuato). The southerners/middle class participants were from the City of Buenos Aires barrios of Once (Balvanera), Caballito, and San Telmo; and from the Greater Buenos Aires partidos of Almirante Brown (Adrogué, Burzaco, Mármol), Ezeiza (Ezeiza), San Vicente (Alejandro Korn). All of the participants in both groups had completed secondary education and those who were not university students worked or had worked in professional, administrative, or educational positions. The individual demographic information for the relevant participants will be provided in the discussion of each experiment.

The following chapter will investigate the effect of these social factors, as well as phonetic factors, on the voicing variation of /ʒ/ in order to determine whether the devoicing change has been completed in the BAS speech community.
Chapter 2

Variation and change in sibilant voicing in Buenos Aires Spanish

2.1 Introduction

This chapter presents an instrumental acoustic analysis of the realization of /ʒ/ in Buenos Aires Spanish (BAS) and of the social and phonetic factors that condition its realization. Comparing the voicing variation of /ʒ/ to those of the phonologically voiceless fricative /s/ will provide a metric in order to determine whether the change is still in progress or has reached completion.

Recall that the devoicing of /ʒ/ in Buenos Aires Spanish had been described by Wolf and Jiménez (1979) as a classic example of change in progress, where the younger females were in the lead, with the younger males lagging behind by a generation. Wolf (1984) conducted a follow up study with some of the participants from Wolf and Jiménez (1979) and found that the younger upper class speakers who had shown some rates of devoicing of /ʒ/ in 1979 showed considerably lower rates or devoicing (or none at all) when restudied in 1984. Wolf reported that these upper class speakers had negative
evaluations of the devoicing of /ʃ/ and were utilizing their higher rates of /ʃ/ voicing as a way to differentiate themselves from the middle class. Given these findings, Wolf (1984) argued that the devoicing of /ʃ/ was no longer a change in progress, but actually a case of stable variation employed to differentiate between social class lines.

Recently there have been studies that employ instrumental measures to examine the current status of the voicing variation of /ʃ/ in BAS. Chang (2008) studied the devoicing of /ʃ/ and observed that the change is already generalized in the younger generation. Chang (2008:26) does not find an effect of sex, but does find a robust effect of age: “speakers in the youngest age bracket are all approaching a ceiling of 100% devoicing,” which seems to suggest that for them the change from /ʃ/ to /ʃ/ has already been completed. In terms of the effect of neighborhood of origin, Chang does not find a significant correlation with the devoicing.

Rohena-Madrazo (2008) also studied instrumentally the devoicing of /ʃ/, examining the social distribution of the change. His data confirmed that the men in the younger generation have already caught up with the women in terms of the usage of the voiceless variant; therefore, for the majority of the younger speakers the postalveolar fricative had already been devoiced. However, contrary to Chang (2008), Rohena-Madrazo found that area of residence, as an index of social class, did have an effect on the devoicing of /ʃ/. Among the older generation, the middle class speakers were in the lead of the innovation, ahead of both the upper and working class, which is a characteristic of changes “from below” (Labov 1972, 2001). However, the difference in voicing between the working and middle class has been eroded for the speakers of the younger generation, which seems to be evidence that the devoicing of /ʃ/ could be nearing completion.
2.2 The voicing feature in Spanish fricatives

It is complex to determine what a voiced fricative is in a language that lacks a phonemic voicing contrast for fricatives, since it is difficult to demarcate the limit that would separate a voiceless fricative from a voiced one, without the aid of native categorical perception. In the case of /ʃ/ in BAS, some previous studies identify several variants: completely voiced [ʒ], partially devoiced [˚ʃ], and completely voiceless [ʃ] (Fontanella de Weinberg 1978:223). These symbols for each category are conceptually interpretable; however, it is not clear what percentage of (de)voicing would correspond to each category.

When previous studies discuss “devoicing” or completion of the change, it is not clear what these terms refer to exactly. One extreme view of completion of the change is once speakers have reached ceiling levels of devoicing, as Chang (2008:62) mentions. In this case there is no doubt that the change has been completed. However, as was discussed in Section 1.2, the fact that /s/ is a phonologically voiceless fricative does not necessarily imply that it exhibits 0% voicing, since voiceless fricatives in Spanish can be realized occasionally with partial or complete voicing (see also Martínez Celdrán and Fernández Planas 2007:69-84, Martínez-Celdrán 2008, Martínez Celdrán 2009). By extension, it should not be taken for granted that a speaker for whom the devoicing of /ʃ/ has reached completion will realize every instance of /ʃ/ with 0%. For this reason, examining the voicing of /ʃ/ against a case of unmarked voicing like /s/ would be more informative than measuring the absolute levels of voicing for /ʃ/ alone, since it would serve as a voicing baseline for voiceless fricatives in BAS. As discussed above, among the voiceless fricatives in BAS, choosing /s/ as a control against which to compare /ʃ/ is most appropriate, since /s/ is most similar to /ʃ/ both articulatorily and acoustically; furthermore, the voicing levels of /s/ have not been associated with other social factors.
2.3 Goals of this study

The present study aims to investigate the devoicing of /zą/ in BAS, using instrumental techniques that capture the variability of the voicing of this fricative, in order to investigate the social and phonetic factors at play. In addition, by comparing the present results to those of previous studies, it will be possible to determine the progress of the change.

Another objective of this study is to propose a novel method of determining the completion of the devoicing of /zą/ in BAS by comparing the voicing levels of the postalveolar fricative /zą/ to those of /s/. If the devoicing change is really completed, then the voicing levels of /zą/ should not be different from the voicing levels of /s/. If this is the case, then perhaps it would be more appropriate to refer to the postalveolar phoneme as /ʃ/ instead of /zą/ in BAS.

In order to investigate the objectives stated above, a production experiment was conducted, which is detailed below.

2.4 Methodology

2.4.1 Materials

The data analyzed in this study were obtained from a larger production and perception experiment. The present data were extracted from the first word list task and the sentence reading task. The word list included 75 cases of the sibilant fricatives in question (35 tokens of /zą/ and 40 tokens of /s/) that were distributed in initial position and intervocalic position. The same words from the word list were also incorporated into the sentence list. The phonological contexts in which the words appeared in the sentence list were
matched as well as possible with those in the word list. The word list and sentence list can be found in Appendix A and B, respectively.

The recordings were made in a silent room with a Sony MZ-RH910 MD recorder and an Audio-Technica ATM75 head-mounted microphone.

2.4.2 Acoustic analysis: percentage voicing

The percentage of voicing during frication, evidenced either indirectly, through acoustic means (e.g. F0, voicing bar, etc.), or directly, through measurement of vocal fold vibration (e.g. electroglottography, throat microphone, etc.), has been used as a dependent measure in several studies of variation in fricative voicing (Haggard 1978, Docherty 1992, Smith 1997, Stevens et al. 1992, Chang 2008); however, percentage voicing is not the only correlate of phonological voicing. Several other studies, particularly conducted in languages with fricative voicing contrasts, have found that other correlates of voicing can be frication duration (Smith 1997, Stevens et al. 1992, Crystal and House 1988, Jongman et al. 2000, Baum and Blumstein 1987), the duration of the preceding vowel (Raphael 1972, Stevens et al. 1992, Smith 1997), average airflow during frication (Smith 1997), the relative intensity with respect to the adjacent vowels (Chang 2008, Colantoni 2006b), the degree of periodicity of the segment (Colantoni 2006a,b), the trajectory of F1 transitions (Stevens et al. 1992). Given the space constraints as well as the lack of information on perceptual cues for /ʒ/ devoicing in BAS, this investigation will focus on the more straightforward measure of percentage of voicing during frication as the primary dependent variable, with frication duration as a secondary dependent measure. The study of other possible voicing correlates involved in the devoicing of /ʒ/ will be left for future investigation.
The percentage voicing during frication for /\textit{z}/ as well as for /s/ was measured instrumentally using Praat 5.1.26 (Boersma and Weenink 2010). The segmentation of the fricatives was carried out by inspecting the onset and offset of high frequency energy. Borzone de Manrique and Massone (1981) found that the noise components of the postalveolar fricative in Buenos Aires Spanish range between 2500-5500Hz. Taking this into consideration, the boundaries of the fricative were placed at the onset and offset of broadband noise at those frequencies or higher, irrespective of any overlap with formant structure, following the segmentation criteria from fricative devoicing studies in English (Frisch and Wright 2002, Smith 1997). For both fricatives, this was the segmentation rubric used instead of one based on changes in formant structure or periodicity in the waveform, because a criterion was needed that could remain agnostic to voicing, since percentage voicing would be the dependent variable. Furthermore there were cases of voiced fricatives where there was clear frication; however F1 and F2 continued throughout the frication duration. An example of this can be seen in Figure 2.1.
If offset and onset of formants had been the segmentation criterion, then the token in Figure 2.1 would have been unsegmentable, since the formant structure continues throughout the duration of the fricative. However, the energy in the higher frequencies is clearly present, indicating that this is indeed a fricative. Figure 2.2 presents a less controversial example of the segmentation of an /s/.

Once the fricatives were segmented, a Praat script was used to mark 21 equidistant points throughout the duration of the fricative. The script queried F0 at each one of these 21 points, which would result either in a numerical value or “–undefined–” if F0 could not be calculated. F0 is directly correlated with the frequency of vocal fold vibration; therefore, if a numerical value for F0 could be queried, then that point of the fricative was voiced. The percentage of voicing of a particular fricative was the proportion of those 21 equidistant points for which F0 could be calculated.
2.4.3 Inter-experimental factors: fricative, position, and style

The factor *fricative* has two levels: /s/ and /ʒ/. If /ʒ/ is still voiced then it should exhibit voicing rates that are higher than those of /s/. If, on the contrary, the voicing levels of /ʒ/ and /s/ are not significantly different from one another, then it would be evidence for the completion of the devoicing change, in other words, that /ʒ/ exhibits the voicing levels of a phonologically voiceless fricative.

The factor *position* has two levels: phrase initial (#CV) and intervocalic (VCV). Assuming an articulatory account, then intervocalic fricatives should exhibit higher voicing rates of voicing, since there will be more opportunity for gestural misalignment through coarticulation (Browman and Goldstein 1990). Furthermore, obstruent voicing should be less likely in phrase initial position than in intervocalic position, because the pressure difference tends to be lower in utterance initial environments (Westbury and Keating 1986:153). Consequently, articulatory and aerodynamic constraints on obstruent voicing converge in favor of voicing in intervocalic position and against voicing in phrase initial position.

The factor *style* also has two levels: words and sentences. These correspond to the tasks in which the fricatives were read. In his style scale, Labov (1972:80-85) describes that the reading of isolated words as a more careful speech style than that of reading sentences in a text; therefore, the latter tends to exhibit higher usage of variants associated with formality or prestige. If we assume that the devoicing of /ʒ/ is not a prestige feature, as has been stated by Fontanella de Weinberg (1978, 1983) and Wolf (1984), then it should be subject to stylistic variation, in which case it would be expected that the voicing levels should be higher for the more careful word list style. On the other hand, this would not be expected for the voicing rates of /s/, since it is not associated
with conscious prestige and thus should not be subject to stylistic shift. Therefore, one might expect an interaction between style and fricative.

### 2.4.4 Participants and social factors

For this study the speech of 16 participants was analyzed. They were all born and raised in the city of Buenos Aires or its metropolitan area. The 16 speakers were equally stratified in terms of age (8 speakers between 18-30 vs. 8 speakers 55 or older), sex (8 males vs. 8 females), and social class (8 upper class speakers vs. 8 middle class speakers). The specific demographic information for each participant is presented in Table 2.1.

<table>
<thead>
<tr>
<th>Subject #</th>
<th>Sex</th>
<th>Age</th>
<th>Neighborhood</th>
<th>Zone/Social Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>F</td>
<td>20</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>27</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>9</td>
<td>M</td>
<td>23</td>
<td>Recoleta</td>
<td>North/Upper</td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>29</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>63</td>
<td>Burzaco</td>
<td>South/Middle</td>
</tr>
<tr>
<td>16</td>
<td>F</td>
<td>66</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>17</td>
<td>M</td>
<td>27</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>19</td>
<td>M</td>
<td>24</td>
<td>Vicente López</td>
<td>North/Upper</td>
</tr>
<tr>
<td>21</td>
<td>F</td>
<td>23</td>
<td>Recoleta</td>
<td>North/Upper</td>
</tr>
<tr>
<td>23</td>
<td>M</td>
<td>67</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>26</td>
<td>M</td>
<td>60</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>28</td>
<td>M</td>
<td>60</td>
<td>San Telmo</td>
<td>South/Middle</td>
</tr>
<tr>
<td>30</td>
<td>F</td>
<td>65</td>
<td>Vicente López</td>
<td>North/Upper</td>
</tr>
<tr>
<td>31</td>
<td>M</td>
<td>24</td>
<td>Olivos</td>
<td>North/Upper</td>
</tr>
<tr>
<td>33</td>
<td>F</td>
<td>58</td>
<td>Martínez</td>
<td>North/Upper</td>
</tr>
<tr>
<td>34</td>
<td>M</td>
<td>63</td>
<td>Martínez</td>
<td>North/Upper</td>
</tr>
</tbody>
</table>

Table 2.1: Demographic information of the participants in the production experiment.

In terms of age, the two extremes (18-30 vs. 55+) were chosen, since that is where the most marked differences are expected to appear. Given the findings of previous stud-
ies, it would be expected that for the speakers of the younger generation the devoicing change should be completed; therefore, they should not exhibit significant differences between the voicing rates of /ʒ/ and /s/. On the other hand, it is expected that for the older generation maintain the voicing of /ʒ/; therefore they should exhibit significantly higher rates of voicing for /ʒ/.

In past studies (Fontanella de Weinberg 1978, Wolf and Jiménez 1979) sex had been found to be a very important factor, whereby the women were ahead of the men. More recent studies (Wolf 1984, Chang 2008, Rohena-Madrazo 2008), however, have found that the differences in voicing based on sex have been eroding. For this reason it is not expected that there will be a significant effect of sex in this study.

Wolf (1984) and Rohena-Madrazo (2008) found that the voiced variant [ʒ] was more frequently produced by upper class speakers. Therefore, it would be expected that the upper class speakers in this corpus show voicing rates for /ʒ/ that are significantly higher than those for /s/. Considering the results of the previous studies, the young middle class speakers should be the group for whom the devoicing change has been completed, since they had been found to be leading the change.

### 2.5 Results and discussion

#### 2.5.1 Voicing and its implementations

During the acoustic analysis of the data, both fricatives exhibited completely voiced realizations as well as voiceless ones (See Figure 2.3 and Figure 2.4). This confirms the existence of voicing variability even in the voiceless category.

Although there were examples of fully voiced and fully voiceless realizations of each fricative, the vast majority of the tokens were partially voiced, particularly since
Figure 2.3: Waveforms and spectrograms for the word *allá*, “there”. Note the completely voiced realization [aZa] to the left by subject 21 (younger, upper class, female) and the mostly voiceless realization [aʃa] to the right by subject 5 (younger, middle class, female).

Figure 2.4: Waveforms and spectrograms for the sequence /asa/ in the word *pasajeros*, “passengers”. Note the completely voiced realization [aza] to the left by subject 5 (younger, middle class, female) and the mostly voiceless realization [asa] to the right by subject 21 (younger, upper class, female).
there was usually some degree of overlap between the onset/offset of frication and the onset/offset of the surrounding vowels. The implementation of this partial voicing varied as well.

Figure 2.5 shows two types of implementation of partial de/voicing. In the word *embellecer*, the */z/* is 38.1% voiced and the */s/* is 9.5% voiced; however, the voicing is not implemented the same way in both fricatives. In the */z/* there is both carryover voicing from the previous vowel as well as anticipatory voicing before the following vowel, whereas in */s/* there is only carryover voicing from the previous vowel. In Figure 2.6, one can see another example of the word *embellecer*, where both */z/* (19% voiced) and */s/* (14.3% voiced) show a similar partial voicing pattern of some carryover voicing and little to no anticipatory voicing.

A question that arises is whether the partial voicing of fricatives occurs more often during the onset of the fricative or during the offset of the fricative. Assuming an aero-dynamic model, one would expect the voicing to occur more frequently during the onset
Figure 2.6: Examples of both /s/ and /f/ in the word *embellecer*, “to beautify”, spoken by subject 6 (younger, middle class, male) in the word list task.

of the fricative, since the conditions for voicing deteriorate the longer that the frication is sustained. This pattern is what one would expect if voicing continues into the constriction and dissipates as the supraglottal air pressure equalizes with respect to the subglottal pressure (see Westbury and Keating 1986, Ohala 1983). On the other hand, if one assumes an articulatory model, where the partial voicing results from gestural mis-alignment, then the partial devoicing should be as common in fricative onset as in offset position. Furthermore, is either pre-voicing or post-voicing correlated with the socially conditioned voicing of /s/ in BAS? Since these are not the central questions being pursed in this dissertation, they must be left for subsequent investigations. Therefore, the percentage voicing values that will be shown in the subsequent analysis will be based on the overall percentage voicing present in the whole duration of the fricative.
2.5.2 Realizations of the sibilants /s/ and /ð/ 

A total of 2637 stimulus sibilants were extracted (1349 for /s/ and 1288 for /ð/). Out of these tokens, 50 were discarded because of ambient or microphone noise, disfluencies or reading errors, or because the token was not clearly segmentable. Although the stimuli were designed to elicit alveolar and postalveolar sibilants, other realizations occurred as well. The number of occurrences of these realizations is shown in Table 2.2 and Table 2.3. The realizations in the tables are shown between bars in order to remain agnostic to their voicing realization.

<table>
<thead>
<tr>
<th>Realization</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>/s/</td>
<td>1139</td>
</tr>
<tr>
<td>/ð/</td>
<td>79</td>
</tr>
</tbody>
</table>

Table 2.2: Realizations for the stimulus obstruent /s/.

<table>
<thead>
<tr>
<th>Realization</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ð/</td>
<td>1139</td>
</tr>
<tr>
<td>/dʒ/</td>
<td>79</td>
</tr>
<tr>
<td>/ʃ/</td>
<td>18</td>
</tr>
<tr>
<td>/j/</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 2.3: Realizations for the stimulus obstruent /ð/.

The two instances of /s/ realized as /ð/ occurred in <cazaba>, see example in Figure 2.7. The pronunciation of <z> as /ð/ is not characteristic of BAS but of Castilian Spanish. This might be a case of hyper-careful speech, such that it becomes essentially non-native.

A similar explanation to that of /ð/ might be appropriate for the approximant realizations of /ð/ as [j] and [ʃ], see Figure 2.8 and Figure 2.9. These marginal approximant realizations of /ð/ were also found in Chang (2008:61). These realizations have
Figure 2.7: Example of an interdental fricative realization of /s/ in the word *cazaba*, “s/he hunted”, spoken by subject 34 (older, upper class, male) in the word list task.

not been characteristic to BAS for well over a century (Fontanella de Weinberg 1978); however, speakers are exposed to these variants, not just in non-Argentinean media, but also in the school system. In a separate interview with subject 15, an elementary school teacher, she discussed the process of early alphabetization. In many schools in Argentina, schoolchildren are exposed to the realizations of <z> as [θ] and <ll> as [ʎ] for the purposes of dictation, in order to facilitate spelling. The combined factors of this being a controlled reading task as well as the investigator being a speaker of a foreign variety of Spanish could have influenced these hyper-correct/non-native realizations; perhaps the participants were aiming for some sort of supra-regional “norm”.¹ A potentially fruitful topic for future investigation might be whether these marginal segmental variants of /s/ and /ʃ/ surface also in the unscripted speech of the picture description task.

¹It should be noted that the investigator did not pronounce <s> as [θ] or <y, ll> as [ʎ, ʎ̃].
Figure 2.8: Example of a palatal lateral realization of /ʎ/ in the word *alla*, “there,” spoken by subject 34 (older, upper class, male) in the word list task.

Figure 2.9: Example of a palatal glide realization of the /ʎ/ in the sequence in the word *embellecer*, “to beautify”, spoken by subject 26 (older, middle class, male) in the word list task.
Figure 2.10: Example of a voiced affricate realization of /ʒ/ in the name Yolanda spoken by subject 30 (older, upper class, female) in the sentence reading task.

The relatively high number of affricate realizations for /ʒ/ is quite interesting, particularly, because they do not seem to be equally distributed. 70/79 affricated cases of /ʒ/ occur in #CV, which might be indicative of a prosodic boundary fortition effect, reminiscent of /ʝ/ being realized frequently as [jj] or [ʝ] in phrase initial position in other varieties of Spanish.

As seen in Figure 2.10 and Figure 2.11, the affricated realizations varied in voicing as well, similarly to /ʒ/ and /s/; however, it should be noted that the same effect is not found with /s/, since there were no affricated realizations of the alveolar fricative. Because this #CV affrication effect was not constant across the stimulus obstruents, the affricated realizations were excluded from the general voicing variation analysis.

Another aspect that should be mentioned is the presence of preaspiration in several of the tokens. These preaspirations appeared as a low amplitude wave with both periodic and aperiodic components. Examples of this are shown in Figure 2.12 and Figure 2.13.
Figure 2.11: Example of a mostly voiceless affricate realization of /ʃ/ in the word *llevale*, “take to him/her”, spoken by subject 26 (older, middle class, male) in the sentence reading task.

Figure 2.12: Example of a preaspirated realization of /s/ in the word *cazaba*, “s/he hunted”, spoken by subject 5 (younger, middle class female) in the word list task.
The mean duration of these preaspirations was 27.8ms (SD=12.1ms), their mean percent voicing was 91.3% (SD=22.6%). This preaspiration only occurred before alveolar or postalveolar fricative realizations. There were 62 preaspirated tokens of /s/ and 34 preaspirated tokens of /ʃ/. 83/96 cases occurred in words where the preceding segment was /a/.

This preaspiration phenomenon might be due to gestural miscoordination, in which the tongue tip takes longer to reach the critical alveolar or postalveolar position in order to produce the high intensity frication. At the offset of the vowel, the amplitude and complexity of the waveform decrease. The air is still being expelled in preparation for the production of the following sibilant; however, since the onset of the tongue tip gesture is not tightly coordinated with the offset of the gestures associated with the vowel, the relatively open vocal tract results in [ʃ]. The fact that this preaspiration occurs most when the preceding segment is /a/ is consistent with this gestural miscoordination.
account, since the /a/ requires the lowest tongue position and the tongue tip would have
to travel the furthest in order to reach the critical alveolar or postalveolar position.

<table>
<thead>
<tr>
<th>Subject #</th>
<th>5</th>
<th>6</th>
<th>9</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>19</th>
<th>21</th>
<th>23</th>
<th>26</th>
<th>28</th>
<th>30</th>
<th>31</th>
<th>33</th>
<th>34</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>5</td>
<td>3</td>
<td>6</td>
<td>16</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 2.4: Distribution of preaspirated tokens by subject.

The distribution of the preaspirated tokens by subject is shown in Table 2.4. Since
the tokens of preaspirated sibilants were relatively well distributed among most of the
subjects and the numbers were rather small, these realizations were included in the gen-
eral analysis.

<table>
<thead>
<tr>
<th>Obstruent</th>
<th>Position</th>
<th>#CV</th>
<th>l/CV</th>
<th>n/CV</th>
<th>V/CV</th>
<th>VCV</th>
</tr>
</thead>
<tbody>
<tr>
<td>/s/</td>
<td></td>
<td>358</td>
<td>32</td>
<td>81</td>
<td>77</td>
<td>785</td>
</tr>
<tr>
<td>/z/</td>
<td></td>
<td>304</td>
<td>36</td>
<td>8</td>
<td>138</td>
<td>653</td>
</tr>
</tbody>
</table>

Table 2.5: Distribution of tokens by obstruent and position. The symbol “#” indicates
phrase initial position and “/” indicates word initial, phrase medial position.

The numbers of the obstruents by position is not well distributed either, particularly
in the cases of n/VC and l/CV, which, aside from only occurring in the sentence reading
task, present comparatively few tokens. The distribution is shown in Table 2.5. Con-
sidering this distribution, only the phrase initial and intervocalic (word internal or word
boundary) will be included in the general analysis.

### 2.5.3 General analyses

The main dependent variable in this investigation is the percentage voicing of the seg-
ment in question, either /z/ or /s/. Since the distribution of the percentage values did not
show a normal distribution, the arcsine transform of the percentage values was used as
Table 2.6: Results of a multiple linear regression presenting the factors that significantly affect the voicing of fricatives. The $r^2$ values are cumulative.

To determine the effect that each one of the factors has on the percentage voicing of the fricatives, a linear regression was performed which included the factors fricative, position, style, age, sex, and class. The results in Table 2.6 show that all of the independent factors significantly affect the voicing levels of the fricatives, except for sex.

Notably, the factors which have the strongest effect on the voicing of the segment are fricative and position, since it would be expected that /ʒ/ exhibit higher percentages of voicing compared to /s/ and also that intervocalic fricatives exhibit higher voicing rates compared to fricatives in phrase initial position. Contrary to what had been found in previous studies (Wolf and Jiménez 1979, Fontanella de Weinberg 1978, 1983), the effect of sex was not significant [t(2417)=-.890, p=.373], which confirms Chang’s (2008) findings. Given these results, sex was excluded as a factor in the subsequent analyses.

A potential issue with position, which should be addressed before further analyses, is the status of the intervocalic positions VCV (phrase medial, word medial) and V/CV (phrase medial, word initial). As can be seen in the boxplots\footnote{In the boxplots, the dark horizontal line in the box indicates the median, the edges of the box indicate the upper and lower quartile, the whiskers indicate the minimum and maximum values, and the circles indicate the outliers.} in Figure 2.14, the per-
Figure 2.14: Boxplots showing the percentage voicing rates of the two stimulus fricatives in three positions.

Percentage voicing values for the sibilants in the two intervocalic positions are very similar to each other, particularly compared to the voicing levels in phrase-initial position.

The V/CV cases only occurred in the sentence reading task; therefore a three-way coding of position along the lines shown in Figure 2.14 is not orthogonal to the coding of task. A two-way ANOVA was used to test whether there were any significant differences in percentage voicing between the two intervocalic categories V/CV and VCV for each of the two fricatives. Neither the main effect of position, F(1,822)=0.194, p=0.660, nor the position*fricative interaction, F(1,822)=2.357, p=0.125, were significant. There was a significant main effect of fricative, F(1,822)=128.846, p<.001, which will be discussed in more detail below. The results from this ANOVA suggest that the differences
Table 2.7: Summary of the linear mixed effect regression model adjusted to the voicing levels of the fricative, with subject as a random factor, for the all the data combined. (Intercept = /ʃ/, sentences, VCV, older, upper class)

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>$\beta$</th>
<th>S.E.</th>
<th>t</th>
<th>$p_{MCMC}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>% voicing</td>
<td>(Intercept)</td>
<td>.235</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fricative</td>
<td>-.109</td>
<td>.030</td>
<td>-3.688</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Style</td>
<td>.077</td>
<td>.027</td>
<td>2.785</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td>Position</td>
<td>.210</td>
<td>.021</td>
<td>10.228</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Class</td>
<td>-.027</td>
<td>.117</td>
<td>-0.234</td>
<td>.896</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>-.115</td>
<td>.117</td>
<td>-0.988</td>
<td>.323</td>
</tr>
<tr>
<td></td>
<td>Fricative:Style</td>
<td>-.015</td>
<td>.024</td>
<td>-0.618</td>
<td>.536</td>
</tr>
<tr>
<td></td>
<td>Fricative:Position</td>
<td>.145</td>
<td>.027</td>
<td>5.360</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Style:Position</td>
<td>.037</td>
<td>.028</td>
<td>1.298</td>
<td>.194</td>
</tr>
<tr>
<td></td>
<td>Fricative:Class</td>
<td>.175</td>
<td>.032</td>
<td>5.451</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Fricative:Age</td>
<td>.329</td>
<td>.033</td>
<td>10.099</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Class:Age</td>
<td>.043</td>
<td>.165</td>
<td>0.258</td>
<td>.796</td>
</tr>
<tr>
<td></td>
<td>Fricative:Class:Age</td>
<td>.048</td>
<td>.046</td>
<td>1.029</td>
<td>.303</td>
</tr>
</tbody>
</table>

In percentage voicing between V/CV and VCV are not significant, and thus these two positions may be collapsed for the following analyses of /s/ and /ʃ/.

In order to investigate not only the main effects but the interaction of these factors, the data were analyzed in a linear mixed effects regression model, with subject as a random factor and fricative and style (within-subject), and class and age (between-subject) as fixed effects. For these statistical models the estimated effect size ($\beta$), its standard error (S.E.) and its corresponding t-value are reported. The p-values were calculated using Markov chain Monte Carlo simulations. If there are significant interactions, the data will be divided to explore their effects further.

The linear mixed effect regression model produced significant fixed effects for fricative, style and position (see Table 2.7). The mean percentage voicing of /ʃ/ (M=.331, SD=.334) is significantly higher than that of /s/ (M=.159, SD=.154). The percentage voicing of the fricatives in the sentence style (M=.303, SD=.284) is significantly higher
than that of the fricatives in the word list style \(M=0.173, SD=0.235\). The percentage voicing of the fricatives in phrase initial position \(M=0.081, SD=0.142\) is significantly lower than that of the fricatives in intervocalic position \(M=0.304, SD=0.282\). There was also a significant interaction between fricative and position. Neither of the main effects for the external factors of age and class were significant, only their interactions with fricative. This is not unexpected, since one would expect differences in voicing rates according to age and class for /ʒ/, but not for /s/. None of the other interactions were significant. It should be noted that the fricative*style interaction was not significant, which suggests that both /ʒ/ and /s/ are affected by style in the same manner. (See discussion of stylistic effects in Section 2.5.5 below.)

2.5.4 Fricative and position

The distribution of percentage voicing was significantly different between /s/ and /ʒ/, as can be seen in Figure 2.15. Overall, /ʒ/ had a wider range of variation in voicing than /s/. From this result we can infer that the devoicing of /ʒ/ has not yet been completed in the BAS speech community as a whole. However, this is not always the case, depending on several social factors, as will be discussed in more detail below.

As seen in the linear mixed effects model results above, the fixed effect of position as well as its interaction with fricative was significant. As Figure 2.16 illustrates that the relatively small (but significant) difference between /s/ and /ʒ/ in phrase initial position (/s/: \(M=0.049, SD=0.039\) vs. /ʒ/: \(M=0.119, SD=0.199\)) is exacerbated in intervocalic position (/s/: \(M=0.205, SD=0.161\) vs. /ʒ/: \(M=0.411, SD=0.340\)).

Recall that Westbury & Keating (1986) state that it is more difficult to stop vocal fold vibration during a singleton stop closure than it is to continue throughout the obstruent segment. Since voicing is the default speech-ready state, then realizing a devoicing
Figure 2.15: Boxplots showing the percentage voicing by fricative. ‘y’ represents /ʒ/.

Figure 2.16: Average percentage voicing rates for /s/ and /ʒ/ in each phrase initial and intervocalic position.

gesture necessarily involves an articulatory change, thus making a voiceless obstruent harder to produce in intervocalic position. Converging with this articulatory effect is the aerodynamic fact that the pressure differential in CVC position is comparatively more favorable for voicing (or less disfavorable) than that in #CV position.

However, if this is the case, then why do the fricatives in intervocalic position not exhibit higher rates of voicing still? This is probably the case when the articulatory and aerodynamic constraints on voicing conflict. The conditions for voicing are still nowhere
near optimal during the production of a fricative, so perhaps the mean voicing rates of
around 20% for /s/ represent a default when articulatory and aerodynamic constraints
on voicing come into play. Furthermore, there might be some phonological pressure
to maintain a faithful realization of voiceless fricative for /s/. The higher voicing rates
for /ʒ/ are likely due to the underlying voicing feature associated with that segment in
BAS; however, this underlying feature might not be the same for all speakers in the BAS
speakers (see discussion of voicers and devoicers in section 2.5.7 below).

2.5.5  **Stylistic variation and phonetic factors**

The effect of style was significant such that the fricatives in the sentence reading task
(M=.303, SD=.284) exhibit higher voicing rates that those in the in the word list task
(M=.173, SD=.235). Interestingly, the interaction between fricative and style was not
significant. Figure 2.17 shows the percentage of voicing by fricative and style. The two
lines for each fricative are nearly parallel between the word list and the sentence styles,
which shows that both /s/ (words: M=.103, SD=.079; sentences: M=.219, SD=.188) and
/ʒ/ (words: M=.261, SD=.321; sentences: M=.388, SD=.334) are affected in the same
manner with the shift in styles.

![Figure 2.17: Average percentage voicing rates for /s/ and /ʒ/ in each style.](image)

52
Fontanella de Weinberg (1978, 1983) found similar stylistic effects for the speakers with secondary education, but not for the speakers with elementary education or with university education, who showed lower devoicing rates in the most careful style (reading isolated words), which is a characteristic pattern of prestige variants (Labov 1972). If the voicing of /ʃ/ were associated with prestige, as Fontanella de Weinberg (1978, 1983) and Wolf (1984) suggest, the word list style - since it elicits the most careful speech - should exhibit higher levels of voicing. However, the opposite effect is seen: both fricatives increase their voicing levels from the word list style to the sentence reading style.

This apparently paradoxical result could be due to the fact that the speech rate in the sentence reading task, since it elicits connected speech, is faster than in the reading of isolated words. Browman and Goldstein (1990) and Byrd (1996) have found that in faster speech there tends to be an increment in gestural misalignment. In the case of BAS sibilants, the faster speech in the sentence reading style could disfavor the precise alignments of the oral gestures with the laryngeal gestures. As a result, the laryngeal gestures for voicing of the vowel that precedes or follows the fricative can encroach on the oral gestures of the fricative, thereby causing its partial or complete voicing.

If the speech rate of the sentence reading task is faster, one would expect both fricatives to have shorter durations. If this were the case, there also should be a significant effect of style on the duration of the fricative.

In fact, a linear mixed-effect regression with duration (in seconds) as a dependent variable (see Table 2.8) gives very similar results to the regression on the percentage voicing (see Table 2.7). There are significant effects of the fixed factors fricative, style, and position, as well as a marginally significant effect of age. The duration of /ʃ/ (M=.122, SD=.038) is significantly shorter than the duration of /s/ (M=.137, SD=.038).
### Table 2.8: Summary of the linear mixed-effect regression model adjusted to the duration of the fricative, with subject as a random factor, for the all the data combined. (Intercept = \(/_3^/\), sentences, VCV, older, upper class)

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>(\beta)</th>
<th>S.E.</th>
<th>t</th>
<th>(p_{MCMC})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td>(Intercept)</td>
<td>.181</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fricative</td>
<td>-.002</td>
<td>.003</td>
<td>-6.415</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Style</td>
<td>-.037</td>
<td>.003</td>
<td>-14.202</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Position</td>
<td>-.024</td>
<td>.002</td>
<td>-12.008</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Class</td>
<td>.003</td>
<td>.013</td>
<td>0.002</td>
<td>.999</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>-.025</td>
<td>.013</td>
<td>1.984</td>
<td>.047</td>
</tr>
<tr>
<td></td>
<td>Fricative:Style</td>
<td>.002</td>
<td>.002</td>
<td>0.944</td>
<td>.345</td>
</tr>
<tr>
<td></td>
<td>Fricative:Position</td>
<td>-.006</td>
<td>.003</td>
<td>-2.332</td>
<td>.020</td>
</tr>
<tr>
<td></td>
<td>Style:Position</td>
<td>.010</td>
<td>.003</td>
<td>3.582</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Fricative:Class</td>
<td>-.006</td>
<td>.003</td>
<td>-1.961</td>
<td>.050</td>
</tr>
<tr>
<td></td>
<td>Fricative:Age</td>
<td>.027</td>
<td>.003</td>
<td>8.583</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Class:Age</td>
<td>-.004</td>
<td>.018</td>
<td>-0.215</td>
<td>.830</td>
</tr>
<tr>
<td></td>
<td>Fricative:Class:Age</td>
<td>-.005</td>
<td>.004</td>
<td>-1.220</td>
<td>.223</td>
</tr>
</tbody>
</table>

In the style factor the duration of the fricatives in the word list (M=.148, SD=.038) is significantly longer than the fricatives in the sentence list (M=.112, SD=.031). With position, the phrase initial fricatives (M=.154, SD=.041) were longer than the ones in intervocalic position (M=120, SD=.034). The only difference between the significant fixed effects in this analysis and the one done for percentage voicing is the factor age: older speakers’ fricatives (M=.137, SD=.042) were longer in duration than those of the younger speakers (M=.123, SD=.035). Shipp et al. (1992) found that the speech rates of speakers below 30 years of age were significantly higher than those of speakers 50 and above while reading the rainbow passage in English. It might be the case that similar differences are found in Spanish. However, whether these differences in speech rate are due to physiological or socio-stylistic effects is an open question.

Similar to the analysis of the percentage voicing, the fricative*position, fricative*class and fricative*age interactions are also significant and, importantly, the interaction of
fricative*style is not significant, thereby demonstrating that both fricatives are affected in the same manner with in the shift form one style to the other (see Figure 2.18).

Contrary to the results of the percentage voicing analysis, there is a significant interaction of style*position, see Figure 2.19. This is not unexpected, given that both intervocalic position and running speech are two factors that influence hyparticulation, especially when compared to utterance initial position and the very careful word list style.

In languages such as English, with a voicing contrast in fricatives, it has been reported that phonologically voiced fricatives exhibit shorter durations compared to their voiceless counterparts (Jongman et al. 2000). Even in a study of BAS Spanish fricatives, Borzone de Manrique and Massone (1981) found differences in frication duration
of voiced and voiceless /ʒ/ (170-210ms for [ʃ] vs. 98-149ms for [ʒ]), it is not clear, however, how they elicited the different voiced and voiceless tokens, since this is not a difference that can be reliably encoded in Spanish orthography and even in loanwords present variable voicing rates (see Fontanella de Weinberg 1978).

A Pearson’s test indicates that there is a significant, and considerably strong, negative correlation between percentage voicing and duration of the fricative, \[ r(2313)=-.512, p < .001 \]. In other words, the longer the fricative is, the lower its percentage voicing will be. Haggard (1978:101) finds a similar negative correlation of frication duration and percentage of voicing for phonologically voiced fricatives in English. The correlations of duration and voicing percentage in BAS are significant for both /ʒ/ \[ r(1093)=-.546, p < .001 \] and /s/ \[ r(1218)=-.460, p < .001 \]. This result is not surprising if one considers the conflicting aerodynamic constraints that exist between the vibration of the vocal folds and the production of frication (see discussion in Section 1.2.1). If the frication duration is long, then this means that the adverse conditions for phonation will also be sustained for a longer period (Ohala and Kawasaki-Fukumori 1997). The faster and shorter the articulation of the fricative, the easier it will be for the fricative to be voiced. These tendencies are increased when the interaction of position and fricative is taken into consideration.

It is possible that the stylistic variation, which formerly would have been motivated by factors of prestige, has given way to phonetic factors. This shows that the devoicing change of /ʒ/, as well as the stylistic variation in fricative voicing, is motivated by the interaction of aerodynamic and articulatory factors.
2.5.6 Social factors and the completion of change

The general results suggest that the devoicing of /ʒ/ has not yet been completed for the speech community as a whole, since /ʒ/ still shows higher voicing rates than the voiceless fricative /s/; however, this might not be the case for every subgroup of the community. In order to explore the interactions of the fricative factor with age and class, the data was divided into 4 subgroups: 1) older upper class speakers, 2) younger upper class speakers, 3) older middle class speakers, and 4) younger middle class speakers. This way it can be determined whether the devoicing change has reached completion in any of these subgroups.

In the four linear mixed-effect regression models shown in Table 2.9, the fixed factors are fricative and style. If in any of the subgroups there is a significant effect of fricative, then the change has not been completed for that group, since the /ʒ/ still maintains higher levels of voicing than those of /s/.

The results in Table 2.9 indicate a significant effect of fricative, where the /ʒ/ shows higher levels of voicing than /s/, in the older upper class speakers (/s/: M=.138, SD=.100; /ʒ/: M=.532, SD=.387), younger upper class speakers (/s/: M=.181, SD=.197; /ʒ/: M=.290, SD=.336) and the older middle class speakers (/s/: M=.121, SD=.116; /ʒ/: M=.348, SD=.337). The only subgroup that did not show the same effect for fricative was the younger middle class speakers (/s/: M=.201, SD=.169; /ʒ/: M=.186, SD=.147). Each of the subgroups showed a significant effect of style, but most of them showed no significant fricative*style interaction, except for older upper class speakers. Although not very robust, this interaction patterns in the same direction as the rest, in fact the difference in voicing rates between /ʒ/ and /s/ is larger in the sentence reading task. Therefore, the stylistic effect is not due to voiced /ʒ/ being associated with more careful speech, thus supporting the phonetic motivation for the stylistic effect.
<table>
<thead>
<tr>
<th>Response</th>
<th>Social Class</th>
<th>Age</th>
<th>Predictor</th>
<th>$\beta$</th>
<th>S.E.</th>
<th>t</th>
<th>$p_{MCMC}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>% voicing</td>
<td>Upper</td>
<td>Older</td>
<td>(Intercept)</td>
<td>.283</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fricative</td>
<td>.493</td>
<td>.044</td>
<td>11.13</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Style</td>
<td>.109</td>
<td>.040</td>
<td>2.70</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fricative:Style</td>
<td>.124</td>
<td>.040</td>
<td>2.05</td>
<td>.040</td>
</tr>
<tr>
<td></td>
<td>Younger</td>
<td>Older</td>
<td>(Intercept)</td>
<td>.286</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fricative</td>
<td>.171</td>
<td>.039</td>
<td>4.44</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Style</td>
<td>.253</td>
<td>.038</td>
<td>6.69</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fricative:Style</td>
<td>-.032</td>
<td>.054</td>
<td>-0.59</td>
<td>.557</td>
</tr>
<tr>
<td></td>
<td>Middle</td>
<td>Older</td>
<td>(Intercept)</td>
<td>.268</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fricative</td>
<td>.322</td>
<td>.039</td>
<td>8.28</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Style</td>
<td>.106</td>
<td>.036</td>
<td>2.92</td>
<td>.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fricative:Style</td>
<td>.007</td>
<td>.054</td>
<td>0.13</td>
<td>.893</td>
</tr>
<tr>
<td></td>
<td>Younger</td>
<td>Older</td>
<td>(Intercept)</td>
<td>.330</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fricative</td>
<td>-.001</td>
<td>.024</td>
<td>-0.06</td>
<td>.951</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Style</td>
<td>.219</td>
<td>.024</td>
<td>9.14</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fricative:Style</td>
<td>-.052</td>
<td>.034</td>
<td>-1.53</td>
<td>.128</td>
</tr>
</tbody>
</table>

Table 2.9: Summary of the linear mixed effect regression model adjusted to the voicing levels of the fricative, with subject as a random factor, for the fours subgroups divided by class and age. (Intercept = /z/, sentences).

Figure 2.20: Average percentage voicing rates for /s/ and /z/ for the subjects stratified by age and social class, pooled across tasks. The asterisks indicate that the difference in percentage voicing between /s/ and /z/ is statistically significant.
These results suggest that the only subgroup for which the change has been completed is the younger middle class speakers, for whom the /z/ exhibits voicing levels of a phonologically voiceless fricative. According to the findings in previous studies (Rohena-Madrazo 2008, Wolf and Jiménez 1979), where younger middle class speakers were found to be leading the change, it is expected that the younger middle class group show the lowest voicing rates of /z/. Actually, this group is beyond just leading the change—they have already completed it (see Figure 2.20).

Chang (2008:62) had noted that the change from [z] to [j] had been completed in the younger generation; however, this is not exactly the case for all younger speakers: the younger upper class speakers still have not devoiced /z/ completely. Furthermore, qualifying Chang’s statements that “speakers in the youngest age bracket are all approaching a ceiling of 100% devoicing” (2008:62), we see that for the change to have been completed, it is not necessary for the /z/ to present 0% voicing, rather that its voicing not be different from the voicing of a phonologically voiceless fricative, which was about 17% for the /s/ in this study.

For the older middle class speakers and for the upper class speakers, the change does not seem to have been completed yet. However, although these groups exhibit a difference in voicing between /z/ and /s/, this does not necessarily mean that that these groups all have the same voicing levels for their /z/. In fact, one can compare the mean voicing percentages for /z/ between the groups for whom the devoicing change has not been completed. Tukey HSD tests reveal that the voicing levels of /z/ for the older upper class speakers (M=.522, SD=.385) are significantly higher than both the younger upper class speakers (M=.288, SD=.331) and the older middle class speakers (M=.348, SD=.335). However, the differences between the voicing levels of /z/ for the younger upper class speakers and the older middle class speakers are not significant. These
results suggest that, even among the groups who still voice /ʒ/ more than /s/, there are indications that the change is still in progress. It is particularly important to note that, contrary to what Wolf (1984) proposed, the devoicing of /ʒ/ does not seem to be a case of stable variation: younger upper class speakers, although they might be voicers of /ʒ/, show lower levels of voicing than /ʒ/ than the older generation.

2.5.7 Completion of change at the individual level

The results of the previous section suggest that among the social subgroups that maintain a significant difference in the voicing of /s/ and /ʒ/ there are differences in the voicing of /ʒ/ across groups. An interesting question that arises is whether a group that is different from another is different because the voicers in one group have similarly lower voicing rates or whether a group is comprised of a variable number of voicers and devoicers. In order to further explore these differences, a one way ANOVA was performed for each subject in order to determine whether their voicing rates of /ʒ/ were significantly different from those of /s/. The results are shown in Figure 2.21.

Nine speakers were voicers (whose percentage voicing of /s/ and /ʒ/ were significantly different) and seven were devoicers (whose percentage voicing of /s/ and /ʒ/ were not significantly different). However, not all the voicers exhibited similar voicing rates of /ʒ/, particularly subject 9 [F(1,150)=5.416, p=.021], subject 33 [F(1,146)=4.812, p=0.030], and subject 16 [F(1,142)=3.989, p=.047], whose differences do not seem to be as robust as the other voicers’. Considering their differences between /s/ and /ʒ/, these voicers seem to be moving in the direction of the devoicing change as well. Even comparing between voicers with robust differences, the voicing ranges for /ʒ/ seem to be highly variable. For example, compare subject 21 for whom more than half of the /ʒ/ tokens were 100% voiced, to subject 15, for whom the majority of the /ʒ/ exhibited
Figure 2.21: Boxplots showing the percentage voicing for /s/ and for /z/ produced by each subject, stratified by age and social class. The values for /s/ are in white and the values for /z/ are in gray (and marked by ‘y’). The asterisks beneath the subjects indicate whether the mean percentage voicing rates of /s/ and /z/ were significantly different (‘*’ p < .05, ‘***’ p < .001).
percentage voicing of 50% or less. Comparatively, the devoicers are a much more stable group with regards to fricative voicing, i.e. the voicing rates for each of the two fricatives are roughly comparable across speakers. This analysis provides more evidence against the claim that the devoicing of /ʒ/ has reached the level of stable variation. In fact, in most of the social group divisions under investigation here there are speakers who are devoicers, and for the group who is comprised only of voicers, the older upper class, there is one speaker who likewise appears to be moving in that direction. The results for the older speakers whose difference between voicing of /ʒ/ and /s/ is not very robust (i.e. subjects 33 and 16) raise the question of whether they have maintained those voicing levels since youth or whether they had had higher voicing rates of /ʒ/ and have decreased them throughout their lifetime, as has been found for other changes in progress (see Sankoff and Blondeau 2007 for [r] > [r] in Montreal French). Unfortunately, since there is no longitudinal data available on variation in percentage voicing of /ʒ/ and /s/, there is no way to tell whether the comparatively lower percentage voicing of /ʒ/ of these non-robust voicers represents apparent time change or actual change across the lifespan.

For the seven devoicers in this study it would no longer be appropriate to posit an underlying voiced fricative /ʒ/ as part of their inventory but rather to posit /ʃ/, since its voicing levels resemble those of a phonologically voiceless fricative. If this is true, then the resulting /ʃ/ should be subject to the same phonetic effects of position that were present for /s/. Recall that in the general analysis there was a significant interaction between fricative and position, such that /ʒ/ showed much more voicing than /s/ in intervocalic position, where the articulatory conditions for voicing are more favorable. For these devoicers, this interaction of position and fricative should no longer be significant.
if the devoicing has been completed. This is exactly what the results show in a linear mixed effect model run on this subset of the speakers, see Table 2.10.

Furthermore, if /ʃ/ is the underlying phoneme for these devoicers, then the effects of fricative and position seen in Table 2.10 should be mirrored in an analysis with fricative duration as the dependent variable. Again, this is exactly the case (see Table 2.11). The resulting phoneme /ʃ/ is affected in the same manner by all the factors that affect the voicing rates and duration of /s/.

Yet another way in which the resulting /ʃ/ of the devoicers is phonologically different from the /ʒ/ has to do with some tokens that were actually excluded from the voicing analysis. Recall that in section 2.5.2, 79 tokens of /ʒ/ were excluded because they were realized as affricates, [dʒ]. It happens to be the case that 78/79 instances of affricated realization of /ʒ/ were produced by speakers who are voicers. Furthermore,
this affrication occurs overwhelmingly in phrase initial position #CV or postconsonantal C/CV positions. This is reminiscent of the description of the variation in production of /j/ in other varieties of Spanish, for example, in Castilian Spanish, affricated or stop realizations of /j/ as [jj] or [j] are more common postconsonantally or in utterance-initially (Hualde 2005:165-6). This also seems to be the case for voicers of /ʒ/ in BAS, who produce [dʒ] realizations in utterance-initial position and postnasally. If this phenomenon were due to a process of articulatory strengthening, then it would be expected that devoicers also show [tʃ] realization of /ʒ/ in those positions. However, this affrication effect is not at all common among the devoicers, who do not produce [tʃ] when the postalveolar fricative occurs in #CV position. For the voicers, the /ʒ/ exhibits contextual variation that is phonologically parallel to its etymological analog in other variants of Spanish, /j/. However, for the devoicers the postalveolar fricative diverges from the phonological behavior of the voicers’ /ʒ/, surfacing mostly as a fricative, regardless of position. This differing phonological pattern is further evidence supporting the claim that the devoicers’ underlying phoneme is no longer /ʒ/, but rather the voiceless fricative /ʃ/.

Interestingly, in the informal metalinguistic commentary that many of the subjects offered after the experimental session, the only phonological feature that several speakers mentioned was this affricated pronunciation, which these speakers associated with upper class speakers. One participant mentioned something along the lines of Las chicas chetas dicen “[dʒo]”, ‘the upper class girls say [dʒo] (I)’. This appears to be, at least for some speakers, a sociolinguistic stereotype (Labov 1972) that identifies upper class speakers, who in this corpus were all voicers. It is interesting that none of the participants identified this affrication, or the voiced realization of /ʒ/ with older speakers, who are overwhelmingly voicers as well. However, subject 30, an older upper class female,
in her comments during the perceptual assimilation rating task, did identify \[f\] as a characteristic of the speech of the younger generation and had a negative opinion of it. The devoicing of \(/z/\) seems to exhibit the pattern of change “from below” in its final stages, perhaps with correction from above (Labov 1966:225). However, the extent to which the (de)voicing of \(/z/\) is a phenomenon that speakers in the BAS community are aware of is not very clearly understood. Since this was not the main objective of my study, the observations made unsystematically by some of the participants in this study are inconclusive at best. More perceptual subjective reaction studies are necessary to determine the degree to which this change, which seems to be nearing completion among the younger speakers, is above the level of consciousness for the speakers in the community.

### 2.6 Conclusion

This chapter has presented an instrumental acoustic analysis of the variable devoicing of the postalveolar fricative in BAS, comparing its voicing levels to those of \(/s/\) in order to determine the stage of the change. The results show that for the younger middle class the postalveolar fricative has devoiced to a level equivalent to that of the other phonologically voiceless sibilant in the system: \(/s/\). This is evidence that the change towards devoicing – a classic example of change in progress – for a group of BAS speakers has already been completed. In terms of the phonological inventory of these speakers, perhaps it would be more appropriate to characterize the postalveolar fricative phoneme not as a \(/\tilde{z}/\) but as a \(/f/\). According to Chang’s (2008) findings, for the speakers from the younger generation, the change had already reached completion. However, given the present results, this does not seem to be the case, since there are two out of the four young speakers from the upper class who still maintain the voicing of \(/\tilde{z}/\); however, most younger speakers seem to have devoiced \(/\tilde{z}/\) or are moving in that direction. These
results also confirm that neighborhood, as a proxy for social class, does seem to affect the voicing levels of /\v/.

Although the devoicing of /\v/ could have resembled a stable variable, functioning as a social differentiator for the upper class, as Wolf (1984) had noted, the instrumental analysis in this study presents evidence that the voicing rates of /\v/ in this class is also ceding to the devoicing change, albeit at a slower rate. Even between the groups that maintain a difference between the voicing of /\v/ and that of /s/, the voicing rates of /\v/ are not the same across the groups: the younger upper class speakers and the older middle class speakers exhibit lower voicing rates than the older upper class speakers. Undoubtedly, the voicing rate of /\v/ is still a sociolinguistic variable but, although the completed change might serve as a social indicator for some speakers, the evidence in the present study suggests that the change is still in progress for a large portion of the BAS speech community.

The results of this investigation suggest that the devoicing of /\v/ is affected by stylistic factors, but not in classic sociolinguistic terms of shifting towards a prestige norm in more careful settings, as had been described by Fontanella de Weinberg (1978, 1983) for Bahía Blanca. Both fricatives exhibit higher voicing rates in the reading of sentences than in the reading of a list of isolated words, which is supposed to elicit more careful speech. These differences in the voicing could depend on the interaction of speech rate and the coordination of the speech gestures needed to produce frication and voicing. Therefore, these stylistic differences could be explained by the differences in speech rates between the word list and sentence list styles. Instead of depending on accommodation to a prestige norm, these stylistic effects seem to be phonetically motivated. The significant effect of position and style show that these are crucial factors that determine the voicing variation not just of /\v/ but of /s/ as well. In fact, this systematic variation
in voicing of /s/, due to the varying articulatory and aerodynamic constraints associated with different positions and styles, serves as a phonological yardstick to determine whether the underlying postalveolar phoneme for some BAS speakers is actually the voiceless fricative /ʃ/.

The present instrumental study of the voicing of /ʒ/ provides information about the social stratification of the devoicing phenomenon, confirming some of the predictions of previous studies, where it had been found that younger middle class speakers were in the lead of the change. A detailed description of the voicing rates of /ʒ/ can reveal much information, especially about their distribution in the speech community. However, it does not provide tools for determining the completion of change, because it does not take into consideration the inherent voicing variability among the voiceless fricatives in the system. The study presented here proposes an independent method to establish the completion of change, by comparing the voicing levels of the sociolinguistic variable /ʒ/ to those of /s/, a phonologically voiceless fricative that is not subject to the same social factors as /ʒ/. Understanding the inherent variability in the system provides firmer footing on which to make determinations about sociophonetic variation and phonological change in the sibilants of Buenos Aires Spanish.
Chapter 3

Perception of non-native obstruent voicing contrasts

3.1 Introduction

The previous chapter analyzed the phonetically and socially conditioned variation in production of voicing in BAS Spanish sibilants, particularly with respect to the devoicing of /ʒ/. The current chapter will examine the ways in which sociophonetic and allophonic voicing variation in the speech community affect how listeners perceive an analogous voicing contrast in another language. Two models of cross-language speech perception, the Perceptual Assimilation Model and the Speech Learning Model, will be discussed in terms of the predictions that they make about how BAS listeners perceive native-like and non-native obstruent voicing contrast. The differing predictions made by these models will be tested in two discrimination experiments in order to shed light on the effects of native phonological factors as well as the effects of sociolinguistic variation on the perception of non-native contrasts.
3.2 Background on cross-language speech perception

Much research in cross language speech perception has revealed that adults have difficulty distinguishing between phonetic pairs that are not contrastive in their native language (Goto 1971, Gillette 1980, Best and Strange 1992). However, many of these investigations have also demonstrated that, although adults do have difficulties distinguishing non-native contrasts, not all non-native contrasts are perceived with equal difficulty (Best et al. 1988, Werker and Tees 1984b, Polka 1991). The performance of these cross-language listeners depends on several factors, among them the age of the listener (Werker 1995, Werker and Tees 1984a, Kuhl et al. 2006, Flege 1995), the type of testing conditions of the tasks (Werker and Tees 1984b, Werker and Logan 1985, Pisoni 1973, Davidson and Shaw submitted), the phonetic differences and similarities between the non-native phones and phonemes in the listener’s native inventory (Best 1995, Flege 1995, Best et al. 1988, 2001), the phonotactic patterns in which the phones are presented (Dupoux et al. 1999, Davidson 2007, Hallé et al. 1998, 2003) as well as the phonological contexts in which the contrasts appear (Logan et al. 1991, Strange et al. 2001, Flege 1989). Several different models of cross-language speech perception have been developed in order to account for these differences in listener performance, for example, the Perceptual Assimilation Model (Best 1995, Best et al. 1988, 2001, Best and Tyler 2007) the Speech Learning Model (Flege 1989, 1995, 1999), and the Native Language Magnet Model (Kuhl 1993, Kuhl and Iverson 1995, Kuhl et al. 2008). These models all propose that the listeners do not perceive non-native contrasts in a native-like fashion because of the influence of their native sound systems. Analyzing the variable performance of listeners in cross-language perception tasks can shed light on the phonological organization of the listeners’ native language. The listeners’ native phonetic and phonological categories affect how they perceive novel sounds; however, the nature and extent of the
effect of these categories is what differentiates these models of cross-language speech perception. For an overview of these and other models, see Strange (1995) as well as Bent (2005).

The present study will mainly be concerned with two of these models, Best’s Perceptual Assimilation Model and Flege’s Speech Learning Model. As will be discussed below, these two models make predictions, sometimes contradictory ones, about the role that positional allophones and phonological categories play in the perception of non-native speech sounds.

### 3.3 Models of cross-language speech perception

#### 3.3.1 Perceptual Assimilation Model

One of the theories that attempts to explain the fact that some non-native contrasts seem to be easier to perceive than other non-native contrasts is Best’s Perceptual Assimilation Model (PAM) (Best 1995, Best et al. 1988, 2001, Best and Tyler 2007). PAM posits that different types of assimilation patterns result from the articulatory similarities of the non-native phonemes to the native categories of the listeners. It is essential to PAM that the perceptual assimilation patterns be based on articulatory similarities between the stimulus language and the native language, because Best’s model is couched in the direct realist approach, which posits that what is perceived in speech is the distal object that produced the sound (i.e. the articulatory gestures), rather than the acoustic properties of the sound.

Best (1995) and Best et al. (1988, 2001) discuss several different types of assimilation that can occur depending on the comparison of the non-native phones and the listeners’ native phonemic inventory. They propose various assimilation types, depending
on the divergence from the native categories. The non-native sounds can be perceived as categorizable with respect to their native categories, in which case the accuracy in discrimination will vary depending on the goodness of fit to the native category that the non-native sounds are assimilated to. Alternatively, the non-native speech sounds are perceived as uncategorizable, where the novel sounds are perceived as speech but cannot be assimilated to any native category, in which case the listeners will use a combination of phonetic and auditory processing to distinguish the sound. Finally, non-native sounds can be perceived as non-assimilable, which means that they are essentially not perceived as speech. In these cases, the phonological knowledge is bypassed and the listeners use language-independent auditory processing exclusively to distinguish between the non-native contrasts.

The contrasts that will be examined in this chapter fall within the realm of categorizable non-native contrasts. Therefore the relevant assimilation types are mainly Two Category Assimilation (TC), Same Category Assimilation (SM) and Category Goodness Assimilation (CG) (but see (Best et al. 1988, 2001) for predictions and results on the other assimilation types).

In the cases of TC assimilation, two non-native contrastive phonemes are mapped onto two natively contrastive phonemes. In this assimilation type PAM predicts good discrimination, because, although the particular contrast does not exist in the native language, it is mapped onto a native pair, whose members are phonemically distinct, thus allowing an easier discrimination between the two. The example that (Best et al. 2001) give for TC assimilation is the Zulu contrast between voiced and voiceless lateral fricatives [ɬ]-[ɮ] as perceived by native English listeners. English does not have any lateral fricatives in its phonemic inventory; however, it does have the contrasting coronal fricative pair /s/-/z/, which are articulatorily closest to the Zulu pair [ɬ]-[ɮ]. Furthermore,
they differ on exactly the same feature, in this case, voicing. For these reasons, PAM predicts high accuracy in the discrimination of this contrast.

SC assimilation, on the other hand, occurs when the two non-native phonemes are mapped equally well onto the same single category in the native inventory. In these cases PAM predicts poor discrimination, since the contrasting non-native phones are perceived, essentially, as equally good exemplars of the same native category. The example that (Best et al. 2001) present is the [b]-[ɓ] contrast in Zulu, which is absent in English and the two phones are mapped onto the same /b/ category in English. The implosive bilabial stop is virtually indistinguishable from the plosive /b/ to native English listeners, which leads to low accuracy rates in the discrimination of the contrast.

In CG Assimilation the two non-native phonemes are mapped onto the same native category, but contrary to SC assimilation, one of the non-native phonemes is a better fit to the native category, i.e. it is a more prototypical exemplar of the native category that it maps onto. In these cases PAM predicts intermediate accuracy in discrimination rates, because, although the two non-native phones are mapped onto the same native category the discrepancy in goodness of fit between the two phones is an aid in distinguishing them. The example of CG assimilation in (Best et al. 2001) is the aspirated vs. ejective contrast [kʰ]-[k’] in Zulu as perceived by English speakers. The two non-native phonemes are mapped onto the native /k/ category in English; however, [kʰ] is a common realization of /k/, whereas the ejective variant [k’] is deemed a worse exemplar of the category. Although the non-native contrasts are assimilated to the same native category, the goodness of fit discrepancy for one member of the pair makes the contrast more distinguishable. Therefore, the listeners’ accuracy rates in discrimination of this contrast should be higher, as compared to the accuracy rates of a SC assimilation pattern.
Given these assimilation types and the accuracy rates predicted for each, PAM predicts the following accuracy scale: TC > CG > SC. According to this scale, listeners should be most accurate at discriminating Two Category Assimilation types, less accurate in Category Goodness Assimilations and least accurate in Same Category Assimilation types.

### 3.3.2 Speech Learning Model

Another theory that attempts to explain and predict the patterns of cross-language speech perception is Flege’s Speech Learning Model (Flege 1989, 1995, 2003). SLM focuses on second language acquisition, particularly on the effects that the native language sound system as well as external factors, such as age of exposure and proportion of use of L1 and L2, have on the perception and production of non-native phonetic and phonological categories. SLM posits that the elements that make up the subsystems of the L1 and L2 all exist within a common phonological space. Although discriminating between non-native categories becomes more difficult as the age of exposure increases, SLM proposes that the basic mechanisms for acquiring new categorical representations for speech sounds are essentially the same throughout the life span. According to SLM, perception of the non-native speech segments will depend on how they compare acoustically to the other native segments in the same position. The aspect of position is crucial, since SLM predicts that the listener should perceptually relate “positional allophones in the L2 to the closest positionally defined allophone (or “sound”) in the L1” (Flege 1995:238).

As mentioned above, one of the main hypotheses of the Speech Learning Model is that “sounds in the L1 and L2 are related perceptually to one another at the position-sensitive allophonic level, rather than at a more abstract phonemic level.” (Flege 1995:239).
In other words, the non-native sounds are related to position sensitive allophones in the native language. Flege finds evidence for this in several studies that have found discrepancies in the perception (and production) of non-native phonemes in initial vs. final position. For example, several studies have found that native Japanese listeners have difficulty distinguishing between the American English phonemes /\alpha/ and /l/; however, the Japanese listeners can discriminate this non-native contrast better in word-final position than in word initial position (Logan et al. 1991, Strange 1992). These discrepancies in accuracy might be due to the different acoustic cues that /\alpha/ and /l/ present in the speech signal in final position that are comparatively more robust in final position than in initial position, particularly since there are acoustic cues present in the preceding vowel for final liquids but not for initial liquids (Sheldon and Strange 1982).

The further away in the acoustic perceptual space that an L2 sound is from the listener’s L1 category, the likelier that the listener will be able to discriminate between the L2 sounds and eventually form a distinct phonetic category for it. Some studies have shown that the acoustic characteristics of Japanese /r/ are closer to English /l/ than to English /\alpha/ (Sekiyama and Tohkura 1993, Takagi 1993); therefore; according to SLM’s predictions an adult native Japanese speaker will be more likely to establish a category for /\alpha/ than for /l/, and thus be able to perceive /\alpha/ more accurately. The results from Flege et al. (1996) confirm these predictions, where native Japanese speakers with English experience identified English /\alpha/ more accurately than English /l/ in naturally produced English minimal pairs.

3.3.3 PAM vs. SLM

The basic hypothesis that non-native sounds will be assimilated by speakers into an already existing category in the native system is common to both SLM and PAM; how-
ever, these two models differ in several ways. One of these differences is that SLM was originally developed to predict and explain perception patterns of listeners involved in second language learning and acquisition, whereas PAM’s main focus is on non-native speech perception by “naïve” listeners that have had no exposure to the language of the contrasts that they are trying to perceive. This might raise the question of whether these two models are entirely comparable, since they are not designed to explain the behavior of the same types of listeners. Although active language learners and completely “naïve” listeners are two different types of listeners, one could interpret “naïveness” as the beginning stage of the speech learning process, at least for the adult learners. If the listeners discussed above were to begin to learn Zulu, for example, they would then be solely within the scope of SLM and, presumably, they might begin to form new categories, depending on factors such as age of arrival, amount of L1/L2 input, etc. (Flege 1995). For a more detailed account of the similarities and differences between PAM and SLM and also for an extension of PAM for the second language learning realm, refer to Best and Tyler (2007).

Another difference between these models is what each one posits as the unit of perception. While the perceptual objects in SLM are the acoustic cues in the non-native speech signal, PAM is based on “the ecological direct-realist premise that the focus of speech perception is on the information about the distal articulatory events that produced the speech signal” (Best and Tyler 2007:22). For this reason, PAM assumes the principles of Articulatory Phonology (Browman and Goldstein 1989, 1990, 1992), which posits that the timing and coordination of articulatory gestures necessary for speech function as the primitives for phonological organization.

Another difference between SLM and PAM, for the purposes of the present investigation, is each model’s stance on the importance of the phonological position or context
in which the non-native contrast is presented and its relationship to the native phonological inventory. In SLM the relationship of the non-native contrast to the native one is position-specific; whereas, in PAM the assimilation of the non-native phone takes into consideration phonetic-articulatory patterning “at both the abstract contrastive level and, importantly, at the level of non-contrastive gradient phonetic detail” in order to predict the assimilation patterns of the non-native sounds (Best and Tyler 2007:22). In SLM the acoustic similarity in the perceptual space will be based on a native phonetic category in the same position, while in PAM the similarities in the higher-order invariants, which should be relatively constant across positions, are the targets for assimilation. The phonetic variation between the native in gestures of the native and the non-native sound determine the goodness of fit of the non-native sound to the native category.

Several studies have uncovered results that are problematic for both models’ interpretation of the role of position. In Flege (1989) Chinese listeners were just as accurate as English listeners perceiving the English /t/-/d/ contrast in word-final position, despite the fact that the same contrast does not occur in final position in Chinese but does occur in initial position. Another problematic result is the dental-to-velar assimilation patterns in that French listeners tend to perceive onset /tl/ and /dl/ clusters as /kl/ and /ql/, respectively, despite the fact that French allows such clusters in word medial position (Hallé et al. 1998). Hallé and Best (2007) find similar results with French and English listeners with similar stop+/l/ clusters in Hebrew, showing that phonological position affects the perception and assimilation of non-native sounds. However, Hallé et al. (2003) find that even Hebrew listeners, for whom onset /tl/ clusters are legal, experience difficulty discriminating them, which might suggest a universal perceptual difficulty particular to this cluster (although, see Hallé and Best 2007:2912 for a description of the gestural
organization differences between /tl/ and /dl/, which might account for this perceptual difficulty).

### 3.3.4 Sociolinguistic variation and cross-language speech perception

One aspect that neither model addresses explicitly is the effect that native sociolinguistic categories might have on the assimilation of non-native contrasts. Best et al. (2009) have shown that at an early age (19 months) listeners are able to disregard non-native dialect-specific phonetic variation in defining a word’s identity, suggesting that they are tuning into the higher order invariants, rather than dialect-based phonetic differences, in categorial discrimination. However, other studies have shown that listeners have categories for dialectal variation, which are affected by the socially indexed variation within the dialect region (Hay et al. 2006, Clopper et al. 2005, Clopper and Pisoni 2006). If this is so, presumably socially dependent categories to which particular phones are associated could also affect the assimilation of non-native phonemes. In essence, a social variant could function as a phonemic category.

![Table 3.1: Schema of the Social Variant Assimilation pattern of a non-native contrast to two non-contrastive but socially distinct sociophonetic variants.](image)

In a case like the one presented in Table 3.1, two contrasting phonemes /rt/ and /a/ in Language A are perceived by listeners who speak Language B. For speakers of Language B, the phones [r] and [a] are not phonemically contrastive, but variants of the phoneme /rt/; however, each of the phones might be associated with speakers of differ-
ing social characteristics or associated with differing speech styles, notated here by the social variant sub-indices. For instance, in American English the segments [r] and [ɹ] are contrastive, distinguishing minimal pairs like Betty [bɛrɪ] and berry [bɜːri]. On the other hand, in Puerto Rican Spanish [ɾ] and [ɹ] are two non-contrastive variants of the same category /ɾ/ in coda position (/ar.ma/: [ar.ma]~[aɹ.ma] “weapon”); however, [ɾ] is characteristic of very careful speech, whereas [ɹ] is characteristic of casual speech (Simonet et al. 2008). In cases like these, will listeners of Language B (PR Sp.) have good discrimination of the Language A (Am. Eng.) contrast? If so, it would be evidence that social variants can, in essence, function as native categories in terms of perceptual assimilation, since the listeners might treat this non-native contrast as they would a TC assimilation. Or is it the case that, as is shown in the Jamaican English results of Best et al. (2009), the listeners will show poor discrimination given that they assimilate the non-native contrasts to one native category (SC Assimilation under PAM)?

Given the distribution of the stop voicing contrasts as well as the sociophonetic variation of /ʒ/, the perception of non-native obstruent voicing contrasts by Buenos Aires Spanish listeners might shed some light on both the effects of phonological position as well as native sociophonetic variation on cross-language speech perception.

### 3.3.5 Test case: voicing contrasts and voicing variation in Spanish

#### 3.3.5.1 Stop voicing contrast

Spanish presents a phonemic distinction between voiced and voiceless stops, where /p, t, k/ contrast with /b, d, g/. However, this voicing distinction is realized differently depending on the phonological context in which the stops occur. In phrase initial position (as well as after nasals and after /l/ for homorganic stops) the phonemic contrast is realized by a difference in voicing, where the phonologically voiceless stop is realized with short
lag VOT and the phonologically voiced stop is realized with negative VOT, e.g. /pV/ vs. /bV/ is realized as [pV] vs. [bV] (Lisker and Abramson 1964, Martínez Celdrán and Fernández Planas 2007). On the other hand, the voicing contrast in intervocalic position is realized not only by a difference in voicing, but also by a difference in the manner of articulation, whereby the phonologically voiced stops spirantize and the phonologically voiceless stops retain their occlusion. For example, the contrast of /VpV/ vs. /VbV/ is realized as [VpV] vs. [VbV], where the voiced stop spirantizes. Traditionally, the allophones of the voiced stops in intervocalic position have been described as voiced fricatives [β, δ, χ] (Lapesa 1981, Quilis 1999); however, these phones are usually realized not as fricatives but as approximants [β, δ, χ] (Martínez-Celdrán et al. 2003, Santagada and Gurlekián 1989). Therefore, regardless of the constriction degree of the intervocalic segment, the realization of /VbV/ as *[VbV] is not a licit realization for native Spanish speakers. For this reason, the canonical voicing distinction for stops in Spanish is present only in phrase initial (or post-nasal) position, voiced stops will spirantize and be realized as voiced approximants elsewhere: /b, d, g/ → [β, δ, χ].

The phenomenon of voiced stop spirantization has been described as a classic example of complementary distribution. However, several studies have shown that the spirantization of voiced stops in Spanish is in fact a gradient process, ranging from occlusion to deletion, and its realization is affected by factors such as speech rate (Soler and Romero 1999), lexical stress (Cole et al. 1999, Ortega-Llebaria 2004) duration of the stop gesture (Parrell 2010), place of articulation (Lavoie 2001, Colantoni and Marinescu 2010) nature of the flanking segments (Hualde et al. 2010, Cole et al. 1999, Kingston 2008), as well as dialect of Spanish (Amastae 1995, Lewis 2001). One point of dialectal variation is which positions favor occlusion vs. spirantization, for example, in Colombian Spanish any voiced stop following a consonant will be occluded, e.g. /desde/ >
In other varieties, like Puerto Rican Spanish, voiced stops are occluded following an aspirated/deleted coda /s/, e.g. /desde/ > [dehde] or [dede]. However, in other /s/ aspirating varieties of Spanish, such as Buenos Aires Spanish, the stop is spirantized in those positions, regardless of the debuccalization of the previous consonant /desde/ > [dehde]. These same spirantization patterns hold in cases where the voiced stop occurs across a word boundary; however, even though there is a greater degree of constriction, i.e. occlusion, in phrase initial position, the same is not found in phrase internal word boundaries (Cole et al. 1999).

There is also variation in the production of phonologically voiceless stops in Spanish as well. Across different varieties of Spanish, voiceless stops are often realized with partial or full voicing during closure (Machuca Ayuso 1997, Lewis 2001, Martínez Celdrán 2009, Parrell 2010, Torreira and Ernestus to appear). Machuca Ayuso (1997) found in her corpus of conversation speech more than 40% of phonologically voiceless stops /p, t, k/ were realized as [b, d, g]. Torreira and Ernestus (to appear) found that in the Nijmegen Corpus of Spontaneous Spanish (Torreira and Ernestus 2010) 32.7% of voiceless stops presented complete voicing during stop closure. Figure 3.1 below shows an example reproduced from Torreira and Ernestus (to appear) of word autopista, where the complete voicing for both the /t/ and /p/ can be seen clearly. This phenomenon of voicing of phonologically voiceless stops has been reported for many varieties of Spanish also; however, it also presents dialectal variation. Colantoni and Marinescu (2010) find relatively few cases of voicing of /p, t, k/ in their analyzed corpus of speech from the Argentinean provinces of San Juan and Corrientes.

Aside from voicing, many studies have also found cases of spirantization of voiceless stops as well, where they are realized as voiceless or voiced approximants (Machuca Ayuso 1997, Lewis 2001, Parrell 2010, Torreira and Ernestus to appear). This is potentially
problematic for the description of the voicing distinction of stops in Spanish, since there 
is an overlap in the realization of stops such that both voiced and voiceless ones can be 
realized as $[\beta, \delta, \chi]$. Martínez-Celdrán (2008) has proposed that the distinction is not 
one based on voicing necessarily but on tenseness; however, see Parrell (2010) for an 
articulatory study showing that $/p/$ and $/b/$ differ in duration and constriction target.

Considering all this research, it cannot be said that voiced stop spirantization is 
categorical or that phonologically voiceless stops in Spanish are always phonetically 
voiceless. However, what can be gleaned from most of the studies is that phrase initial 
position and intervocalic positions seem to be the two poles in the continuum. What is 
common to all varieties of Spanish is that voiced stops in $#CV$ position favor occlusion 
and voiced stops in $VCV$ position favor spirantization. $^1$ Similarly, for voiceless stops, 
$#CV$ position should favor voiceless realizations, whereas in $VCV$ position is where le-

$^1$However, see Cole et al. 1999 for unexpected results showing that flanking $/a/$ favors more constricted 
realizations of $/g/$.
nition to a voiced stop (or even approximant) would be more likely to occur (although the lenition of voiceless stops might not be as prevalent cross-dialectically as the lenition of voiced stops, Colantoni and Marinescu 2010).

3.3.5.2 Variable fricative voicing in Spanish

The phonological description of voicing for fricatives is relatively simple, when compared to the situation of stops described above. As opposed to stops, there is no voicing contrast in fricatives; General Spanish has only phonologically voiceless fricatives in its inventory: /f, s, x/ (and, in peninsular Spanish, also /θ/). There is no phonological process analogous to spirantization of voiced stops that occurs with fricatives in intervocalic position. There is allophonic variation in that the voiceless fricatives are usually realized as voiced in coda position when the onset of the following syllable is a voiced consonant, e.g. /mis.mo/ > [miz.mo] “same”, /mas.de/ > [maz.de] “more than” (Hualde 2005:159-161). It should be noted that this allophonic /s/>[z] variation is not common in BAS because of the pervasiveness of coda /s/ aspiration as the educated norm.

Despite this general description, phonologically voiceless fricatives, similarly to voiceless stops, are sometimes produced as voiced (Martínez-Celdrán 2008). In the previous chapter, it was shown that in Buenos Aires Spanish variable voicing of phonologically voiceless /s/ is influenced by phonological position as well as by speech style/rate. Buenos Aires Spanish presents a particularly interesting example, because aside from the phonetically conditioned variation in fricative voicing, it also presents socially conditioned voicing in the postalveolar fricative /ʃ/, affected by factors such as age and social class, as seen in Chapter 2.

Given the complex phonological and sociolinguistic interactions of obstruent voicing in Buenos Aires Spanish, how would BAS listeners perceive a canonical voicing
distinction in stops in initial vs. intervocalic position? Moreover, how would they perceive an analogous voicing distinction in fricatives, where there is no native voicing distinction but rather sociolinguistically differentiated variation? The two models of cross-language speech perception discussed above would make differing predictions as to the discrimination rates and assimilation patterns of the non-native voicing contrasts as perceived by BAS listeners.

### 3.4 Goals of the study

The voicing discrepancy in the obstruent inventory in BAS, as well as the positionally determined process of spirantization of voiced stops, presents a testing ground for the predictions made by SLM and PAM in terms of the position specificity of perception as well as the different possible perceptual assimilation patterns that can result from the perception of non-native voicing contrasts that intersect with patterns of socially determined voicing variation in the native speech community. The present study attempts to answer the following questions:

1. Given that in Spanish the canonical voicing distinction is only implemented in phrase initial position, is the accuracy of discrimination affected when listeners are presented with an occluded stop voicing contrast in a non-native intervocalic position?

2. Given that Buenos Aires Spanish exhibits a voicing distinction only for stops, can listeners transfer this native voicing distinction and discriminate between analogous non-native voiced and voiceless fricatives? If not, are all fricative contrasts discriminated equally poorly?
3. Does the presence of socially stratified native variation, particularly with respect to /ʒ/ ([ʒ]~[ʃ]), affect the discrimination of an analogous non-native fricative voicing contrast?

4. Are the speakers who are more likely to be voicers themselves (i.e. upper class speakers and older speakers) more accurate at distinguishing the non-native contrast /ʃ/-/ʒ/?

3.5 Perception experiment 1: AX discrimination

In order to answer the questions posed above, an AX discrimination experiment was conducted using Brazilian Portuguese (BP) as the non-native stimulus language. The reason for using BP is that it is relatively similar to Spanish, i.e. most of the phonemes of Spanish are contained in the phonemic inventory of Brazilian Portuguese (Mattoso Câmara 1972). Most important for this study is that BP has a voicing distinction not only for stops but also for fricatives. Also, voiceless stops are described as voiceless unaspirated and voiced stops as prevoiced, i.e. with negative VOT, in both initial and intervocalic position, i.e. phonologically voiced stops in intervocalic position are always realized as occlusives, they do not spirantize as in Spanish (Barbosa and Albano 2004).

3.5.1 Materials

For this experiment six obstruent voicing pairs were used, three stop pairs (p/b, t/d, k/g) and three fricative pairs (f/v, s/z,ʃ/ʒ). The stimuli included obstruents in initial position (CV, CVIV: sa/za, kili/gili) as well as in intervocalic position (VCV, IVCV: asa/aza, liki/ligi). The disyllabic stimuli with the filler consonant /l/ were included in order to
make the initial and intervocalic tokens more comparable in length and to make them more word-like. The stimuli were produced in 4 vocalic contexts: aCa, aCi, iCi, iCa (e.g. asa/aza, asi/azi, isi/izi, isa/iza). All the disyllabic stimuli were stressed on the penultimate syllable.

The stimuli for the experiment were recorded by a female native speaker of Brazilian Portuguese, 27 years old, from São Paulo, Brazil. She had been living and working in Buenos Aires for several years and was Portuguese-Spanish bilingual. The only sequences which merited special attention were those containing /ti/ and /di/ sequences. In most varieties of Brazilian Portuguese dental plosives are realized as affricates before high front vowels, e.g. /ti/ > [tʃi], /di/ > [dʒi] (Barbosa and Albano 2004). For the stimuli containing /ti/ and /di/ sequences, the speaker was instructed to produce the consonants as plosives ([ti] and [di]), both in initial and intervocalic position.

The stimuli were recorded in a quiet room using a Sony MZ-RH910 Hi-MD recorder and Shure WH20XLR head-mounted microphone. She read each stimulus 8 times, out of which two tokens were selected that were the most similar in terms of length, intonation, voicing duration, etc. These two tokens were matched as closely as possible with the 2 corresponding tokens of the opposite voicing category. The tokens were checked by 3 native Brazilian Portuguese speakers to ensure that they were good exemplars of the phonemic category in Brazilian Portuguese.

### 3.5.2 Participants

Twenty-one speakers of Buenos Aires Spanish were recruited for this experiment (See Table 3.2). They were all native speakers of Buenos Aires Spanish, i.e. they were all born and raised in the Buenos Aires metropolitan area. There were 9 males and 11 females, 9 of the participants were from northern Buenos Aires and 11 were from
southern Buenos Aires. In accordance to the discussion in Section 1.3.1.2, the speakers’
neighborhood was used as a proxy for social class (north = upper class, south = middle
class). Their ages ranged from 21-65, with a mean age of 33. One speaker (subject U)
was excluded because of near-native fluency in French. Out of the remaining 20 speak-
ers, none spoke other languages natively. Most of them had studied other languages in
school (mostly English, Italian, and French); however, none of these participants had
previously studied Portuguese.

<table>
<thead>
<tr>
<th>Subject #</th>
<th>Sex</th>
<th>Age</th>
<th>Neighborhood</th>
<th>Zone/Social Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>M</td>
<td>26</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>B</td>
<td>F</td>
<td>65</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>C</td>
<td>M</td>
<td>34</td>
<td>Mármol</td>
<td>South/Middle</td>
</tr>
<tr>
<td>D</td>
<td>M</td>
<td>33</td>
<td>Munro</td>
<td>North/Upper</td>
</tr>
<tr>
<td>E</td>
<td>M</td>
<td>31</td>
<td>San Isidro</td>
<td>North/Upper</td>
</tr>
<tr>
<td>F</td>
<td>M</td>
<td>38</td>
<td>Once</td>
<td>South/Middle</td>
</tr>
<tr>
<td>G</td>
<td>F</td>
<td>31</td>
<td>Burzaco</td>
<td>South/Middle</td>
</tr>
<tr>
<td>H</td>
<td>F</td>
<td>28</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>I</td>
<td>F</td>
<td>32</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>J</td>
<td>M</td>
<td>25</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>K</td>
<td>F</td>
<td>62</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>L</td>
<td>F</td>
<td>29</td>
<td>Barrio Norte</td>
<td>North/Upper</td>
</tr>
<tr>
<td>M</td>
<td>M</td>
<td>27</td>
<td>Caballito</td>
<td>South/Middle</td>
</tr>
<tr>
<td>N</td>
<td>F</td>
<td>21</td>
<td>Ezeiza</td>
<td>South/Middle</td>
</tr>
<tr>
<td>O</td>
<td>F</td>
<td>26</td>
<td>Retiro</td>
<td>North/Upper</td>
</tr>
<tr>
<td>P</td>
<td>M</td>
<td>33</td>
<td>San Isidro</td>
<td>North/Upper</td>
</tr>
<tr>
<td>Q</td>
<td>F</td>
<td>24</td>
<td>Retiro</td>
<td>North/Upper</td>
</tr>
<tr>
<td>R</td>
<td>M</td>
<td>24</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>S</td>
<td>F</td>
<td>34</td>
<td>Villa Crespo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>35</td>
<td>Barrio Norte</td>
<td>North/Upper</td>
</tr>
<tr>
<td>U</td>
<td>F</td>
<td>59</td>
<td>Recoleta</td>
<td>North/Upper</td>
</tr>
</tbody>
</table>

Table 3.2: Demographic information of the participants in the AX experiment.
3.5.3 Experimental task

The designed experiment was an AX discrimination task, in which the participants had to listen to two tokens that belong to either the same or different phonological categories in the non-native language. The listeners then had to determine whether the two tokens they heard were the same or different from each other. This was a categorial perception task, since both the “same” and the “different” trials were physically different tokens of the same non-native phonemic category. An AX paradigm was chosen, as opposed to an AXB or oddity discrimination paradigm, in order to keep the memory demands on the subjects relatively low. In earlier studies (Pisoni 1973, Werker and Logan 1985) found that subjects’ accuracy rates in discrimination for short (250ms) interstimulus intervals (ISI) was much higher when compared to longer (1500ms) interstimulus intervals. They concluded that in the shorter ISI the subjects were making use of acoustic perception, which accounts for their relatively higher accuracy, whereas in the longer ISI, because of necessarily higher memory demands, the perceived lower-level acoustic differences were degraded or associated to higher-level phonological categories.\(^2\) The aim of this experiment is somewhat intermediate, since it attempts to tap into the subjects’ sub-phonemic categories (e.g. sociolinguistic variants). Therefore it was not intended that the listeners rely either solely on the acoustic information, since all the tokens were physically different, or solely on the phonemic categories (since the sociolinguistic variants fall within that category). For this reason an AX paradigm was chosen but with an intermediate ISI (500ms), to facilitate discrimination but not to overly encourage higher level-phonemic classification.\(^3\)

\(^2\)However, Davidson (2011) suggests that the high discrimination rates based on the availability of lower-level acoustic information may not hold for longer, word-like stimuli. See Section 3.5.6 for further discussion.

\(^3\)Although the AX task has been reported as the paradigm that enables listeners to discriminate more easily between the stimuli, Gerrits and Schouten (2004:362) suggest that subjects might be biased to answer “same” unless they are very sure of their decision. See Section 3.5.6 for further discussion.
The AX discrimination task was presented to the participants using Praat MFC Experiment (Boersma and Weenink 2010). Participants had to choose whether the two tokens that they heard were the same or different by clicking on the corresponding box on the screen. They first completed a training session with nasal place contrasts native to Spanish (e.g. nu/mu), which was also a categorial discrimination task, since the “same” tokens were physically different tokens of the same category. During this training session, the listeners were presented the correct answer on the screen, either SAME or DIFFERENT, as they listened to each pair. They were instructed to focus on the consonant, ignoring any intonation or vowel length differences. After the training session they were asked whether they had any questions about the instructions.

For the experimental session with the Brazilian Portuguese stimuli, participants were instructed to listen to pairs of words in a foreign language. They had to indicate whether the pair they heard consisted of two examples of the same word or whether they were different words. It was stressed to the participants that the stimuli that they were hearing were in a foreign language, in an attempt to encourage them to abstract away from their native Spanish phonology and pay attention to phonetic cues that might be different in the obstruent, but might not be phonemic in Spanish. (See Appendix C for the instructions for the AX task).

The stimuli were presented in two blocks, one in intervocalic and one in initial position. The order of these blocks was counterbalanced. The trials were blocked by position in order to explain the task more easily to the participants and to ensure that they attend to the differences in the obstruent that always occurred in the same position, rather than to any differences in the surrounding segments. There were 4 trials for each position and vowel combination (2 same and 2 different) for each of the six voicing pairs, for a total of 368 trials per listeners. The trials within blocks were randomized. The experiment
was run on a Dell Inspiron 600m laptop and the stimuli were played over circumaural Sennheiser HD 202 headphones in a quiet room.

### 3.5.4 Variables and predictions

#### 3.5.4.1 Position

One of the independent factors in this investigation is the position of the obstruent, in order to determine whether discrimination accuracy is affected by a stop voicing contrast presented in a non-native intervocalic position. The two possibilities for the phonological position of the obstruent in this experiment were initial (#CV, #CVtV) and intervocalic (VCV, lVCV).

The Perceptual Assimilation Model and the Speech Learning Model each make different predictions for the accuracy rates of the BAS listeners making these discriminations in BP stop voicing contrasts in intervocalic position. (Since SLM and PAM would not make differing predictions for the stop pairs in different places of articulation, in the discussion of position, “t”,”d”, and “ð” will be used to symbolize voiceless occlusion, voiced occlusion, and spirantization, respectively, for all stop voicing pairs.)

For the stop voicing contrast in initial position SLM would predict high accuracy in discrimination, since the acoustic cues of the voicing contrast in BP are similar to the ones in Spanish for that position. For the stop voicing contrasts in intervocalic position, the voiceless stops in BP will be related to the voiceless stops in intervocalic position in Spanish; however, the voiced stops in intervocalic position in Brazilian Portuguese [VdV] are incongruous with the spirantized positional allophone of /VtV/ that is most common in Spanish: [VðV]. The BP [VdV], although voiced, because of its occlusion and burst, would be acoustically more similar to the positional allophone for Spanish /VtV/. Furthermore, as discussed in Section 3.3.5, the fact that in intervocalic position
phonologically voiceless stops often present voicing during closure, presents another reason why BP [VdV] would be associated with the positional allophone [VdV] of the Spanish in the sequence /VtV/. This would increase the perceived acoustic similarity of BP [VtV]- [VdV] for Spanish listeners, thereby making the discrimination harder. For these reason, the Speech Learning Model would predict that Spanish listeners would have lower accuracy rates for the BP stop voicing contrast in intervocalic position.

The predictions that Best’s Perceptual Assimilation Model would make for the phrase initial stops are similar to the ones made by the Speech Learning Model: initial stops in Brazilian Portuguese, [tV]-[dV], should assimilate to the analogous categories in Spanish as a Two Category Assimilation, /t/-/d/. For this reason, listeners should exhibit high levels of accuracy. For the stop voicing contrast in intervocalic position, however, PAM’s predictions differ from those of SLM. In PAM, positionally-defined allophones do not play such a defining role as in SLM, since the listeners are also attending to higher-order invariants that distinguish the articulation of the voiced and voiceless stops, which would be constant across position. Therefore, the BP [VdV] should be assimilated to the voiced stop Spanish category /d/, regardless of the fact that occluded allophones [VdV] for native /VdV/ do not occur in that position in Spanish. In these cases, PAM would predict levels of accuracy in discrimination similar to those for the stop voicing contrast in initial position, which is analogous in Portuguese and Spanish. This type of assimilation would also be a Two Category Assimilation, where BP [VtV] maps onto Spanish /t/ and BP [VdV] maps onto Spanish /d/; therefore, listeners should exhibit high accuracy rates. The obstruent in [VdV], perhaps might be a worse fit for the Spanish /d/ category in that position; however, since the stimulus that the listeners are comparing it to belongs to another category, this should not affect their accuracy in discrimination between /t/ and /d/.
3.5.4.2 Voicing pair

The other linguistic variable under investigation is voicing pair, which will help determine whether listeners are able to transfer a voicing distinction in stops to aid in the discrimination of a non-native voicing contrast in fricatives. Here “voicing pair” is used as a cover term to convey different combinations of place of articulation, manner of articulation, and voicing. In this study there are six voicing pairs, three stops and three fricatives: p/b, t/d, k/g vs. f/v, s/z, ʃ/ʒ. The general prediction here is that listeners should perceive the voicing contrast in stops more accurately than the voicing contrast in fricatives, since Spanish maintains a voicing contrast in stops, but not in fricatives. However, SLM should predict an interaction between position and voicing pair for the stops, since Spanish lacks an occlusive allophone intervocalic ally for the voiced stops, whereas there should not be an interaction with position for the fricative voicing pairs, since the contrasts are non-existent in BAS regardless of position.

Best’s PAM would also predict that BAS listeners should perform better on stops than on fricatives, because all of the Brazilian Portuguese stop contrasts are cases of TC assimilation. The non-native Portuguese [p, t, k] map onto the native BAS categories /p, t, k/ and the non-native Portuguese [b, d, g] map onto the native categories /b, d, g/. However, PAM would not predict an interaction between position and voicing pair, predicting high accuracy rates for all the stop contrast in both positions.

In the case of the fricative voicing pairs, however, PAM would not make one uniform prediction for their accuracy rates, since the voiced fricatives could have differing assimilation patterns based on the articulatory similarities to other Spanish phonemic categories. The possible assimilation patterns for each fricative voicing pair are discussed below.
Considering the labiodental fricative voicing pair fl/v, [v] is not described as a member of the Spanish phonological inventory; however, it is articulatorily most similar to /fl/; therefore, one could predict that it will be assimilated to the native /fl/ category, but not as a very good fit to that category. This would be a case of Category Goodness Assimilation, where the listeners should exhibit intermediate levels of accuracy that should be significantly lower than those for the stop voicing contrasts.

There is another possible prediction for the type of assimilation that the BP voicing pair fl/v could have. The graphemes <f, v, b> all exist in Spanish orthography, where <f> represents the phoneme /fl/ and both <v> and <b> represent the phoneme /b/. This is one of the relatively few points in Spanish orthography that actually causes confusion for Spanish speakers while spelling, since there are usually no phonetic cues in casual speech for the listener to determine whether a word containing the phoneme /b/ should be written with a <b> or a <v>. However, [v] is an unlikely, yet possible, allophone of /b/. It may occur in very extreme cases of dictation or when the speaker wants to emphasize the spelling of a certain word to their interlocutor. Hualde (2005:5) mentions that “some school teachers, especially in Latin America, also insist on artificially introducing a distinction in pronunciation between orthographic v and b as a way to aid in the memorization of the standard spelling of words.” This practice was confirmed in a separate interview conducted with subject 14 from the production study, who is an elementary school teacher in Buenos Aires. Furthermore, Lustig (2011) conducted a preliminary study on the realization of orthographic <b> and <v> in this corpus of BAS. She found that, although not categorical, in reading tasks the cases of orthographic <v> in intervocalic position tend to be realized as actual fricatives (as opposed to the more canonical approximant) more frequently than the cases of <b> in intervocalic position. Although this result is not compared to spontaneous speech, this greater in-
cidence of fricative realizations for <v> than for <b> could presumably be due to an increased degree of attention paid to orthography in read speech. Taking these facts into consideration, then the BP f/v contrasts could result in a Two Category Assimilation, in which BP [f] is mapped onto native Spanish /f/ and BP [v] gets mapped onto native Spanish /b/. If this is the case, then listeners should exhibit high levels of accuracy in distinguishing this non-native voicing pair, comparable to those seen for the stop voicing pairs.

In the case of the BP voicing pair s/z, BP [s] would map onto its BAS counterpart /s/, since they are realized similarly in both languages. BP [z] is not part of the native Spanish phonemic inventory; however, it is most similar articulatorily to BAS /s/. This could be a case of Category Goodness assimilation, where non-native [z] is assimilated to the native /s/ category but as a worse fit. In these cases of Category Goodness Assimilation, according to PAM, listeners should exhibit intermediate accuracy rates in distinguishing the non-native contrast. If this is the case, listeners discriminating the s/z voicing pair should have lower accuracy rates than the rates for the stop voicing pairs.

In Chapter 2 it was shown that /s/ varied in production, such that it exhibited voicing partial or complete voicing during frication (recall the example in Figure 2.4 of subject 5’s pasajero, where the /s/ in /asa/ was realized as fully voiced, i.e. [z]). For this reason, it could be postulated that [z] be assimilated as an equally good example of the /s/ category in Spanish, in which case the s/z voicing pair would a SC assimilation. If

---

4Both in BAS and in BP coda /s/ is a very variable in terms of its (non)realizations; however, all of the experiments conducted focus on the realization of obstruents in syllable onset position, where the realization of /s/ is much more stable in both languages.

5Recall that in many varieties of Spanish [z] is an allophone of coda /s/ whenever it is followed by a voiced consonant, e.g. /mis.mo/ > [miz.mo] ‘same’. However, this is not a characteristic of BAS, since preconsonantal coda /s/ is consistently aspirated: [mihmo]. In BAS, as opposed to other /s/-aspirating varieties of Spanish, aspiration of preconsonantal /s/ is the educated norms and occurs consistently in both casual as well as very formal speech styles (Lipski 1994); therefore, the likelihood of a [z] occurring in a word like /mismo/ is very low.
this were the case, participants would be expected to show low accuracy rates for this voicing pair. Although, this assimilation pattern is a possibility, it might not be very likely, given the limited range of variation for voicing of /s/ as compared to the voicing of /ʃ/ in the community as a whole. It is true that for some speakers the voicing rates of /ʃ/ are not different to those of /s/, but even those speakers are exposed to much more voicing variation for /ʃ/ than they are for /s/ in the speech community. Given this relatively limited range of exposure to [z], as compared to [ʒ], it is less likely that they would associate both members of the s/z voicing pairs as equally good exemplars of the /s/ phonemic category in BAS.

The case of the postalveolar fricative voicing pair is different from the other fricatives. The phonemic voicing contrast present in Brazilian Portuguese between /ʃ/ and /ʒ/ does not exist in Buenos Aires Spanish; however, both the voiceless and the voiced postalveolar fricatives are allophones of the same postalveolar fricative phoneme in Buenos Aires Spanish. Although speakers might vary in their own voicing rates for /ʒ/, as demonstrated in Chapter 2, all speakers are exposed to variants ranging from /ʃ/ to /ʒ/ in daily interactions and in the media. Taking this variation in the speech community into consideration, PAM would predict that the listeners would have a Same Category Assimilation pattern with the non-native voicing pair /ʃ/-. Listeners should assimilate BP [ʃ]-[ʒ] to the same postalveolar fricative category equally well and, thus, should exhibit relatively low accuracy rates distinguishing between the /ʃ/ voicing pair.

Given the hypothesized assimilation patterns under PAM, several discrimination accuracy scales could be postulated, as shown in the Tables 3.3-3.6 below.

\[
\begin{array}{cccc}
p/b, t/d, k/q & f/v & s/z & /ʃ/ \\
TC & CG & CG & SC \\
\end{array}
\]

Table 3.3: Discrimination accuracy scale (with f/v as CG).
\[
p/b, \ t/d, \ k/g = f/v > s/z > \text{f/3}
\]
Table 3.4: Discrimination accuracy scale (with f/v as TC).

\[
p/b, \ t/d, \ k/g \geq f/v \geq s/z > \text{f/3}
\]
Table 3.5: Discrimination accuracy scale (with s/z as CG).

\[
p/b, \ t/d, \ k/g \geq f/v > s/z = \text{f/3}
\]
Table 3.6: Discrimination accuracy scale (with s/z as SC).

The stop voicing pairs, since they all occur in Spanish, should be Two Category Assimilation types and thus have the highest accuracy rates for all the accuracy scales. The case of f/v, as discussed above, could have two options. It could be a case of Category Goodness assimilation (if the BP [v] maps onto the articulatorily closest fit in the Spanish inventory: /f/), in which case lower levels of accuracy should be expected as compared to the stop voicing pairs, as shown in Table 3.3. Alternatively, the f/v voicing pair could be a Two Category Assimilation (if the BP [v] maps onto the native category /b/, which is clearly different from the native /f/ category that BP [f] maps onto), in which case listeners should exhibit high levels of accuracy comparable to those of the stop voicing contrasts, as shown in Table 3.4. If the s/z is a Category Goodness assimilation it should exhibit lower levels of accuracy as compared to the Two Category Assimilation types (p/b, t/d, k/g) and higher rates compared to \text{f/3}, as seen in Table 3.5. If both f/v and s/z turn out to be Category Goodness assimilation patterns, then they should show lower accuracy rates than the stops but higher rates than \text{f/3} (and perhaps similar rates compared to each other.) However, if s/z turns out to be Same Category assimilation, then it should exhibit similar accuracy rates to \text{f/3}, as seen in Table 3.6. Finally, in the case of \text{f/3}, the two non-native phones would be mapped onto the same
phonemic category as a Same Category Assimilation, and therefore the listeners should exhibit the lowest accuracy rates discriminating this non-native BP contrast, as pictured in Tables 3.3-3.6.

An alternate hypothesis for the assimilation of /ʃ/ could be one that takes into account an effect of the sociolinguistic variation in the native speech community for the fricative voicing pair /ʃ/ and /ʒ/. It is true that in Buenos Aires Spanish [ʃ] and [ʒ] are two possible realizations of the same postalveolar fricative phonemic category; however, these two phones have a different social distribution. As discussed in Chapter 2, age and social class were significant factors determining the variation of /ʒ/. Given these results, one could posit a social variant hypothesis, where one member of the voicing pair is associated with certain social characteristics and the other member of the pair with other characteristics. In essence the social categories could act as native phonetic categories for purposes of discrimination.

<table>
<thead>
<tr>
<th>BP</th>
<th>BAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-native</td>
<td>Native</td>
</tr>
<tr>
<td>/ʃ/</td>
<td>[ʃ]middle/younger</td>
</tr>
<tr>
<td>/ʒ/</td>
<td>[ʒ]upper/older</td>
</tr>
</tbody>
</table>

Table 3.7: Schema of the Social Variant Assimilation pattern of a non-native BP contrasts by BAS listeners.

In the schema in Table 3.7, listeners might discriminate /ʃ/ more accurately by associating the non-native contrast to a native social distinction, because it would essentially function as a Two Category Assimilation in PAM. Therefore, PAM and the Social Variant Assimilation would predict two differing voicing perception scales, where the crucial difference lies in the accuracy of discrimination of /ʃ/, where it should be higher than 6

6The voice of the stimulus reader, a 27-year-old female, could potentially bias the speakers toward one category, mainly that of younger speakers. However, recall that in the production results there were young female voicers and devoicers depending on class, presumably the stimulus reader could be associated with either class group.
s/z in the social variant assimilation hypothesis because the assimilation pattern of /ʃ/ would mimic that of a TC category assimilation, much like what was posited for f/v as a TC, in Table 3.4.

\[
p/b, t/d, k/q = f/v \geq s/z > /ʃ/
\]

Table 3.8: A discrimination accuracy scale under Perceptual Assimilation Model.

\[
p/b, t/d, k/q = f/v = /ʃ/ > s/z
\]

Table 3.9: A discrimination accuracy scale incorporating Social Variant Assimilation Hypothesis.

If the results support a discrimination scale such as the one shown in Table 3.9, then it would be evidence for sub-phonemic sociolinguistic allophonic variants can function as separate perceptual categories in cross-language speech perception.

### 3.5.4.3 Social class

In Chapter 2 it was shown that class was a significant factor affecting the voicing rates of /ʃ/, as had been shown in some of the previous investigations of the phenomenon (Wolf 1984, Rohena-Madrazo 2008). In their study of the devoicing of the postalveolar fricative in Buenos Aires, Wolf & Jiménez (1979) state that speakers in their sample were not aware of the devoicing phenomenon. However, in a later study Wolf (1984) suggested that upper class speakers did show awareness of the devoicing of /ʃ/, and would monitor their voicing levels accordingly in opposition to the devoicing of the middle class. Furthermore, Colantoni (2008) found that speakers from higher social strata had an awareness of the voicing variation and attributed the voiced variants more frequently to fellow upper class speakers. She argues that for the upper stratum speakers
of BAS the voicing of /ʃ/ is a case of macro variation, i.e. they are conscious of the phenomenon, whereas for the lower stratum speakers it is a case of micro variation, i.e. they are unaware of it. Given this information, one might expect upper class speakers (i.e. northerners) to exhibit higher accuracy rates than the non-upper class speakers at distinguishing the /ʃ/ contrast.

3.5.4.4 Age

Age has long been associated as a factor affecting the devoicing of /ʃ/, which was confirmed in the production results of Chapter 2. It is not clear the extent to which the speakers in the community are aware of the devoicing of /ʃ/; however, if this change from below is indeed in its final stages, there might be signs of correction from above, which would probably come from the older speakers who are likelier to be voicers themselves. If this is the case, there should be a positive correlation between age and the accuracy rates in distinguishing the /ʃ/ contrast.

3.5.5 Results and discussion

For the statistical analyses of the AX discrimination experiment, the dependent variable used was d’ (d-prime). The d’ statistic serves as a sensitivity index that computes how easily a signal can be detected (Macmillan and Creelman 2005). In this experiment the signal is present in the “different” trials ([aɬa]-[aʃa]) and absent in the “same” trials [aɬa]-[aɬa]. A d’ score of 4.65 indicates that listeners have perfect discrimination, while a d’ score of 0 indicates that the listeners’ cannot discriminate at all and are effectively performing at chance.

---

7Although King (2009) found that among some younger speaker there is awareness of the voicing variable.
A mixed-design ANOVA with d-prime scores as the dependent measure was run with voicing pair and position as within-subject factors and class as the between-subjects factor. The results show a significant main effect of voicing pair, $F(5,90)=68.038, p<.001$, a non-significant main effect of position, $F(1,18)=1.723, p=.206$, and a non-significant effect of class, $F(1,18)=0.290, p=.597$. Of the interactions, only position*voicing pair was significant, although not very robustly, $F(5,90)=2.320, p=.050$. However, Tukey HSD post-hoc comparisons do not show any significant differences between #CV and VCV positions for any of the voicing pairs. The Tukey HSD post hoc comparisons for significant main effect of voicing pair reveal two groups of voicing pairs that are significantly different from each other: (p/b, t/d, k/g, f/v) vs. (s/z, S/Z). All the differences between the voicing pairs were significant with a $p<.001$ and none of the differences within the groups were significant with $p\geq .236$.

### 3.5.5.1 Position and voicing pair

Position was not found to be a factor significantly affecting the sensitivity of discrimination of non-native obstruent voicing contrasts, (#CV: $M=3.166$, $SD=1.507$; VCV: $M=3.061$, $SD=1.419$). This would seem to be evidence for PAM’s predictions; however, recall that SLM would predict an interaction between position and voicing pair where position would be a significant factor for the stops but not for the fricatives. This interaction was significant, the d’ results are shown in Figure 3.2 and the accuracy results are shown in Figure 3.3.

Post-hoc comparisons reveal that, although the voicing pair*position interaction was significant, there were no significant differences between #CV and VCV position within any of the stop voicing pairs or the fricative voicing pairs. All of the significant differences in the interaction were between either of sibilant voicing pairs and the group
Figure 3.2: Boxplots showing the sensitivity rates (d’) for the voicing pairs in initial and intervocalic positions. #CV is shown in white and VCV is shown in gray. “x/j” represents the #/3 voicing pair.

Figure 3.3: Accuracy rates for voicing pairs in initial and intervocalic positions. Results shown for % correct of the different trials.

formed by the stops+f/v, which were all at ceiling level sensitivity. For s/z and #/3, the mean sensitivity for VCV position (s/z: M=1.438, SD=0.847; #/3: M=1.280, SD=0.937) were lower than for #CV position (s/z: M=1.835, SD=1.212; #/3: M=1.394, SD=1.427);
however, none of these differences between or within these two voicing pairs was significant with $p \geq .708$. The fact that there is no difference between listeners’ sensitivity rates discriminating stops, regardless of position, is in line with the predictions of Best’s Perceptual Assimilation Model.

Figure 3.4 and Figure 3.5 show the sensitivity and accuracy rates, respectively, for each fricative voicing pair, collapsed across position, since the interaction did not reveal any significant differences in the pertinent voicing pairs. All the stop voicing pairs were perceived at ceiling level, with very high sensitivity rates and accuracy rates of 95% and above. This is expected, given that Spanish also has in its phonemic inventory these stops, so they are assimilated to their corresponding native voiced and voiceless stop categories. The accuracy in discrimination of fricative voicing contrasts is not as uniform as the accuracy in discrimination in the stop voicing contrasts. Notably, the accuracy of discrimination for f/v is at ceiling level, at 99%, significantly higher than for both s/z and f/z. In fact, there is no significant difference between the sensitivity rates for f/v and the sensitivity rates for any of the stop voicing contrasts. This supports the account of the f/v voicing contrast as a case of Two Category Assimilation rather than Category Goodness assimilation. Instead of assimilating to the corresponding native categories (as seen for p/b, t/d, k/g), the BP f/v voicing pair is assimilated to non-corresponding categories in the native inventory: [f] is mapped onto /f/, whereas [v] is mapped onto /b/.

Listeners’ levels of accuracy of discrimination for both sibilant voicing pairs (s/z and f/z) were quite low, significantly lower than the stop voicing pairs or f/v, but not significantly different from each other. The fact that one of these two sibilant voicing pairs shows socially differentiated variation does not seem to affect the discrimination of the voicing contrast. The resulting accuracy scale for voicing perception in the AX
Figure 3.4: Boxplots showing the sensitivity rates (d’) for the voicing pairs in initial and intervocalic positions. “x/j” represents the /ʃ/voicing pair. The two significantly different groups are (p/b, t/d, k/g, f/v) vs. (s/z, /ʃ/).

Figure 3.5: Accuracy of discrimination of non-native voicing contrasts, collapsed across position. Results shown for % correct of the different trials. An asterisk indicates a significant difference.
experiment can be seen in Table 3.10 below, which is quite similar to the accuracy scale shown in Table 3.6.

\[
\text{p/b, t/d, k/g} = \text{f/v} > \text{s/z} = \text{f/3}
\]

Table 3.10: Discrimination accuracy scale for voicing pairs in the AX experiment.

Recall that in the Social Variant Assimilation Hypothesis, the different sub-indices associated with each member of the postalveolar voicing pair would presumably aid in the discrimination; however, there does not seem to be any evidence for this. However, the voicing perception scale above does not coincide exactly with what would be predicted if s/z were a Category Goodness assimilation type. Judging by their accuracy rates, it seems that listeners are treating the s/z contrasts as a Same Category assimilation. Perhaps the variation in voicing found in the production of /s/, regardless of its more limited range as compared to BAS /3/, is enough for BP [z] to be assimilated to BAS /s/.

3.5.5.2 Social class and age

Neither social class nor any of the interactions with the other factors showed a significant effect on the accuracy of discrimination. If there is some sort of sociolinguistic awareness by part of the upper class on the devoicing of /3/, as proposed by Wolf (1984), then one would expect the upper class speakers to discriminate the f/3 more accurately than the speakers of the middle class. However, the means and standard deviations of the sensitivity rates in f/3 discrimination were not significantly different between the upper class and middle class listeners (upper: M=1.422, SD=1.216) vs. middle: M=1.267, SD=1.197). Both social class groups show equivalently poor rates of discrimination of the postalveolar fricative voicing pair. Furthermore, no significant correlation was found
between the listeners’ accuracy (d’) and their age, neither for the discrimination of all of the voicing pairs, r(240)=.109, p=.091, nor for the discrimination of l/3, r(20)=.264, p=.099.

### 3.5.6 Interim conclusion

Given the results of this AX experiment, the accuracy of discrimination of a native voicing contrast does not seem to be affected by the fact that it occurs in a non-native position. BAS listeners were not significantly less accurate at distinguishing fully occluded voiced stops from voiceless stops, despite the non-native status of intervocalic occluded voiced stops. These results are in line with the Perceptual Assimilation Model, since the higher-order invariants constant across positions would influence the assimilation patterns of the non-native phonemes to the native categories.

The fact that Spanish has a voicing contrast in stops does not seem to aid in the discrimination of a voicing contrast in fricatives. However, not all fricative voicing pairs have similar accuracy rates. Listeners have very poor discrimination rates for s/z and l/3. The exception in this case is the labiodental contrast f/v, presumably because in Spanish each element of this non-native voicing pair is assimilated to a different native category; therefore, listeners are more accurate in distinguishing it.

The presence of sociolinguistically conditioned variation in the speech community (i.e. l/3) does not seem to aid listeners at distinguishing the l/3 voicing contrast as compared to s/z. Variation in l/3 is differentially distributed in the speech community according to age and social class; however, listeners do not seem to be tapping into these socially differentiated phonetic cues, at least not in this type of task. Furthermore, there seems to be no effect of social class or age on the discrimination of 3/f or any other contrast. Even the upper class speakers, who had been described as the ones conscious
of the devoicing change Wolf (1984), do not have an advantage in the discrimination of
the voicing contrast. This goes against the Social Variant Assimilation Hypothesis as
stated in Table 3.7. BAS speakers might be losing awareness (or have lost it already)
of the devoicing change in their speech community. If they have no awareness of the
devoicing of /ʒ/, it could be that the fact that they are exposed to voicing variation both
in /ʒ/ and, to a lesser extent, /s/, influences the assimilation of the voiced BP sibilant
stimuli to the native sibilant categories (SC assimilation). Therefore, social variation
does not seem to affect, either positively or negatively, the discrimination of non-native
contrasts.

Although these conclusions might be valid, there are some problematic aspects to
the methodology used to investigate the perception of the voicing contrast. After the
experiment was over, several participants remarked that some contrasts were more dif-
ficult than others, which was expected. However, some of the participants remarked
that in some cases they did notice a difference between the relevant sounds in the two
stimuli but, since they were both deemed as “the same word”, they chose the option
SAME. Recall that the instructions specified that they were to listen to pairs of words in
a foreign language and that they had to indicate whether the pair they heard consisted
of two examples of the same word or whether they were different words. This is prob-
lematic, because, although the intent of the instructions was to induce the participants to
abstract away from their native concept of “word” by thinking of the stimuli as “foreign”
words, the participants might have still been, consciously or unconsciously, disregar-
ding perceivable sub-phonemic differences because the stimulus pairs would have been
both categorized as “the same word” in Spanish. For example, they might have heard
[iʃi]-[iʒi] and noticed a difference, but since both of those would be considered the same
word in Buenos Aires Spanish, they might have pressed SAME. This presents a possible confound in the experimental design.

It is very challenging to fashion instructions for a categorial AX task for participants without linguistic knowledge and who might not be consciously aware of the fact that physically different exemplars of a particular sound can be associated as same in a phonological category. The best way to explain this in a clear and concise manner seemed to be to ask the participants to determine whether the two foreign words that they heard were the same or different. The reason why the term “word” was used is that it evokes the sense that a word can be produced many different ways and is still the same word. The term “letter” could not be used because it would be confounded by orthography, since there is no way to reliably distinguish between voiced and voiceless fricatives in Spanish orthography. The word “sound” was not used, lest it condition the participants to pay more attention to the purely acoustic differences and disregard any phonetic differences based on socially differentiated distributions of the fricatives in their native speech community.

The reason an AX task was chosen was that it reportedly gives the listeners a better chance at discriminating between the stimuli, because there is less of a cognitive demand compared to other experimental paradigms (Pisoni 1973, Polka 1991, Werker and Logan 1985, Strange and Shafer 2008). The words “foreign word” were used to induce the listeners to abstract away from their native phonemic categories and the AX design, to allow them to have access to phonetic cues, in this case voicing, that differentiate subphonemic social variants in the native speech community. However, the commonly held assumptions that an AX task poses less cognitive demands on the listeners and that it allows them to have easier access to acoustic detail are also somewhat problematic. These generalizations were reached based on studies of monosyllabic stimuli; however,
results from Davidson (2011) suggest that the facilitation in discrimination evidenced in AX tasks with shorter ISI might not hold for longer, word-like stimuli. Most of the trials used in the current experiment were in fact disyllabic. Moreover, Gerrits and Schouten (2004:364) argue that in AX tasks listeners may be biased towards answering “same”, and will only answer “different” if they are very confident of their choice. These factors combined with the fact that the listeners might have been using their native language specific concept of “word”, thereby disregarding possibly relevant sociophonetic cues, call into question the results of this experiment. In order to circumvent this problem, an ABX task was designed to corroborate these findings.

3.6 Perception experiment 2: ABX discrimination

An ABX discrimination experiment was conducted to confirm the results obtained in the AX experiment. In an ABX task (or AXB or XAB) the participants actually listen to three stimuli (A, B, and X), two of which are explicitly from different categories (A vs. B). Then the listeners need to determine the category to which the X stimulus belongs, either to A or to B. Similar paradigms that have been used for cross-language speech perception are AXB (Best et al. 1988, 2001, Best and Strange 1992) in which both comparisons (A to X and X to B) are equidistant in time; and also XAB (Escudero 2007), which Escudero argues to be more ecologically sound, since as a listener one hears something novel and has to assign it to one of the two following categories. For this study the ABX paradigm was chosen, because, although the A and B stimuli are not equidistant to the X stimulus, the ABX configuration allows for the calculation of reaction times straightforwardly, from the moment the X stimulus is presented to the
participants.\(^8\) Using ABX, the reaction time would only calculate the time it takes the listener to make a determination, whereas in AXB or XAB from the moment of presentation of the X stimulus the listeners still have to process the phonological information of the other stimuli.

### 3.6.1 Stimulus materials

The same Brazilian Portuguese stimuli from the AX discrimination task were used for the ABX experiment. The same six obstruent voicing pairs were used, three stop pairs (p/b, t/d, k/g) and three fricative pairs (f/v, s/z, f/s). The stimuli also included obstruents in initial position (e.g. sa/za, kili/gili) as well as in intervocalic position (e.g. asa/aza, liki/ligi). However, only two flanking vowels were used: /a/ and /i/ (e.g. asa/aza, isi/izi).

### 3.6.2 Participants

38 speakers of Buenos Aires Spanish participated in this experiment, shown in Table 3.11. They were all native speakers of Buenos Aires Spanish, i.e. they were all born and raised in the Buenos Aires metropolitan area. 4 participants (Subject #: 7, 8, 11, 38) were excluded for either having previous knowledge of Portuguese or for having near-native knowledge of English. Of the remaining 34 speakers, most of them had studied other languages (mostly English, French or Italian), but none had studied Portuguese before. Among the remaining participants, there were 18 males and 16 females; 21 of the participants were from northern Buenos Aires and 13 were from southern Buenos Aires. Their ages ranged from 20-67, with a mean age of 43. It should be noted that 16

\(^8\)The presentation method used for the AX experiment (Praat MFC Experiment) did not allow for the recording of reaction times.
of these participants (Subject #: 5, 6, 9, 14, 15, 16, 17, 19, 21, 23, 26, 28, 30, 31, 33, 34) are those whose production data was analyzed in Chapter 2.

3.6.3 Experimental task

The stimuli were presented in an ABX discrimination task using the program DMDX (Forster and Forster 2003). This presentation software was used instead of Praat MFC Experiment because, in addition to being free of cost and PC compatible, it allows for the collection of reaction time data. Participants went through a practice session with stimuli that presented phonemic native contrasts that were not part of the experimental phonemes (i.e. nasal place contrasts). For the practice session the participants were told that they would be listening to trios of nonce words (“palabras inventadas”) and that the first two words were always going to be different from each other. They had to choose whether the last word in the trio would be another example of either the first or the second word by pressing the shift keys, which were marked “ABA” or “ABB”. It was stressed that these would be the only available options and that they had to pay close attention to the stimuli, since they would only be able to hear them once. Also, they were told that they should answer as quickly as possible, since after 4 seconds had elapsed the experiment would continue on to the next trio. Contrary to the AX experiment, no feedback on the correct answer was provided during the presentation of the practice trials. After the practice session, however, their attention was called to the fact that the intonation differences were not important and that the crucial difference between the words was in the sound of the first consonant of the word, or in the sound of the consonant in the middle of the word, depending on the block. They were told that the experimental sessions would follow the same format. The instructions for the experimental sessions were identical to the instructions in the practice sessions, except
<table>
<thead>
<tr>
<th>Subject #</th>
<th>Sex</th>
<th>Age</th>
<th>Neighborhood</th>
<th>Zone/Social Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>37</td>
<td>Nordelta</td>
<td>North/Upper</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>38</td>
<td>Nordelta</td>
<td>North/Upper</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>39</td>
<td>Once</td>
<td>South/Middle</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>29</td>
<td>Belgrano</td>
<td>North/Upper</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>20</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>27</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>35</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>38</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>23</td>
<td>Recoleta</td>
<td>North/Upper</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>32</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>72</td>
<td>Acassuso</td>
<td>North/Upper</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>45</td>
<td>Olivos</td>
<td>North/Upper</td>
</tr>
<tr>
<td>13</td>
<td>F</td>
<td>32</td>
<td>Burzaco</td>
<td>South/Middle</td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>29</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>63</td>
<td>Burzaco</td>
<td>South/Middle</td>
</tr>
<tr>
<td>16</td>
<td>F</td>
<td>66</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>17</td>
<td>M</td>
<td>27</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>18</td>
<td>F</td>
<td>33</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>19</td>
<td>M</td>
<td>24</td>
<td>Florida</td>
<td>North/Upper</td>
</tr>
<tr>
<td>20</td>
<td>M</td>
<td>25</td>
<td>Tigre</td>
<td>North/Upper</td>
</tr>
<tr>
<td>21</td>
<td>F</td>
<td>23</td>
<td>Recoleta</td>
<td>North/Upper</td>
</tr>
<tr>
<td>22</td>
<td>F</td>
<td>63</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>23</td>
<td>M</td>
<td>67</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>24</td>
<td>M</td>
<td>65</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>25</td>
<td>M</td>
<td>36</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>26</td>
<td>M</td>
<td>60</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>27</td>
<td>M</td>
<td>35</td>
<td>Alejandro Korn</td>
<td>South/Middle</td>
</tr>
<tr>
<td>28</td>
<td>M</td>
<td>60</td>
<td>San Telmo</td>
<td>South/Middle</td>
</tr>
<tr>
<td>29</td>
<td>M</td>
<td>32</td>
<td>Caballito</td>
<td>South/Middle</td>
</tr>
<tr>
<td>30</td>
<td>F</td>
<td>65</td>
<td>Florida</td>
<td>North/Upper</td>
</tr>
<tr>
<td>31</td>
<td>M</td>
<td>24</td>
<td>Olivos</td>
<td>North/Upper</td>
</tr>
<tr>
<td>32</td>
<td>F</td>
<td>54</td>
<td>Belgrano</td>
<td>North/Upper</td>
</tr>
<tr>
<td>33</td>
<td>F</td>
<td>58</td>
<td>Martínez</td>
<td>North/Upper</td>
</tr>
<tr>
<td>34</td>
<td>M</td>
<td>63</td>
<td>Martínez</td>
<td>North/Upper</td>
</tr>
<tr>
<td>35</td>
<td>F</td>
<td>51</td>
<td>Béccar</td>
<td>North/Upper</td>
</tr>
<tr>
<td>36</td>
<td>F</td>
<td>66</td>
<td>Martínez</td>
<td>North/Upper</td>
</tr>
<tr>
<td>37</td>
<td>M</td>
<td>66</td>
<td>Martínez</td>
<td>North/Upper</td>
</tr>
<tr>
<td>38</td>
<td>M</td>
<td>25</td>
<td>Martínez</td>
<td>North/Upper</td>
</tr>
</tbody>
</table>

Table 3.11: Demographic information of the participants in the ABX discrimination experiment.
that it was stressed that they would be listening to words in a foreign language. For a detailed transcription of the instructions for the ABX discrimination experiment see Appendix D.

As in the AX task, the X token and its matching token were always physically different exemplars of the same item; therefore, no two stimuli in a trial were ever identical.

The experiment was presented in two blocks, one with the stimulus obstruents in intervocalic position and another in phrase initial position. The order in which the blocks appeared was counterbalanced. The trials were counterbalanced in terms of the presentation order of the stimuli, i.e. there were an equal number of ABA, ABB, BAA, and BAB trials. The trials within the blocks were randomized. After randomization, care was taken to ensure that no adjacent trials were testing the same voicing pair. The inter-stimulus interval was kept at 500ms in order to compare the results to those of the AX task. There were a total of 192 experimental trials per subject. The experiment was run on a Dell XPS M1330 laptop computer and the stimuli were played over circumaural Sennheiser HD 202 headphones in a quiet room.

3.6.4 Variables and predictions

3.6.4.1 Position

The two possibilities for the phonological position of the obstruent in this experiment were initial (e.g. pa/ba) and intervocalic (e.g. apa/aba). One of the goals of the ABX experiment is to confirm the results obtained in the AX experiment; therefore, the hypotheses for the variable of Position are the same as described in Section 3.5.4.1 above. Mainly, SLM would predict significantly lower accuracy rates for VCV stops than #CV stops, whereas PAM would predict similarly high accuracy rates for stops across positions.
3.6.4.2 Voicing pair

The six voicing pairs under investigation remain the same, three stops and three fricatives: p/b, t/d, k/q vs. f/v, s/z, /ʃ/. Recall the reports of the participants of the AX task, who claimed that they classified as SAME two stimuli that they heard as different because they would belong to the same word in Spanish. Given the ABX task, where the phonetic differences of the voicing pair contrast are modeled in each trial, if the listeners indeed notice the sub-phonemic difference in the voicing pairs, then one would expect different accuracy rates for the two sibilant fricatives, s/z, /ʃ/. The possible discrimination accuracy scales are shown below (since the accuracy rates of f/v were at ceiling level in the AX task, the accuracy scales assume it as a TC assimilation pattern).

\[ \begin{array}{c|c|c|c} \text{p/b, t/d, k/q} & \text{f/v} & \text{s/z} & \text{/ʃ} \\ \hline \text{TC} & \text{TC} & \text{SC} & \text{SC} \end{array} \]

Table 3.12: Discrimination accuracy scale (with s/z as SC).

\[ \begin{array}{c|c|c|c} \text{p/b, t/d, k/q} & \text{f/v} & \text{s/z} & \text{/ʃ} \\ \hline \text{TC} & \text{TC} & \text{CG} & \text{SC} \end{array} \]

Table 3.13: Discrimination accuracy scale (with s/z as CG).

\[ \begin{array}{c|c|c|c} \text{p/b, t/d, k/q} & \text{f/v} & \text{/ʃ} & \text{s/z} \\ \hline \text{TC} & \text{TC} & \text{Social Category} & \text{CG/SC} \end{array} \]

Table 3.14: Discrimination accuracy scale incorporating Social Variant Assimilation Hypothesis.

Table 3.12 shows the accuracy scale corresponding to whether the listeners truly were not able to discern the phonetic differences in voicing in either sibilant voicing pair, and thus would mirror the AX results. Table 3.13 shows the accuracy rates if the listeners perceive the BP [z] as a deviant member of the BAS /s/ category, but perceive BP [/ʃ] and [ʃ] as equally good exemplars of the variable BAS /ʃ/. Table 3.14 shows the
accuracy rates assuming Social Variant Assimilation Hypothesis, under which BP [ʃ] and [ʒ], due to their differing social distributions in BAS, are assimilated to two different social categories; therefore, listeners are more accurate at distinguishing between ʃ/ʒ than s/z. No evidence to support this hypothesis was found with the AX task; however, this ABX task might be more sensitive in terms of detecting the differences that some listeners claimed to have perceived but later disregarded.

3.6.4.3 Social class

The null effect of social class in the AX experiment might have been due to the fact that subjects were dismissing their perceived differences in the obstruents as “the same word”. Again, ABX task might be better at uncovering any differences that one social group might have in the perception of the postalveolar voicing contrast. This would be expected particularly for the upper class speakers, since that is the group that has been previously described as having consciousness of the devoicing change (Wolf 1984, Colantoni 2008).

3.6.4.4 Age

An effect of age might have also been masked by the AX experimental paradigm; therefore, it was also included as a variable to investigate. If there is any correction from above in the devoicing change, one would expect the older speakers to be more attuned to the voicing difference and, thus, there should be a positive correlation between age and accuracy rates.
3.6.4.5 Voicers vs. devoicers

As noted in Section 3.6.2, 16 of the 34 listeners in the ABX experiments were the subjects whose production data was analyzed in Chapter 2 (the subject numbers for the participants in the production tasks and ABX correspond to the same people). Recall that 9/16 speakers were “voicers” of /z/ and 7/16 were “devoicers” of /ʒ/. The particular listener’s production status for /z/ was included as a factor in order to investigate whether the actual production of a voiced fricative for /ʒ/ might facilitate the discrimination of the /ʃ/ voicing contrast. If this is so, one would expect the voicers to exhibit higher accuracy rates than the devoicers.

3.6.5 Results and discussion

A mixed-design ANOVA was run with arcsine transform of the mean proportion correct of trials as the dependent measure, with voicing pair and position as within-subject factors and class as the between-subjects factor. The results show a significant main effect of voicing pair, $F(5, 153) = 58.642$, $p < .001$, a significant main effect of position, $F(1, 47) = 7.734$, $p = .009$ (#CV: $M = .883$, $SD = .154$; VCV: $M = .851$, $SD = .173$), and a non-significant effect of social class, $F(1, 30) = 1.081$, $p = .307$, (upper: $M = .875$, $SD = .162$; middle: $M = .854$, $SD = .168$). The only significant interaction was voicing pair*position, $F(5, 155) = 4.421$, $p = .001$. A Tukey HSD post-hoc test was performed in order to determine which voicing pair differed from each other in terms of listeners’ accuracy in discrimination. The results of the post hoc tests reveal a different pattern in the differences between the voicing pairs as was shown for the AX experiment, in which there were 2 significantly different groups: (p/b, t/d, k/ŋ, f/v) vs. (s/z, ʃ/z).
3.6.5.1 Position and voicing pair

The ABX results show that, overall, listeners exhibit significantly lower accuracy rates for obstruent voicing contrasts in intervocalic position (85% accuracy) versus in initial position (88%). However, upon closer examination, it seems that this effect of position is not present in all voicing pairs (see Figure 3.6).

A Tukey HSD post-hoc test revealed that only the sibilant voicing pairs s/z and f/ʃ exhibit a significant difference in accuracy rates (for both p<.001) when comparing the results from intervocalic and phrase initial position within each voicing pair. The accuracy rates for all the stop voicing pairs (as well as for the labiodental voicing pair f/v) are not significantly different in intervocalic and initial position (for all p≥.480).

The fact that there is no difference between listeners’ accuracy rates discriminating stops, regardless of position, appears to support PAM’s predictions. This does seem to be a case of Two Category Assimilation, since the listeners seem to be assimilating...
the voiced stops in a non-native intervocalic position to their native voiced stop category with little problems, disregarding the differences in production exhibited in the BP stimuli. This suggests that the higher level phonological structure might have more weight in terms of discrimination ability than factors such as position specificity.

If there is a case for position specificity, then it would apply to voicing pairs that are not in a Two Category assimilation relationship, as in this case s/z and j/ʃ. In the case of these sibilant voicing pairs, listeners were significantly more accurate discriminating the voicing contrast in initial position than in intervocalic position (s/z: 81% vs. 74%, j/ʃ: 72% vs. 66%). Recall that in production, BAS sibilants exhibited significantly more voicing in intervocalic position than in phrase initial position, in fact as seen in Figure 2.16, the voicing rates for both /#s/ and /#ʃ/ were very low. It may be the case that the intervocalic BP [z] and [ʃ] are actually more similar to the BAS /s/ and /ʃ/ in VCV position, thereby making the contrast harder to distinguish. It seems that these position-specific phonetic differences come to have an effect only when the two elements in the non-native contrast map onto the same native category, which would explain why this position effect is absent for the stops and f/v: these are TC assimilations.

The post hoc comparisons of the significant main effect of voicing pair show results that are different from those obtained in the AX experiment.

As can be seen from Figure 3.7, all the stop voicing pairs show high accuracy rates, at ceiling levels, (all at 90% accuracy or above). The results are very similar to those of the AX task, save for the minor difference in the p/b contrast. According to the Tukey HSD post hoc test listeners had high accuracy rates distinguishing the p/b contrast (90%); however, this was significantly different from all other voicing pairs, including stops, all with p<.01. This result is unexpected, particularly since neither PAM nor SLM would predict any place of articulation asymmetries in accuracy rates for the stop voicing pairs.
However, given that the accuracy rate for p/b contrast is still close to ceiling level, this significant difference should not be given much weight. Furthermore, with the exception of a few outliers, the distribution of the p/b trials, as seen in Figure 3.8, is quite similar to the distribution for the other stop voicing contrasts t/d and k/g (as well as the TC assimilation voicing contrast f/v). For this reason, it seems sound to consider p/b along the same lines as the other two stop voicing contrasts t/d, k/g.

As expected, the labiodental fricative voicing pair f/v is also perceived by listeners with very high accuracy. In fact, listeners’ accuracy discriminating f/v (94%) was not significantly different from the stops (except for p/b, for the reasons discussed above). This is further evidence in support of f/v being a case of Two Category Assimilation.

Given the reports of the participants after the AX task, that they were disregarding noticeable differences since they belonged to the same “word”, it would be expected that
Figure 3.8: Boxplots showing the distribution of the results of the ABX experiment, with Mean % correct scores by voicing pair (a) and the arcsine transform scores by voicing pair (b). (“x/j” represents the f/ʃ contrast.)

The listeners in the ABX task exhibit increased accuracy for the voicing contrasts where they were not already at ceiling levels. This is indeed the case, since participants show much higher discrimination rates for s/z (78%) and f/ʃ (69%) as compared to those in the AX task s/z (43%) and f/ʃ (36%). These results seem to go against the traditional description of AX as the task that is most conducive for allowing listeners to detecting a difference. Davidson and Shaw (submitted) find a similar increase in accuracy from AX to ABX in the discrimination of non-native Slavic consonant clusters by English listeners. This suggests that assigning labels to the two different AB stimuli actually facilitates the identification of X, since the listeners are forced to create and refine labels from tokens that they already know are different, thus increasing accuracy. Moreover, as Gerrits and Schouten (2004) argue, and as supported by Davidson and Shaw (submitted), listeners tend to be more conservative in that they must reach a higher threshold before they feel confident to answer “different” in the AX task. This conservatism might have been exacerbated in the current AX task, since the “same” stimulus pairs were always physically different tokens of the same category.
In the AX experiment there was no significant difference in the accuracy rates between the two sibilant voicing pairs s/z and f/\(\mathfrak{z}\); in the ABX results this is no longer the case. The s/z and f/\(\mathfrak{z}\) voicing pairs still exhibit significantly lower accuracy rates than the other voicing pairs; however, there is a significant difference between them: s/z was perceived with 78% accuracy and f/\(\mathfrak{z}\) with 69% accuracy. The accuracy scale that results from the ABX data is the one that was presented in Table 3.13, repeated as Table 3.15.

\[
p/b, t/d, k/g = f/v > s/z > f/\mathfrak{z}
\]

TC TC CG SC

Table 3.15: Discrimination accuracy scale from the ABX experiment

There does not seem to be any evidence for a Social Variant Assimilation Hypothesis, since the postalveolar voicing pair is not discriminated more accurately than s/z, despite it having a differentiated social distribution. The accuracy scale is much like the one that would be predicted by the Perceptual Assimilation Model. The stop voicing pairs are a Two Category Assimilation, whereby each element of the pair assimilates to the corresponding category in Buenos Aires Spanish. The f/v voice pair also seems to be a case of Two Category Assimilation; however, the two phones in the pair assimilate to non-corresponding categories: BP [f] maps onto BAS /f/ and BP [v] maps onto BAS /b/. The status of s/z as Category Goodness assimilation is supported by the intermediate accuracy rates that the listeners exhibited while discriminating this contrast. And finally, listeners’ accuracy rates for f/\(\mathfrak{z}\) seem to confirm its status as a Same Category Assimilation type, since it is discriminated with the lowest levels of accuracy.

3.6.5.2 Social class and age

Similar to the results in the AX experiment, social class was not a significant factor. The difference in accuracy of discrimination of f/\(\mathfrak{z}\) for the upper class listeners, those
who Wolf (1984) described as having consciousness and even negative evaluations of the devoicing of /ʒ/, showed slightly higher accuracy levels (M=.718, SD=.204) than middle class listeners (M=.669, SD=.169). However, this difference was not significant, F(1, 120)=0.949, p=.332.

Similarly to the AX results, no significant correlation was found between the listeners’ accuracy (arcsine transform of percent accuracy) and their age, neither for the discrimination of all of the voicing pairs, r(1608)=−.0007, p=.978, nor for the discrimination of /ʃ/, r(268)=.029, p=.637.

3.6.5.3 Voicers vs. devoicers

It was expected that if (de)voicer status does have an effect on the discrimination of the /ʃ/ voicing contrast, then the voicers should exhibit higher accuracy rates than the devoicers. A mixed-design ANOVA was run on the /ʃ/ voicing pairs only, with arcsine transform of the mean proportion correct of trials as the dependent measure, with position as within-subjects factor and (de)voicer status as the between-subjects factor. The results show a significant main effect of position, F(1, 13)=7.212, p=.019 (#CV: M=.761, SD=.165; VCV: M=.652, SD=.162), which is expected and follows the pattern of the general results. However, there is no significant main effect of (de)voicer status, F(1,12)=0.054, p=.821, or of its interaction with position. Despite these listeners’ differences in their native fricative inventory (i.e. whether they have /ʃ/ or /ʒ/), this does not seem to affect their accuracy in discrimination of the BP voicing pair /ʃ/ (voicers: M=.714, SD=.171; devoicers: M=.701, SD=.173).
3.6.6 Reaction times

The results from subjects’ accuracy rates in both the AX and ABX experiments showed no significant differences in position for any of the stop voicing pairs. However, the perception of all stop contrasts was at ceiling level in both positions. This ceiling effect might be obscuring more subtle difficulties in the discrimination of occluded stop voicing contrasts in intervocalic position. The reaction time results for the ABX task were analyzed in order to uncover any of these differences. SLM would predict that the voicing contrast in intervocalic should show longer reaction times than the contrast in initial position.

Figure 3.9: Reaction time results for each voicing pair in initial and intervocalic position in the ABX experiment. Asterisks mark the significant differences in position within the voicing pair. Significant differences between voicing pairs are summarized as follows: No differences between any of the stop voicing pairs. All fricative voicing pairs are different from each other. f/v is different from all voicing pairs except k/g.

A two-way ANOVA was run with reaction time as the dependent measure and voicing pair and position as the independent factors. The results showed a significant main effect of voicing pair, F(5,6318)=34.260, p<.001, a significant main effect of position,
F(1,6318)=83.440, p<.001, and a significant interaction, F(5,6318)=6.252, p<.001. The reaction time results are shown in Figure 3.9.

Tukey HSD post hoc tests of the main effect of voicing pair revealed that there are no differences between any of the stop voicing pairs. This is congruent with the fact that they are treated in the same manner, as Two Category assimilation. All the fricative voicing pairs were different from each other. f/v had the fastest reaction times and was significantly different from every other voicing pair, except for k/g. Listeners might have been faster at discriminating this contrast, perhaps because of the fact that the two categories that the member of the f/v voicing pair maps onto are /f/ and /b/, respectively. This assimilation crosses not just the voicing contrast boundary (as do p/b, t/d, and k/q), but also a manner of articulation contrast, since the [f] maps onto the fricative /f/ but the [v] maps onto the stop /b/. This might decrease their confusability, thereby allowing the listeners to discriminate faster. The difference between s/z and ñ/ñ reflects the accuracy results in that the reaction times for ñ/ñ are the slowest, significantly slower than for s/z. This is further evidence showing that ñ/ñ is a SC assimilation; if both [ñ] and [ñ] are equally acceptable members of the postalveolar category, then they will be hardest to tell apart.

Tukey HSD post-hoc tests investigating the interaction between voicing pair and position revealed that although there is a significant main effect of position, the reaction times between CV and VCV position are not significant for any of the stop voicing pairs, only for the fricatives. Participants were significantly slower with fricative voicing contrasts in intervocalic position compared to initial position. It is not clear why they should be significantly slower discriminating the f/v in intervocalic position, particularly since f/v has patterned with the stops in most analyses thus far. For the sibilant voicing pairs, however, the longer reaction times might be related to the fact that in intervocalic posi-
tion, voiceless fricatives are more likely to exhibit higher voicing rates, which might add to the confusability. The difference between a voiced and a voiceless fricative appears to be more salient in #CV position and, therefore, more easily distinguishable. More work is needed to determine whether /f/, a phonologically voiceless fricative, also exhibits position effects in production as do /s/ and /z/.

3.7 Conclusions

This chapter has investigated the effects of phonological, phonetic and sociolinguistic factors on the perception of perception non-native voicing contrasts by Buenos Aires Spanish listeners. Of particular interest was the testing the differing predictions made by the Perceptual Assimilation Model and the Speech Learning Model on the perception of native contrasts in positions where they do not occur natively. The results from both AX and ABX, as well as reaction time, show that BAS listeners can distinguish a canonical stop voicing contrast in BP, regardless of the fact that this contrast does not occur natively in intervocalic position. These results support the predictions of the Perceptual Assimilation Model and suggest that the ability of a listener to discriminate a non-native contrast might be most affected by the relationship of the non-native phonemes to the native phonological categories existent in the phoneme inventory.

Although BAS listeners had little difficulty distinguishing the non-native stop voicing contrast in BP, they had lower accuracy rates with the voicing contrast in sibilants. The ABX paradigm helped reveal differences in accuracy that the AX task did not, thus providing more evidence for the proposal that listeners might actually be more sensitive to acoustic differences in a task where they are forced to label two different stimuli. The results from the ABX experiment helped elucidate the possible effect of sociolinguistic variation on the perception of non-native contrasts, by revealing a difference in accu-
racy between s/z and \( f/z \). Despite the fact that \( f/z \) shows socially stratified variation in the speech community, this did not facilitate the discrimination of \( f/z \), it actually hindered the perception of the postalveolar voicing pair as compared to s/z. Listeners do not seem to tap into these stratified distributions to aid them in these perception tasks. On the contrary, the effect that native voicing variation seems to have is that of widening the acceptability range of the native category. This wider range of acceptability within the same native \( f/z \) category seems to increase the confusability of discrimination of \([f]-[z]\) in a non-native language.

The assimilation patterns that have been discussed in this chapter are based on relative accuracy rates in the discrimination experiments. However, in order to confirm whether these Two Category, Category Goodness, and Same Category assimilation patterns accurately reflect listeners’ classification, a separate transcription and rating tasks need to be performed. This will be further explored in the following chapter.
Perceptual assimilation of non-native obstruent voicing contrasts

4.1 Introduction

The previous chapter examined the effects of native allophonic and sociophonetic variation on the discrimination of non-native obstruent voicing contrasts by listeners who are native speakers of Buenos Aires Spanish (BAS). This chapter will analyze the data from their transcription and rating tasks in order to shed light on the inferences that were made about the assimilation patterns based on the accuracy rates in the discrimination tasks. The results from the experiments in this chapter suggest that looking at the accuracy results and even at the transcription data is not enough to determine the nature of the assimilation patterns. A more detailed analysis of the goodness ratings for the non-native sounds are crucial in order to distinguish between the assimilation patterns proposed by Best’s Perceptual Assimilation Model.
4.2 Background

In order to corroborate the assimilation patterns that they predicted on the basis of articulatory similarities, Best et al. (2001) further analyzed the assimilation patterns of native English listeners for the non-native Zulu contrasts (/ʌ/-/ʊ/, /b/-/ɓ/, and /kʰ/-/kʼ/) in their investigation by administering questionnaires after the discrimination tests. The participants were asked to transliterate, using English spelling, the experimental stimuli, which were replayed to the participants. The questionnaire instructed the participants to spell the syllables that they heard as they would in English if and only if they sounded like English consonants and, if they so chose, they could write any comments or supplementary description on the sounds (Best et al. 2001:782). In their study, the type of assimilation was deemed Same Category if the participants spelled both consonant contrasts identically or with phonologically equivalent transliterations. If the contrast was spelled with a common letter, but one of the contrasting elements had an additional letter to qualify it or if they commented on how it sounded differently, then that was considered a Category Goodness assimilation type. If the contrasting pairs were spelled with two different letters representing a phonological contrast in English, then that was considered a Two Category Assimilation type.

This assimilation pattern rubric, although it provides clear directives to assess the assimilation types, is somewhat problematic. This method assumes that by transliterating something with the same orthographic symbol, it is necessarily implying that both sounds are perceived as equally good exemplars of the native phonological category. However, is a change in orthography necessarily the threshold between an acceptable example of a category and a noticeable worse fit for it? It could be the case that there is no evident way to change the orthography or describe the difference. The additional letter or the supplementary commentary might be a clear sign of the relatively more de-
graded status of that particular stimulus to the native phonological category represented by the sole orthographic symbol; however, it does not clarify the boundary between Category Goodness and Two Category assimilation, since the “addition” of a segment might represent a categorical difference as well.

Other studies investigating the assimilation and identification of non-native stimuli (Strange et al. 2004, 2005) provide participants with a series of predetermined responses to choose from and then a Likert scale in order for them to rate how good an instance of that labeled sound it was. This approach is very useful, since it allows for the quantification of the goodness ratings. One drawback might be that the set of responses is predetermined and it would be complicated for listeners to learn the categories if their native orthography has no conventionalized way to encode these differences. A combined approach, with open ended responses and a rating scale would not only offer more flexibility in the investigation of the assimilation patterns but also provide a way to measure the level of “Category Goodness”, so to speak.

4.3 Methodology

4.3.1 Participants

The 38 participants who completed this task were the same speakers that participated in the ABX experiment (a subset of which provided the data that is presented in the production chapter). They were all native Spanish, born and raised in the Buenos Aires metropolitan area. The same 4 participants who were excluded in the ABX task were also excluded for this analysis, on account of having previous knowledge of Portuguese (subjects 7 and 38) or for having near-native knowledge of English (subjects 8 and 11). Of the remaining subjects, most had studied other languages to varying degrees (mostly
English, French or Italian), but none had previous exposure to Portuguese. There were 18 males and 16 females, 21 of the participants were from northern Buenos Aires and 13 were from southern Buenos Aires. Their ages ranged from 20-67, with a mean age of 43. The perceptual assimilation and rating tasks immediately followed the ABX task; therefore, the participants are the same in both tasks. As a reference, their demographic information in repeated in Table 4.1 (which is the same as Table 3.11).

4.3.2 Stimuli

The Brazilian Portuguese stimuli used for the present transcription and rating task was a subset of the stimuli used in the ABX experiment. The relevant obstruents were the same as in the previous experiments: 6 stops /p, b, t, d, k, g/ and 6 fricatives /f, v, s, z, ʃ, ʒ/. The obstruents were presented in the following to templates: /Cala/ and /Cili/ for initial position and /laCa/ or /liCi/ for intervocalic position. These two types of templates were chosen so that the intervocalic and initial tokens would both appear in similar contexts in terms of length and syllabic structure. The stimuli were presented in two different blocks, either intervocalic or initial position, and these blocks were counterbalanced.

4.3.3 Procedure

After the participants had completed the ABX discrimination experiment, they were presented with a transcription and rating questionnaire (see Appendix E). The instructions stated that they would be listening to some of the words that had appeared in the experiment and that they should transcribe each word as they would write it in Spanish. All of the letters corresponding to the vowels and the /l/ in the template were already provided, so that they would only have to focus on the obstruent in question. An open-
<table>
<thead>
<tr>
<th>Subject #</th>
<th>Sex</th>
<th>Age</th>
<th>Neighborhood</th>
<th>Zone/Social Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>37</td>
<td>Nordelta</td>
<td>North/Upper</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
<td>38</td>
<td>Nordelta</td>
<td>North/Upper</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>39</td>
<td>Once</td>
<td>South/Middle</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>29</td>
<td>Belgrano</td>
<td>North/Upper</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>20</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>6</td>
<td>M</td>
<td>27</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>35</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>8</td>
<td>F</td>
<td>38</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>23</td>
<td>Recoleta</td>
<td>North/Upper</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>32</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>11</td>
<td>F</td>
<td>72</td>
<td>Acassuso</td>
<td>North/Upper</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>45</td>
<td>Olivos</td>
<td>North/Upper</td>
</tr>
<tr>
<td>13</td>
<td>F</td>
<td>32</td>
<td>Burzaco</td>
<td>South/Middle</td>
</tr>
<tr>
<td>14</td>
<td>F</td>
<td>29</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>15</td>
<td>F</td>
<td>63</td>
<td>Burzaco</td>
<td>South/Middle</td>
</tr>
<tr>
<td>16</td>
<td>F</td>
<td>66</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>17</td>
<td>M</td>
<td>27</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>18</td>
<td>F</td>
<td>33</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>19</td>
<td>M</td>
<td>24</td>
<td>Florida</td>
<td>North/Upper</td>
</tr>
<tr>
<td>20</td>
<td>M</td>
<td>25</td>
<td>Tigre</td>
<td>North/Upper</td>
</tr>
<tr>
<td>21</td>
<td>F</td>
<td>23</td>
<td>Recoleta</td>
<td>North/Upper</td>
</tr>
<tr>
<td>22</td>
<td>F</td>
<td>63</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>23</td>
<td>M</td>
<td>67</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>24</td>
<td>M</td>
<td>65</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>25</td>
<td>M</td>
<td>36</td>
<td>Palermo</td>
<td>North/Upper</td>
</tr>
<tr>
<td>26</td>
<td>M</td>
<td>60</td>
<td>Adrogué</td>
<td>South/Middle</td>
</tr>
<tr>
<td>27</td>
<td>M</td>
<td>35</td>
<td>Alejandro Korn</td>
<td>South/Middle</td>
</tr>
<tr>
<td>28</td>
<td>M</td>
<td>60</td>
<td>San Telmo</td>
<td>South/Middle</td>
</tr>
<tr>
<td>29</td>
<td>M</td>
<td>32</td>
<td>Caballito</td>
<td>South/Middle</td>
</tr>
<tr>
<td>30</td>
<td>F</td>
<td>65</td>
<td>Florida</td>
<td>North/Upper</td>
</tr>
<tr>
<td>31</td>
<td>M</td>
<td>24</td>
<td>Olivos</td>
<td>North/Upper</td>
</tr>
<tr>
<td>32</td>
<td>F</td>
<td>54</td>
<td>Belgrano</td>
<td>North/Upper</td>
</tr>
<tr>
<td>33</td>
<td>F</td>
<td>58</td>
<td>Martínez</td>
<td>North/Upper</td>
</tr>
<tr>
<td>34</td>
<td>M</td>
<td>63</td>
<td>Martínez</td>
<td>North/Upper</td>
</tr>
<tr>
<td>35</td>
<td>F</td>
<td>51</td>
<td>Béccar</td>
<td>North/Upper</td>
</tr>
<tr>
<td>36</td>
<td>F</td>
<td>66</td>
<td>Martínez</td>
<td>North/Upper</td>
</tr>
<tr>
<td>37</td>
<td>M</td>
<td>66</td>
<td>Martínez</td>
<td>North/Upper</td>
</tr>
<tr>
<td>38</td>
<td>M</td>
<td>25</td>
<td>Martínez</td>
<td>North/Upper</td>
</tr>
</tbody>
</table>

Table 4.1: Demographic information of the participants in the perceptual assimilation task (same as in the ABX task).
ended transcription questionnaire was chosen, because there is no unambiguous way to transcribe the voiced fricatives in Spanish orthography.

After every word there was a space for comments where the participants could write whatever information they wanted about the particular sounds. After transcribing the obstruent and writing any comments, they were asked to look back at the letter (or letters) that they transcribed and think about how they are most commonly pronounced in Spanish and then determine whether the sound that they had heard was a good example of that most usual pronunciation. They were provided with a 5-point Likert scale and were asked to give a rating depending on how good of an example of that most usual pronunciation it was (1 = very bad example, 5 = very good example).

Since the instructions could prove confusing to people without linguistics background, two example stimuli were provided at the beginning of each block. The example phoneme was /r/ in either initial or intervocalic position depending on the block. The example nonce words were either <rele> /rele/ or <lerru> /leru/ for each position. The first example was produced as [rele] or [leru], with an alveolar trill, and it was aimed that the speaker would rate this as a 5. The second example was produced as [rele] or [leru] with a uvular trill, to which they would give a much lower rating. These examples were both produced and recorded by the investigator, a native speaker of Puerto Rican Spanish. This dorsal pronunciation, although not a phoneme of Spanish, is a variant of /r/ that speakers of Spanish are aware of, since it is considered to be a deviant variant of /r/ produced by people who, for developmental reasons, cannot produce an alveolar trill. It is also the variant associated French-accented speakers of Spanish. Although this sound is not “native” to the Spanish phonemic inventory, speakers are aware that it is not an uncommon pronunciation in disordered speech (and many of them noted this

1[^1] is also a native sociolinguistic variant of /r/ for some speakers of Puerto Rican Spanish (see Lipski 1994:333-4), although it is highly unlikely that the participants in this study ever encountered it as such.
in their comments, for example *gangoso* “nasal/twangy”). Furthermore, *<r>* or *<rr>* would be the only ways of representing [r] orthographically in Spanish as a category of /r/, thereby familiarizing them with the concept of better and worse examples of the same category. After hearing and rating the example cases, they were played the rest of the 24 experimental stimuli for each block.

The stimuli were played as an experiment file in Praat (Boersma and Weenink 2010) on a Dell XPS M1330 laptop over Sennheiser HD202 circumaural headphones in a quiet room. Once the stimulus was played, the participants had to transcribe the obstruent, give a rating, and write comments if they so chose. Once the participant signaled to the investigator that they were finished with that trial, the investigator played the subsequent one. Participants could take a break between the blocks if they wanted.

### 4.4 Predictions

Recall that in the AX and ABX experiments, the listeners’ discrimination of the stop voicing pairs was at ceiling levels, regardless of the position they occurred in. The BP stops /p, t, k, b, d, g/ in initial position should be identified as their corresponding stops in Spanish; however, in intervocalic position there could potentially be more chance for confusion, since phonologically voiced stops are not occluded natively, but rather spirantized. If we assume that listeners assimilate the occluded intervocalic stops to the corresponding native stop category, regardless of the lack of spirantization, then they should not have major problems identifying them as such. However, as was discussed in Section 3.3.5, in casual speech it is not uncommon for phonologically voiceless stops to be realized as voiced occluded stops. Given this information, it could be the case that subjects listen to cases of occluded voiced stops in BP and misidentify them as instances of phonologically voiceless fricatives, since they share many of the acoustic details.
This confusion might not have affected the accuracy in the discrimination experiments, because they always had at least one other stimulus to compare non-canonical tokens to; however, in the transcription task subjects are presented with isolated stimuli and might impair their identification of potentially non-canonical tokens.

In terms of the goodness ratings, the voiceless stop stimuli should have high ratings in initial as well as in intervocalic position, but for the voiced stop stimuli there might be some discrepancy in the ratings for initial position vs. intervocalic position. The non-spirantized voiced stops of BP will probably be given lower ratings than their corresponding voiceless stop, regardless of whether they are identified as an example of a BAS voiced stop or a voiceless stop, e.g. [lada], although it might not be infrequent in casual speech, would be a non-canonical example of both /lada/ and /lata/.

As for the fricative stimuli, BP /f/ and /s/ should be identified as their corresponding categories in Spanish (represented orthographically as <f> and <s>, respectively), since they exist as such in BAS. Given the ceiling-level accuracy results for the f/v voicing pair and the interpretation of that as a two category assimilation, the prediction for BP /v/ would be that it would be identified as Spanish /b/ (represented most commonly as <b, v>), although perhaps with a lower goodness of fit rating than for BP /f/. If the BP /s/-/z/ voicing pair is a case of Category Goodness Assimilation, then it would be predicted that BP /s/ would be identified as BAS /s/ (represented orthographically most often as <s, z>), whereas BP /z/ would also be identified as BAS /s/ but with a significantly lower goodness rating.

The prediction for the BP postalveolar fricatives /ʃ/ and /ʒ/ is that they will both be transcribed as either <y> or <ll>; however, the goodness ratings might vary depending on social factors or on whether the listener is one for whom the devoicing of /ʒ/ has been completed. According to Wolf (1984), the upper class speakers in her study were aware
of the devoicing and move away from it; therefore, one would expect that they rate BP /f/ as a worse exemplar of their native postalveolar category than BP /ʒ/. Colantoni’s (2008) findings that speakers from the upper social strata perceived variation while the speakers from the lower strata did not, also suggest that there should be a difference in the ratings by social class. The younger middle class speakers might rate BP /ʒ/ lower than BP /f/, as would be suggested by King’s (2009) findings. Another possibility is that the listeners rate the goodness of fit of the postalveolar stimuli with respect to their own production, in which case the ones for whom the devoicing of /ʒ/ has been completed might find BP /f/ as a more acceptable exemplar of their native category, while those who were “voicers” in Chapter 2 might find /ʒ/ as the better exemplar.

Recall that in the ABX experiment the only instances where the effect of position was significant were in the discrimination of the s/z and S/Z voicing pairs; therefore, one might expect effect of position in the rating of the BP voiced sibilant stimuli as well.

4.5 Results and discussion

The responses and ratings for each participant were compiled for the analysis. Figure 4.4 shows the raw rating distribution for each of the stimulus obstruents, pooled across all subjects. A comparison of the beanplots in Figure 4.4 already reveals some interesting tendencies. There seems to be a discrepancy in the ratings of the voiced stops /b, d, g/ between initial and intervocalic position. The ratings for the voiced stop stimuli seem to be lower in intervocalic position, whereas, the same effect is not apparent in the voiceless stop stimuli /p, t, k/ which appear to be more symmetrical on either side. In general, the voiced fricative stimuli show a more diffuse distribution in their goodness ratings as compared to the stops, particularly in initial position, where ratings are concentrated in the upper ranges. Although informative, this plot might prove to be
somewhat misleading, since it contains the ratings for the stimuli regardless of how the listener identified the stimulus. Inspecting the responses given for each obstruent will provide a clearer picture of the goodness ratings.

4.5.1 Stop stimuli

The overall responses for the orthographic transcription of the stop stimuli can be seen in Figure 4.2. From the inspection of these graphs it seems that in most cases there are one or more clear options that were favored by most listeners as the response, and the rest of the responses were very infrequent.

<table>
<thead>
<tr>
<th>Response</th>
<th>b</th>
<th>d</th>
<th>df</th>
<th>f</th>
<th>p</th>
<th>t</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>122</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.2: Orthographic responses for BP /p/ stimuli.

The responses for BP /p/ were overwhelmingly <p> (Table 4.2), so the other responses were eliminated from the analysis.
Figure 4.2: Distribution of the transcription responses for each of the stop stimuli collapsed across position and vowel contexts.

### Table 4.3: Orthographic responses for BP /t/ stimuli.

<table>
<thead>
<tr>
<th>Response</th>
<th>c</th>
<th>k</th>
<th>p</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>11</td>
<td>3</td>
<td>1</td>
<td>120</td>
</tr>
</tbody>
</table>

The responses for BP /t/ were overwhelmingly <t> (Table 4.3), so the other responses were eliminated from the analysis.
Table 4.4: Orthographic responses for BP /k/ stimuli.

For the BP /k/ stimuli, there is more than one predominant response, <k>, <c>, and <qu> being the most frequent (see Table 4.4). However, this diversity in responses is due to the orthographic conventions of Spanish. Recall that the /k/ in the stimuli was always followed either by an <a> or an <i>. In Spanish, the orthographic sequence <ca> can be used to represent /ka/; however, <ci> does not represent /ki/ but rather /si/. The Spanish spelling convention to represent /ki/ is <qui>. The grapheme <k> is relatively infrequent in Spanish, occurring most commonly in foreign proper names or borrowings, also in the metric compounds with kilo-; however, <k> always represents /k/, regardless of the following vowel, which might account for its high use here. For these reasons, all responses <c>, <k>, <qu>, or any combination thereof were collapsed and the remaining responses were eliminated from the analysis.

The three subsets for each voiceless stop stimulus were combined and a two-way ANOVA was run with goodness of fit ratings as the dependent variable and stimulus obstruent (/p, t, k/) and position (initial vs. intervocalic) as independent factors, with subject as a random factor. The results yielded no significant main effect of obstruent, F(2,28)=0.222, p=.802, no significant main effect of position, F(1,28)=2.326, p=.139, and no significant interaction, F(2,28)=1.043, p=0.366.

Table 4.5: Orthographic responses for BP /b/ stimuli.

Among the responses for BP /b/, <b> was the most frequent identification; however, <v> was also a considerably frequent response as well (see Table 4.5). As has been
discussed before, in Spanish orthography both <b> and <v> are representations of the phoneme /b/, so it would be a valid transcription as well. Responses other than <b> or <v> were eliminated from further analyses. Interestingly, the distribution of <b> and <v> responses is not uniform across position (see Table 4.6).

<table>
<thead>
<tr>
<th></th>
<th>b</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>57</td>
<td>7</td>
</tr>
<tr>
<td>Intervocalic</td>
<td>43</td>
<td>21</td>
</tr>
</tbody>
</table>

Table 4.6: Distribution by position of <b> and <v> responses for BP /b/.

A chi-square test of independence shows that the relation between the listener’s response (<b> vs. <v>) and the position of the stimuli is significant, \( \chi^2 (1, N=128)=7.726, p=0.005 \). Initial /b/ stimuli were less likely to be identified as <v> than as <b>, whereas intervocalic stimuli of /b/ were more likely to be identified as <v> than as <b>. Although, there seems to be a bias towards identifying intervocalic [b] as <v>, this has no effect on the ratings of the stimuli. To investigate this effect on the subjects’ ratings of BP /b/, a linear mixed effects regression model\(^2\) was conducted with response (<b> vs. <v>) and position (initial vs. intervocalic) as fixed effects, and subject as a random factor, (shown in Table 4.7 below).

The results yielded a significant main effect for position, such that the average rating for the stimuli with BP [b] was significantly higher for initial position (M=4.30, SD=0.81) than for intervocalic position (M=3.52, SD=1.10). The main effect of response was not significant and the interaction was not significant either, indicating that the effect of position holds, regardless of the subject’s response.

\(^2\)A mixed-effect regression model was conducted rather than a two-way ANOVA, because of the unbalanced distribution of the responses by subject, i.e. not every subject gave both orthographic variants in their responses. For this reason subject could not be included as a random factor in the ANOVA; however, a mixed-effect model can take subject as a random factor under these conditions.
Table 4.7: Results of the linear mixed effect regression model adjusted to the ratings of BP /b/, with response and position as fixed factors and subject as a random factor. (Intercept = <v>, intervocalic).

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>$\beta$</th>
<th>S.E.</th>
<th>$t$</th>
<th>$p_{MCMC}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>(Intercept)</td>
<td>4.367</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>response</td>
<td>-0.410</td>
<td>0.344</td>
<td>-1.194</td>
<td>0.235</td>
</tr>
<tr>
<td></td>
<td>position</td>
<td>-0.844</td>
<td>0.148</td>
<td>-5.685</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>response:position</td>
<td>0.337</td>
<td>0.382</td>
<td>0.881</td>
<td>0.380</td>
</tr>
</tbody>
</table>

It might be the case that the listeners notice a discrepancy in the non-native occluded voiced stops and they identify it with <v>, the non-canonical spelling of BAS /b/. However, the occluded voiced stops are not a better exemplar of <v> than of <b>. Given these results, it is warranted to collapse the responses of <b> and <v>; for both responses, [VbV] is rated lower than [bV].

Table 4.8: Orthographic responses for BP /d/ stimuli.

<table>
<thead>
<tr>
<th>Response</th>
<th>d</th>
<th>dd</th>
<th>td</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>132</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

The responses for BP /d/ were overwhelmingly <d> (see Table 4.8), so the other responses were eliminated from the analysis. The results of a t-test show that the goodness ratings for /d/ as <d> in intervocalic position (M=3.662, SD=1.189) are significantly lower than in initial position (M=4.224, SD=0.997), t(124)=2.939, p=0.004.

Table 4.9: Orthographic responses for BP /g/ stimuli.

<table>
<thead>
<tr>
<th>Response</th>
<th>d</th>
<th>g</th>
<th>gh</th>
<th>gu</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1</td>
<td>68</td>
<td>1</td>
<td>64</td>
<td>1</td>
</tr>
</tbody>
</table>

Most of the responses for the BP stimuli /g/ are distributed between <g> and <gu> (see Table 4.9). Similar to the case of /k/, this is due to the same spelling rule that affects the orthographic representation of /g/ depending on the following vowel. In Spanish,
the orthographic sequence <ga> can be used to represent /ga/; however, <gi> does not represent /gi/ but /xi/. The Spanish spelling convention to represent /gi/ is <gui>. For these reasons, the responses <g> and <gu> were collapsed and the remaining responses were eliminated from the analysis. The results of a t-test show that ratings for intervocalic /g/ stimuli are significantly lower (M=3.682, SD=0.994) than the ratings for the initial /g/ stimuli (M=4.106, SD=1.242), t(124)=2.166, p=0.032.

One very salient finding of this analysis of the stop stimuli is the effect of position for each of the voiced stops. Although listeners did not show any effect of position in the discrimination experiments, either in accuracy or in response times, in this goodness of fit task, participants gave significantly worse ratings for the intervocalic voiced stops than to initial voiced stops. Judging from the orthographic responses, it seems that the listeners do assimilate the occluded voiced stops to their native voiced stop category, regardless of the lack of spirantization. Furthermore, listeners do not seem to associate occluded intervocalic voiced stops with the common positional variant of voiceless stops in intervocalic position. This is precisely the point where the Perceptual Assimilation Model’s predictions would differ from the Speech Learning Model’s predictions. [ada] as a realization of Spanish /ata/ would be much more frequent in casual speech than [ada] as a realization of Spanish /ada/. Under SLM, if the obstruent in BP [ada] were to relate to a positional variant, it would be much more plausible for it to relate to /ata/, since the acoustic characteristics of it are similar to how /ata/ is usually produced. However, this does not seem to be the case. The data support PAM’s predictions in that the non-native [ada] seems to assimilate to the higher order invariants associated with Spanish /ada/, however, the stimuli were all isolated words and thus evidently not casual, connected speech, so this might have inhibited the association to the casual speech variant. Neither PAM nor SLM address how stylistic or environmental effects might affect the patterns of association of the non-native sounds; however, it has been shown that this does affect listeners’ perceptual boundaries (see Hay et al. 2006).
with the qualification that it is a worse fit to that category, at least as compared to a phrase initial tokens.

4.5.2 Fricative stimuli

The overall responses for the orthographic transcription of the fricative stimuli can be seen in Figure 4.3.

Figure 4.3: Distribution of the transcription responses for each of the fricative stimuli collapsed across position and vowel contexts.
Similar to the situation with the BP stop stimuli, it seems that in most cases there are one or two clear options that were favored by most listeners as the response transcription to the fricative stimulus, while the rest of the responses were very infrequent.

<table>
<thead>
<tr>
<th>Response</th>
<th>dz</th>
<th>f</th>
<th>s</th>
<th>sh</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1</td>
<td>130</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.10: Orthographic responses for BP /f/ stimuli.

The responses for BP /f/ were overwhelmingly <f>, so the other responses were eliminated from the analysis. The results from a t-test show that ratings for initial /f/ (M=4.250, SD=0.976) and intervocalic /f/ (M=4.303, SD=0.784) are not significantly different from each other. Listeners rate intervocalic and initial /f/ equally well.

<table>
<thead>
<tr>
<th>Response</th>
<th>b</th>
<th>b/v</th>
<th>d</th>
<th>v</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>41</td>
<td>1</td>
<td>1</td>
<td>87</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 4.11: Orthographic responses for BP /v/ stimuli.

Among the responses for BP /v/, <v> was the most frequent identification; however, <b> was also a considerably frequent response as well. Responses other than <b> or <v> were eliminated from further analyses. As discussed in Chapter 3, BP /v/ is articulatorily most similar to Spanish /f/ and could theoretically be assimilated to that category by BAS listeners, as a case of Category Goodness. However, BP /v/ was not once identified as <f> by the listeners. This, together with the fact that it was most often identified as an orthographic variant of BAS /b/, indicates that the f/v voicing pair is indeed a case of Two Category assimilation. Furthermore, it is interesting that <v> was the most common response, since supposedly /v/ is not a part of the Spanish phonological inventory. It could be the case that the subjects are recognizing this realization as the realization associated with the very careful speech and dictation style of elementary school. If this is indeed the case, then one would expect the <v> cases to show higher
goodness of fit ratings. Interestingly, as we saw for BP /b/ stimuli, the distribution of 
<\textit{v}> and <\textit{b}> responses is not uniform across position (see Table 4.12)

<table>
<thead>
<tr>
<th>Position</th>
<th>b</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>30</td>
<td>34</td>
</tr>
<tr>
<td>Intervocalic</td>
<td>11</td>
<td>53</td>
</tr>
</tbody>
</table>

Table 4.12: Distribution by position of <\textit{b}> and <\textit{v}> responses for BP /\textit{v}/.

A chi-square test of independence shows that the relation between the listeners’ response (<\textit{v}> vs. <\textit{b}>) and the position of the stimuli is significant, \(\chi^2(1, N=128)=11.627, p < .001\). Initial /\textit{v}/ stimuli were more often identified as <\textit{b}> than as <\textit{v}>, whereas intervocalic stimuli of /\textit{v}/ were more often identified as <\textit{v}> than as <\textit{b}>. Interestingly, Lustig (2011) recently conducted a study on the realization of orthographic <\textit{b}> and <\textit{v}> in reading tasks in this corpus of BAS and found that the cases of orthographic <\textit{v}> in intervocalic position tend to be realized as actual fricatives (as opposed to the canonical approximant) more frequently than the cases of <\textit{b}> in intervocalic position. Perhaps this might be influencing the bias towards responding <\textit{v}> in intervocalic cases of /\textit{v}/.\(^4\) To investigate this effect on the subjects’ ratings of BP /\textit{v}/, a linear mixed effects regression model\(^5\) was conducted with response (<\textit{b}> vs. <\textit{v}>) and position (initial vs. intervocalic) as fixed effects, and subject as a random factor, (shown in Table 4.13 below).

Despite this positional bias, there were no significant effects of response, position, and the interaction was not significant either. These results indicate that, although many listeners identified BP /\textit{v}/ as <\textit{v}>, they do not necessarily consider it to be a “better”

\(^4\)In phrase-initial cases <\textit{v}> did not favor fricative realizations when compared to <\textit{b}>, because of the overriding tendency for fortition of /\textit{b}/ to an actual occluded voiced stop, regardless of orthographic representation.

\(^5\)See footnote 2
Table 4.13: Results of the linear mixed effect regression model adjusted to the ratings of BP /v/, with response and position as fixed factors and subject as a random factor. (Intercept = <v>, intervocalic).

Table 4.14: Orthographic responses for BP /s/ stimuli.

Table 4.15: Distribution by position of <s> and <z> responses for BP /s/.

The relation between orthographic variant and position was not significant, $\chi^2(1, N=122)=0.041, p=0.840$. A linear mixed effects regression model was conducted with

---

6See footnote 2
Table 4.16: Results of the linear mixed effect regression model adjusted to the ratings of BP /s/, with response and position as fixed factors and subject as a random factor. (Intercept = <z>, intervocalic).

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>β</th>
<th>S.E.</th>
<th>t</th>
<th>pMCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>(Intercept)</td>
<td>3.778</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>response</td>
<td>-0.260</td>
<td>0.363</td>
<td>-0.714</td>
<td>0.477</td>
</tr>
<tr>
<td></td>
<td>position</td>
<td>-0.093</td>
<td>0.197</td>
<td>-0.471</td>
<td>0.638</td>
</tr>
<tr>
<td></td>
<td>response:position</td>
<td>0.493</td>
<td>0.473</td>
<td>1.042</td>
<td>0.300</td>
</tr>
</tbody>
</table>

The results did not yield any significant effects of response (i.e. <s> or <z>), or of position, and the interaction was not significant either. This suggests that by collapsing the responses <s> and <z>, no significant differences in ratings will be obscured.

Table 4.17: Orthographic responses for BP /z/ stimuli.

<table>
<thead>
<tr>
<th>Response</th>
<th>c</th>
<th>dz</th>
<th>s</th>
<th>sh</th>
<th>ss</th>
<th>ts</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>zs</th>
<th>zy</th>
<th>zz</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>2</td>
<td>4</td>
<td>49</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>64</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Similar to what we saw for /s/, the most common responses for BP /z/ were <s> and <z>; however, in this case <z> is the most frequent one. The remaining responses were eliminated from the analysis. It is interesting that the <z> response is the most frequent one, given that <z> in BAS never represents an actual phoneme /z/, since Spanish lacks it. And if <z> is realized as [z], because of the coarticulatory factors discussed in Chapter 2, presumably it should behave in the same way as orthographic <s>. In this case the distribution of response by position was also rather even between orthographic responses (see Table 4.18).

A chi-square test of independence revealed no significant relation between the listener’s responses and the position of the stimuli $\chi^2(1, N=113)=0.039, p=0.844$. A linear
Table 4.18: Distribution by position of <s> and <z> responses for BP /zl/.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>β</th>
<th>S.E.</th>
<th>t</th>
<th>pMCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>(Intercept)</td>
<td>2.517</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>response</td>
<td>.036</td>
<td>.283</td>
<td>.126</td>
<td>.900</td>
</tr>
<tr>
<td></td>
<td>position</td>
<td>-.011</td>
<td>.234</td>
<td>-.046</td>
<td>.963</td>
</tr>
<tr>
<td></td>
<td>response:position</td>
<td>.397</td>
<td>.321</td>
<td>1.237</td>
<td>.219</td>
</tr>
</tbody>
</table>

Table 4.19: Results of the linear mixed effect regression model adjusted to the ratings of BP /zl/, with response and position as fixed factors and subject as a random factor. (Intercept = <z>, intervocalic).

mixed effects regression model\(^7\) was conducted with response (<s> vs. <z>) and position (initial vs. intervocalic) as fixed effects on the rating of BP /zl/, and subject as a random factor (shown in Table 4.19 below).

The results did not yield any significant effects of response, or of position, and the interaction was not significant either. As with BP /s/ if the <s> and <z> responses were collapsed, it does not seem that significant differences would be obscured.

Table 4.20: Orthographic responses for BP /ʃ/ stimuli.

For the BP /ʃ/ stimuli, the most common response, as expected, was <y>; however, <sh> had considerable representation in the responses (see Table 4.20). Although not native to Spanish, <sh> is used in common English loanwords such as shorts and show, where it is realized as a postalveolar fricative. Although, originally borrowed as a voiceless postalveolar fricative /ʃ/, the degree to which speakers treat this as a different

\(^7\)See footnote 2
phoneme from the native /ʒ/ is also variable: some speakers realize it consistently as [ʃ], while others treat it as they would native /ʒ/, with all the variation that it entails (see Fontanella de Weinberg 1978). For further analyses <sh> and <y> were kept and the other responses were eliminated (<ll> and <y> responses were combined since <ll> is an orthographic representation of the postalveolar fricative in BAS).

<table>
<thead>
<tr>
<th></th>
<th>sh</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>13</td>
<td>49</td>
</tr>
<tr>
<td>Intervocalic</td>
<td>9</td>
<td>55</td>
</tr>
</tbody>
</table>

Table 4.21: Distribution by position of <sh> and <y> responses for BP /ʃ/ (<ll> and <y> combined).

Despite the apparent bias of <sh> responses by position, the relation between these variables was not significant, $\chi^2(1, N=126)=0.618$, $p=0.432$; therefore, the responses are not differentially distributed by position. A linear mixed effects regression model was conducted with response (<sh> vs. <y>) and position (initial vs. intervocalic) as fixed effects on the rating of BP /ʃ/, and subject as a random factor (shown in Table 4.22).

The results yielded a significant main effect of position, such that the ratings were higher in initial position (M=3.903, SD=1.051) than in intervocalic position (M=3.500, SD=1.260). Neither the main effect of response nor the interaction were significant.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>$\beta$</th>
<th>S.E.</th>
<th>t</th>
<th>$P_{MCMC}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>(Intercept)</td>
<td>3.944</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>response</td>
<td>-.114</td>
<td>.349</td>
<td>-.326</td>
<td>.745</td>
</tr>
<tr>
<td></td>
<td>position</td>
<td>-1.054</td>
<td>.375</td>
<td>-2.809</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td>response:position</td>
<td>.761</td>
<td>.415</td>
<td>1.833</td>
<td>.069</td>
</tr>
</tbody>
</table>

Table 4.22: Results of the linear mixed effect regression model adjusted to the ratings of BP /ʃ/, with response and position as fixed factors and subject as a random factor. (Intercept = <y>, intervocalic).
The subjects who wrote <sh> for the stimulus /ʃ/ are almost all from the north (8 northerners vs. 2 southerners), i.e. upper class speakers, presumably voicers themselves. Although it might be the case that upper class speakers are more likely to identify this sound as <sh>, the goodness ratings that they gave varied widely and actually do not differ significantly from the ratings that the other participants gave for their <y> responses. For this reason, it would be warranted to collapse these two responses.

For the BP /ʒ/ stimuli, the most common response was <y>, while <sh> came up only twice (see Table 4.23). Interestingly, the two participants who responded <sh> also responded <sh> for /ʃ/, one middle class and one upper class (the upper class participant gave <sh> a score of 5 as an example of /ʒ/). Given the trivial number of <sh> responses, only the <y> responses were used for the subsequent analyses (combined with the <ll>, as was done with the /ʃ/ stimuli). A t-test shows that there are no significant differences for the ratings of /ʒ/ with respect to position, t(125)=0.161, p=0.873.

In the results of the ABX experiments, there was an effect of position for the sibilant voicing pairs s/z and /ʃ/ only, such that they had lower accuracy rates in intervocalic position than in initial position. This is not mirrored in the goodness ratings, since the participants rated most of the intervocalic sibilants similarly to the initial sibilants. The only exception was /ʃ/, which was rated as a worse example of the BAS postalveolar category in intervocalic position.

<table>
<thead>
<tr>
<th>Response</th>
<th>da</th>
<th>g</th>
<th>gu</th>
<th>j</th>
<th>ll</th>
<th>sh</th>
<th>sh/ll</th>
<th>y</th>
<th>y/ll</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>2</td>
<td>1</td>
<td>117</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.23: Orthographic responses for BP /ʒ/ stimuli.
4.5.3 Cross-obstruent comparisons

Figure 4.4 displays the rating data for each stimulus obstruent and position, eliminating the infrequent responses and collapsing over the responses that did not significantly affect the goodness ratings. The effect of position can clearly be seen in the responses for /b, d, g/. Also noticeable are the very low ratings for /z/, which are not unexpected, considering the non-native status of this phoneme in BAS. Comparing across the results for different fricative stimuli can prove very informative in terms of the relative goodness of fit ratings for native, non-native and variable contrasts.

![Figure 4.4: Beanplots showing the rating distribution for each stimulus obstruent for initial and intervocalic position after the elimination of infrequent responses. Horizontal lines indicate the mean. “j” corresponds to /ʒ/ and “x” corresponds to /ʃ/.](image)

The data for each voiceless fricative were compared with the data of its corresponding voiced fricative, in order to investigate whether the voiced fricative stimuli were rated worse than their voiceless counterparts. It would be expected that there be differences for the labiodental and alveolar fricatives, since the /v/ and /z/ are not part of
the phonemic inventory of BAS; however, the same might not hold for the postalveolar stimuli, since these are both possible realizations of /ʒ/ in BAS.

Comparing the goodness ratings for /f/ and /v/, the results of a t-test show that ratings for BP /f/ stimuli are significantly higher (M=4.277, SD=0.881) than the ratings for BP /v/ stimuli (M=3.820, SD=1.038), t(248)= 3.807, p<.001. This is expected, since /f/ is part of the BAS Spanish phonemic inventory, and [v] is a very infrequent and marked realization of BAS /b/, so it would be a bad exemplar even for that category.

Comparing the goodness ratings for /s/ and /z/, the results of a t-test show that ratings for BP /s/ stimuli are significantly higher (M=3.730, SD=1.061) than the ratings for BP /z/ stimuli (M=2.619, SD=0.160), t(226)= 7.638, p<.001. This result is expected since /s/ is part of the BAS phonemic inventory, while /z/ is not and is assimilated to BAS /s/ as a very bad exemplar. However, an interesting fact is that the distribution of responses <s> vs. <z> is not uniform across stimulus obstruent, as seen in Table 4.24 below.

<table>
<thead>
<tr>
<th>BP Stimulus</th>
<th>/s/</th>
<th>/z/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response</td>
<td>&lt;s&gt;</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>&lt;z&gt;</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 4.24: Distribution of <s> and <z> responses for BP stimuli /s/ and /z/.

A chi-square test of independence showed that the relation between the listeners’ responses (<s> and <z>) and the type of stimulus (/s/ and /z/) is significant, χ²(1, N=235)=34.319, p<0.001. The <z> responses are more likely with BP /z/ than with /s/. The listeners in this task hear instances of [s] and [z], and when forced to choose a symbol to represent them, they choose <s> to represent the more canonical [s] and to distinguish the two fricatives, then choose <z> for the non-canonical [z]. They know that <z> is pronounced differently in other varieties of Spanish (such as /θ/, which they are sometimes told in school for purposes of dictation) and thus are associating it with
this strange and non-ideal exemplar of BAS /s/. Or at least that they know that <z> is somehow “other” and they mark the stranger stimulus with this “other” symbol.

Comparing the goodness ratings for /ʃ/ and /ʒ/, a t-test shows that there are no significant differences for the ratings between BP /ʃ/ and /ʒ/ stimuli, t(250)=0.624, p=0.531. This follows from the fact that both [ʃ] and [ʒ] are possible realizations of BAS /ʒ/. It would be interesting to see how the ratings for these two fricatives compare to the ratings of the other two fricatives present in the BAS inventory: /ʃ/ and /s/. For this purpose, a subset of the fricative rating data was extracted just for these four BP obstruents: /ʃ, s, ʃ, ʒ/. The results of a one-way ANOVA show that the goodness ratings differ significantly depending on the stimulus obstruent, F(3, 502)= 10.619, p<0.001. A Tukey HSD test was performed to explore the difference between stimulus groups and it revealed that the ratings for /ʃ, ʃ, ʒ/ were not significantly different from each other (all p values >.80), whereas the ratings for /ʃ/ were significantly higher than those for the three sibilants (all p values <.001).

Parallel to what was shown for production in Chapter 2, where the voicing levels of the postalveolar fricative are approximating those of the alveolar fricative /s/, the three sibilants all seem to be equivalently acceptable exemplars of their native categories. However, it is somewhat perplexing that the ratings for BP /s/ are relatively low, particularly compared to /ʃ/, since they both form part of the phonemic inventory of BAS. Perhaps the /s/ in BP, although identifiable as /s/ to BAS listeners, has some different spectral qualities than the /s/ in BAS. It might also be that the stimuli, although carefully selected, differed on a parameter that was not anticipated to affect the goodness ratings.

The salient result from these comparisons is that it seems to be the case that the ratings for /ʒ/ and /ʃ/ do not differ significantly from each other, which could indicate that they are equally acceptable variants to speakers. However, the comparisons discussed in
this section do not take into account the possible effect of social factors on the goodness of fit ratings for BP /ʃ/ and /ʒ/. These effects are discussed below.

4.5.4 Cross-group comparisons

4.5.4.1 Class

One of the main themes of this investigation is the social variation in production and perception of the postalveolar fricatives in BAS; for this reason it is particularly interesting to compare the ratings of the postalveolar fricative stimuli between social groups. Recall that the goodness ratings for /ʃ/ and /ʒ/ in the whole dataset are not significantly different from each other, t(250)=0.624, p=0.531. However, given Wolf’s (1984) and Colantoni’s (2008) results, one would expect the upper class speakers to rate /ʒ/ stimuli higher than /ʃ/ in terms of goodness of fit.

The ratings of /ʃ/ and /ʒ/ in Figure 4.5 show similar distributions between upper class and middle class speakers. To investigate the effect of class on the ratings of BP /ʃ/ and /ʒ/, a linear mixed effects regression model was conducted with class (upper vs. middle) and stimulus fricative (BP /ʃ/ vs. BP /ʒ/) as fixed effects, and subject as a random factor (shown in Table 4.25).

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>β</th>
<th>S.E.</th>
<th>t</th>
<th>p_{MCMC}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>(Intercept)</td>
<td>3.736</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>class</td>
<td>-.102</td>
<td>.298</td>
<td>-.341</td>
<td>.724</td>
</tr>
<tr>
<td></td>
<td>stimulus fricative</td>
<td>-.054</td>
<td>.187</td>
<td>-.289</td>
<td>.797</td>
</tr>
<tr>
<td></td>
<td>class:stimulus fricative</td>
<td>-.047</td>
<td>.238</td>
<td>-.199</td>
<td>.823</td>
</tr>
</tbody>
</table>

Table 4.25: Results of the linear mixed effect regression model adjusted to the goodness ratings, with response and position as fixed factors and subject as a random factor. (Intercept = BP /ʒ/, upper class).
Figure 4.5: Beanplots showing the goodness rating distribution for postalveolar stimuli divided by social class. The dark horizontal line indicates the mean. “j” corresponds to /ʃ/ and “x” corresponds to /ʃ/.

The results did not yield any significant effects of class or of stimulus obstruent, the interaction was not significant either. This seems to indicate that, contrary to what one would predict on the basis of Wolf’s (1984) and Colantoni’s (2008) findings, speakers’ social class does not have an impact on whether they rate /ʃ/ or /ʃ/ as a better exemplar of the native postalveolar category.

4.5.4.2 Age

In Chapter 2, age was found to be a significant factor influencing the completion of the change from /ʃ/ > /ʃ/. Furthermore, the devoicing of /ʃ/ has been a change in progress in the Buenos Aires speech community for almost half a century. Perhaps the ratings
of the postalveolar fricatives might be correlated with age. If the younger speakers find /ʃ/ as a better exemplar of their native postalveolar category, then one would expect that the goodness ratings would be negatively correlated with age. If the older speakers find /ʒ/ as a better exemplar of their native postalveolar category, then one would expect that the goodness ratings be positively correlated with age.

A Pearson’s test shows that age and rating for /ʃ/ were significantly negatively correlated, r(124)=-0.24, p=.007. However, age and rating for /ʒ/ were not significantly correlated, r(126)=0.07, p=0.434. (See Figure 4.6)

![Figure 4.6: Correlations of rating and age for the postalveolar fricative BP stimuli (the correlation for /ʃ/ is significant).](image)

Although the effect of social class was not significant, age does seem to be a significant factor affecting the rating of the /ʃ/ stimuli. It seems that the older the speaker, the less likely they are to rate /ʃ/ as a good exemplar of their postalveolar category.\(^8\) This re-

\(^8\)It should be noted that the correlation is still significant when excluding the <sh> responses, r(102)=-0.288, p=0.003.
result does not necessarily conform to King’s (2009) findings, since most the respondents that were aware of the variation considered /ʃ/ negatively as a cheto (“snobby”) feature, and in these findings there is no effect of age in the rating of /ʃ/.

4.5.4.3 Voicing status: voicers vs. devoicers

In order to investigate whether variation in production of the postalveolar fricative in BAS affects the rating of the postalveolar BP stimuli, a subset of the ratings data was selected that corresponded to the speakers analyzed in Chapter 2. All the speakers whose mean voicing levels for /ʃ/ were not significantly different from the voicing levels for /s/ were considered the devoicers (7/16 participants); whereas, those speakers whose voicing levels for /ʃ/ were significantly higher than for /s/ were considered the voicers (9/16 participants). The distribution of the ratings can be seen in Figure 4.7.

Figure 4.7: Beanplots showing the goodness rating distribution for postalveolar stimuli divided by voicers and devoicers. The dark horizontal line indicates the mean; “j” corresponds to /ʃ/ and “x” corresponds to /ʃ/.  

154
In order to investigate the effect of the listeners voicing status on the ratings of BP /ʃ/ and /ʒ/, the data were analyzed in a linear mixed effects regression model, with subject as a random factor and stimulus fricative (within-subject), and (de)voicer (between-subject) as fixed effects.

<table>
<thead>
<tr>
<th>Response</th>
<th>Predictor</th>
<th>β</th>
<th>S.E.</th>
<th>t</th>
<th>p MCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>(Intercept)</td>
<td>3.254</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>stimulus fricative</td>
<td>.674</td>
<td>.231</td>
<td>2.925</td>
<td>.009</td>
</tr>
<tr>
<td></td>
<td>(de)voicer</td>
<td>.524</td>
<td>.449</td>
<td>1.167</td>
<td>.159</td>
</tr>
<tr>
<td></td>
<td>(de)voicer:stimulus fricative</td>
<td>-1.348</td>
<td>.307</td>
<td>-4.385</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Table 4.26: Results of the linear mixed effect regression model adjusted to the ratings of BP /ʃ/ and /ʒ/, with subject as a random factor. (Intercept = /ʃ/, voicer).

The results in Table 4.26 show that there is a significant effect of stimulus fricative and a significant interaction of stimulus fricative and listener’s voicing status. The /ʃ/ stimuli (M=3.500, SD=1.186) received slightly, yet significantly, lower ratings than the /ʒ/ stimuli (M=3.590, SD=1.116). To investigate the effects of the interaction a TukeyHSD test was conducted and it revealed that the only significant comparison is between the voicers’ and the devoicer’s ratings of /ʃ/. For BP /ʃ/, the voicers’ ratings (M=3.125, SD=1.289) were significantly lower than the devoicers’ ratings (M=3.929, SD=0.900). This significant interaction suggests that, while voicers and voicers do not differ in their ratings of /ʒ/, they do differ on how they rate the BP obstruent that corresponds to the innovation in BAS /ʃ/. Furthermore, a Pearson’s test reveals a significant negative correlation between the listeners’ mean voicing rates for /ʒ/ in their production and their ratings of BP /ʃ/, r(60)=−.500, p<.001. However, the same correlation was not significant for the BP /ʒ/ stimuli, r(61)=.109, p=.405.

These correlations in Figure 4.8 suggest that the speakers who are lagging in the change might consider the innovation a worse fit for their native postalveolar category.

155
Figure 4.8: Correlations of the listeners’ mean % voicing of /ʒ/ in production and their ratings of the two postalveolar fricative BP stimuli (The correlation for /ʃ/ is significant).

The more the speaker voices their /ʒ/, the worse that they will judge /ʃ/. This could be indicative of the correction from above, which is predicted by Labov (1966:225) for changes “from below” that are nearing completion. Furthermore, it might provide evidence for the fact that the devoicers are not aware of the voicing difference in the postalveolar category; or if they are aware, that they accept either variant, regardless of their production.

4.6 Conclusion

The results from these transcription and goodness rating tasks provide the necessary information in order to understand and fully interpret the results from the discrimination experiments. All the BP stops were assimilated to their corresponding stop categories in BAS, even in intervocalic position. This is in accordance with Best’s PAM, since,
despite the occlusion in intervocalic position, the listeners were able to perceive the articulatory invariants in the production of the voiced stop category rather than rely on the acoustic cues, as Flege’s SLM would predict, which were more similar to those of phonologically voiceless (but phonetically voiced) occluded stops in Spanish. Additionally, the rating task is the only task where position was a significant factor for the stops. Although voiced occluded stops in intervocalic position were unmistakably assimilated to their corresponding voiced stop category, they were nevertheless a worse fit for that category; listeners do perceive this difference.

These results also helped elucidate the assimilation patterns for the fricative voicing pairs. The transcription patterns and lower ratings for BP /v/ suggest that the discrimination of the f/v voicing pair is indeed a case of Two Category assimilation, where BP /f/ maps onto BAS /f/ and BP /v/ maps onto BAS /b/. The transcription and rating results for the alveolar fricatives suggest that the s/z voicing pair is a case of Category Goodness assimilation, since both BP /s/ and BP /z/ were assimilated to BAS /s/; however, /z/ was a much worse fit for that category. This wide divergence in goodness of fit is probably what helped the listeners in the ABX experiment in distinguishing that contrast more accurately than the postalveolar contrast. The current data seem to confirm that the f/ʒ voicing pair is a case of Same Category assimilation, since both BP /f/ and BP /ʒ/ were equivalently fit for the native postalveolar category, whose voicing is variable in the Buenos Aires speech community. It is not the case, however, that the ratings were equivalent for all speaker groups.

Contrary to what would be expected from Wolf’s (1984) and Colantoni’s (2008) findings, social class was not a significant factor affecting the ratings of either of the postalveolar fricative stimuli; however, age and devoicing status were significant. In terms of the significant social factors, it seems to be the case that older speakers and
voicers of /ʃ/ give BP /ʃ/, the fricative associated with the native innovation, a lower rating. However, the devoicers do not rate BP /ʒ/ any lower than /ʃ/. This could be evidence that the younger speakers, particularly those for whom the devoicing of /ʒ/ has reached completion, have become unaware of the voicing difference in the postalveolar fricative, or if they are aware of the difference, they might deem both variants as equally acceptable. It is not that they are not sensitive to voicing in fricatives across the board, because they can tell the difference more accurately between /s/-/z/ as opposed to /ʃ/-/ʒ/, but rather that the sociophonetic variation in the speech community has expanded their postalveolar category such that they cannot distinguish a non-native voicing contrast that they might be able to otherwise.
Conclusions & future research

5.1 On variation in production

This dissertation contributes to the Spanish sociolinguistics literature by re-examining the classically studied voicing variation of /ʒ/ in Buenos Aires Spanish using instrumental measures in order to determine the progress of the devoicing change. By incorporating into the analysis not only the socially conditioned voicing variation in /ʒ/ but also the phonetically conditioned voicing variation in /s/, it proposes a methodologically innovative instrumental criterion to determine whether the devoicing change has reached completion. This novel /s/-criterion provides a system internal benchmark for the completion of change, which does not depend on the perceptual biases of the investigator and allows for gradient phonetic variation that is inherent to the system.

The results from the instrumental analysis of the devoicing of /ʒ/ show that it is affected by phonetic factors such as position and speech rate in a similar way to /s/: these sibilants exhibit more voicing in intervocalic position and in more connected speech. In addition to these phonetic effects, the devoicing of /ʒ/ is affected by social factors such as age and social class, whereby younger speakers and middle class speakers present per-
percentage voicing rates that are more similar to the voicing rates of the voiceless fricative /s/. As would be expected, given the results of previous studies (Wolf and Jiménez 1979, Wolf 1984, Chang 2008, Rohena-Madrazo 2008), the younger, middle class speakers were those for whom the devoicing change had been completed. However, contrary to Chang (2008), the change has not been completed for all younger speakers, since there are still younger upper class speakers who maintain the voicing of /ʃ/. Furthermore, although social class is a significant factor determining the voicing of /ʃ/, contrary to Wolf (1984), the voicing variation of /ʃ/ does not appear to be stable. The three other groups that still maintain higher voicing rates for /ʃ/ than for /s/ (younger upper class, older upper class, and older middle class) all present evidence that the change is still ongoing. In fact, an analysis of speakers on an individual level shows that there are devoicers in every social group studied here. For these devoicers, the more accurate description of the postalveolar sibilant in their phonemic inventory is a /ʃ/, since it exhibits voicing rates that are not significantly different from that of phonologically voiceless sibilant /s/. The findings from the production study show that there is still much variation in the production of /ʃ/ in the Buenos Aires Speech community that is both socially and phonetically conditioned. The examination of the social distribution of the variation suggests that the devoicing change of /ʃ/ is still ongoing, but it might be in the advanced stages, since for a sizeable portion of the population, particularly for the younger speakers, the change has already reached completion.
5.2 On variation in perception, informed by variation in production

The results from the production experiments raise the question of how speakers represent or categorize this socially determined voicing variation in the speech community (/f~ʒ/). Phonologically, both voiced and voiceless variants are realizations of the same phonemic category; however, the differing distribution of the variants, indexed by age and by social class, might allow the speakers to develop sociophonetic categories at the subphonemic level. If this is the case, then the categorization of the different variants might affect the discrimination of an analogous non-native contrast /f/-/ʒ/: if they have two socially motivated categories, listeners should discriminate more accurately than if the two variants are treated as equally acceptable exemplars of the same category. Cross language discrimination experiments were conducted in order to explore the effects that socially conditioned variation and positionally defined allophonic variation may have on the delimitation of perceptual categories used to discriminate non-native obstruent voicing contrasts. Two discrimination experiments, AX and ABX, were conducted testing the discrimination of six Brazilian Portuguese obstruent voicing pairs (p/b, t/d, k/g, f/v, s/z, f/ʒ), in intervocalic and in phrase initial position.

The primary result from the discrimination of stop voicing contrasts showed that positionally conditioned allophones are not the relevant unit of representation for the perception of voicing contrasts in non-native positions. Each of the three stop voicing pairs are phonemic in BAS; however, in intervocalic position the implementation of the stop voicing contrast is different in Spanish than in Portuguese. In Portuguese the intervocalic voiced stops are realized as such (e.g. /VdV/ > [VdV]), whereas in Spanish intervocalic voiced stops undergo spirantization (e.g. /VdV/ > [VðV]); there-
Therefore, [VdV] would be a case of a native contrast occurring in a non-native position. Two models of cross-language speech perception (Flege’s Speech Learning Model and Best’s Perceptual Assimilation Model) make different predictions on the accuracy of discrimination of these contrasts in intervocalic vs. initial position. SLM, which posits that the non-native sounds are associated to positionally defined allophones in the native language, predicts lower accuracy rates for the stop voicing contrasts in intervocalic position ([VtV]-[VdV]) as opposed to initial position ([tV]-[dV]). The fact that BP [VdV] is not associated with the BAS positional allophone [VdV] and acoustically is more similar to positional allophones of /VtV/ would increase the confusability of the BP contrast [VtV]-[VdV]. PAM, on the other hand, predicts high accuracy in both intervocalic and initial contrasts, since the listeners perceive the articulatory invariants present in the articulation of /d/ common across positions. The results from both the AX experiment and the ABX support the predictions made by PAM, since the listeners do not show any significant differences in discriminating the stop voicing contrast in initial vs. intervocalic position. Furthermore, the results from the perceptual assimilation tasks confirm that the three stop voicing contrasts were perceived as Two Category assimilation types, where the BP stop maps onto the corresponding BAS stops, regardless of position. Interestingly, the rating task was the only experiment which showed a difference between the stops by position. The listeners consistently rated the intervocalic voiced stops as worse exemplars of their native categories; however, under PAM this should not have affected the discrimination of an example like [VtV]-[VdV], since the [VdV], although a bad exemplar of /VdV/ is still mapped onto a category that is different from /t/, thus ensuring high discriminability in the AX and ABX tasks. These results suggest that the ability of a listener to discriminate a non-native contrast might be most affected by the
relationship of the non-native phonemes to the native phonological categories, rather than to position specific allophones.

For the fricative contrast, results showed that socially conditioned variation did not aid the listeners in distinguishing fricative voicing distinctions. The fricative voicing contrast is not native to Spanish; however, the Perceptual Assimilation Model makes differing predictions for each of the fricative voicing pairs: \( f/v \) as Two Category, \( s/z \) as Category Goodness, and \( f/z \) as Same Category assimilation, in decreasing order of discrimination accuracy. The main comparison to determine the effect of social variation was between \( s/z \) and \( f/z \), where the postalveolar fricative is the one that presents socially conditioned variation. The results from the ABX task show that, instead of improving the discrimination of BP \( f/z \), having social variation in the native speech community within that category (i.e., \( f/\sim z/ \)) actually hinders the discrimination of this contrast. Listeners do not seem to tap into the socially conditioned differential distribution of \( f \) and \( z \) in BAS in order to discriminate the non-native contrast. Instead, it seems to be that the exposure that the listeners have to the different variants actually widens the range of voicing acceptability for the postalveolar category, compared to a category like the BAS /s/, where the voicing variation found in the community is not as ample as the variation for /z/. This in turn increases confusability and interferes with discrimination.

One revealing finding from both the AX and ABX tasks is that the listeners were equally inaccurate at discriminating /f/-/z/ across ages and social classes, which were precisely the factors that conditioned the voicing variation in production. This finding contrasts with previous accounts (Wolf 1984, Colantoni 2008, King 2009) where speakers were aware of the devoicing of /z/, associating it with upper class speakers, and, in cases like Wolf (1984), the upper class speakers even actively increased their voicing rates of /z/ to mark the social distinction. In the current data, even when the subset of
voicers and devoicers are examined, there was no effect of the listener’s own voicing status on the accuracy rates of the postalveolar voicing contrast. This suggests that there is not necessarily a direct, straightforward link between production of sociolinguistic variants and the perception of an analogous non-native contrast. Although the voicers produced significant voicing differences between /ʒ/ and /s/, the same voicers showed very low accuracy rates in the discrimination between BP /ʃ/. This seemingly contradictory fact is reminiscent of the near merger phenomenon. In a case of a near merger listeners do not perceive a difference between a historically phonemic contrast; however, they produce the contrast reliably (Labov et al. 1972, Yu 2007). In these cases, the expected outcome is one similar to the results of the discrimination tasks, particularly for the /ʃ/ contrast. The main difference between the present case and a near merger is that the distinction is not a phonemic one, but one that is socially differentiated. However, this interpretation of the lack of effect of native production on cross-language perception is further complicated by the results of the perceptual assimilation task.

In the perceptual assimilation task, the predictions made by PAM in terms of assimilation types were confirmed. The BP voicing pair /f/ was mapped onto two different categories, BAS /f/ and /b/ respectively, which explains the ceiling level accuracy rates. The members of the s/z voicing pair were both mapped onto BAS /s/, but the goodness of fit of BP /z/ to the BAS category was very low, the lowest among all of the stimuli. This confirms the Category Goodness status of the s/z voicing pair. Both members of the BP /ʃ/ voicing pair were assimilated to the same category: the BAS postalveolar fricative. Given this assimilation pattern and the lowest accuracy rates for BP /ʃ/, the case for this contrast as a Same Category assimilation seems well founded; however, social factors affecting the ratings of BP /ʃ/ complicate the SC interpretation. There are significant effects of age and the listener’s voicing status that affect their ratings of /ʃ/
but not of /\z/\/. Both age and percentage of voicing of /\z/ in production negatively correlate with the ratings of the voiceless postalveolar fricative; in other words, the older the listeners and the more that they voice their /\z/, the lower they will rate BP /\f/.

It is particularly perplexing that despite the fact that listeners showed low accuracy in discrimination of the /\f/\z/ contrast in the ABX results (i.e. they do not seem to distinguish the contrast) the voicers give significantly lower ratings to the /\f/ stimuli than they do to the /\z/ stimuli. The same effect is not found for the BP /\f/ stimuli. Given the results from the perceptual assimilation task, it does seem that the two BP postalveolar fricatives map onto the same native postalveolar fricative category. However, the type of assimilation pattern might differ depending on the voicing status of the listener and depending on age, which was shown to be significantly positively correlated with the voicing of /\z/, although it should be noted that the listener’s average percentage of voicing of /\z/ had a higher correlation coefficient (r=−.50) than age did (r=−.24). It seems that for the devoicers the /\f/\z/ contrast is clearly a Same Category assimilation pattern, but the fact that the voicers rate the /\f/ stimuli significantly lower than the /\z/ indicates that for them it is a Category Goodness assimilation. These results partially confirm Wolf’s (1984) claims that some speakers perceive and negatively judge the devoicing change. The results also partially confirm Colantoni’s (2008) findings, where only the listeners from the upper social strata identified /\z/ with speakers of their own social class, whereas the listeners from the lower classes seemed unaware of the variation. However, the key difference in the current study is that the awareness of the devoicing does not seem to be a class effect, but rather it depends on the speaker’s production of voicing in the postalveolar fricative. In Wolf’s sample all of the upper class speakers were voicers themselves; therefore, that class effect may have been due to the speakers’ voicing status factors. What seems clear is that the devoicers in the current study do not share the voicers’ lower rating of /\f/; in

165
fact, contrary to what King’s (2009) findings would suggest, devoicers seem to rate both /ʃ/ and /ʒ/ as equally good exemplars of their native postalveolar category. One possibility is that the devoicers are aware of a difference in voicing but they do not judge it differently. However, considering their low discrimination rates for the BP /ʃ/–/ʒ/ voicing pair, it could be, as suggested by Wolf and Jiménez (1979), that the devoicers are not aware of the voicing difference in the production of /ʒ/ in the BAS speech community.

According to PAM, if the voicers are aware of the variation in their speech community, and therefore judge /ʃ/ as a worse fit for the postalveolar category, this should aid the voicers in discriminating the /ʃ/–/ʒ/ in the ABX experiments. However, this was not the case. One possible explanation for this seemingly contradictory result is that it is due to a task effect. The nature of the ABX discrimination task and the perceptual assimilation/rating task are somewhat different. In the ABX task there are three stimuli that the listener must compare against each other in an attempt to label them, whereas in the perceptual assimilation/rating task they only hear one and compare it to some idealized representation. It may be that, in the ABX task, the discrimination is not as fine-grained as intended, perhaps the emphasis placed on the “foreign word” might still have an effect on the listeners disregarding some relevant differences, since in BAS the three stimuli would still be the same acceptable word. On the other hand, in the perceptual assimilation task, the listeners hear only one stimulus and they have to transcribe the single obstruent sound within the stimulus and compare it to some abstract representation of what they consider a prototypically voiced exemplar of the native postalveolar category. This could make the differences in the /ʃ/ with respect to their own category more evident, hence the lower ratings. Furthermore, there might be some effect of orthography in the perceptual assimilation/rating task that is reinforcing the voicing difference for the voicers’ idealized representation, which is not present in the ABX task. It could
be that for the voicers the orthography is more directly mapped to their voiced variant /z/ than for the devoicers, and thus the voicers are more likely to reject the [j] variant, which might be less congruent with their phonological-orthographical representation. Regardless of the reasons for this contradictory effect, it calls into question the clear distinction between Category Goodness and Same Category assimilation types, particularly, how much worse of an exemplar of a native category does a non-native phoneme need to be before it has an effect on the discrimination of a non-native contrast? In Best et al. (2001), goodness scales were not used, since the determination between CG and SC was determined by categorical orthographic means. Best and colleagues have used goodness-of-fit Likert scales in other studies of perceptual assimilation (Best et al. 2003, Hallé et al. 2003); however, they do not provide a criterion to determine the cutoff point between CG and SC assimilations. Instead, they make the determination by comparing the relative difference between two ratings, assigning category goodness to the relatively lower one. Strange (2007:54), in her review of the theoretical and methodological issues in the study of cross-language phonetic similarity of vowels, concludes that this indeterminate nature of the boundary between Same Category and Category Goodness assimilation types remains an unresolved question and that further research is needed on how listeners make use of category-goodness rating scales in order to appropriately interpret their responses in the light of these issues.

What can be concluded from the examination of the production, discrimination and perceptual assimilation results is that the relationship between variation in production and variation in cross-language perception is quite complex. Being exposed to a wider range of variation that is socially conditioned, such as /ʃ~ʒ/ is in BAS, does seem to have an effect on the discrimination of an analogous non-native distinction. However, the effects of socially conditioned variation in production on cross-language perception,
at least for BAS listeners, only surface in a perceptual assimilation rating task and seem to go only in one direction: the devoicers do not rate /ʂ/ significantly differently from /ʃ/, whereas the voicers do, giving /ʃ/ lower ratings. This seems to indicate that a more inclusive norm might be emerging, particularly as the devoicing change reaches completion, a norm that includes both [ʃ] and [ʂ] as acceptable variants. In other words, the more innovative speakers (the devoicers) are the ones who have a more expansive postalveolar fricative category, whereas the conservative speakers (the voicers) have a smaller range of acceptable productions. However, any awareness of this smaller range of acceptable variation, which might be causing the lower ratings for BP [ʃ], is not aiding the voicers in the differentiation of a non-native /ʃ/-/ʂ/ contrast along those same lines. This effect warrants more exploration, both in terms of the effects of sociophonetic variation on the representation of phonological categories, as well as in terms of the social evaluation of the variable voicing of /ʂ/ in Buenos Aires Spanish.

5.3 Directions for future research

One logical next step for future investigation would be to extend the analysis of the devoicing of /ʂ/ to other speech styles, such as a picture description task or an unscripted conversation. Presumably, in more casual connected speech styles there should be more gestural overlap, and therefore more voicing, especially in intervocalic position. However, a more detailed examination of speech rate would be needed. In Section 2.5.5, the preliminary metric of fricative duration was used as a possible correlate of speech rate. A future study would need to take into consideration other factors such as number of syllables per second, pause duration, pause frequency, among other factors, which may reveal a clearer pattern in the variation in order to tease apart the effect of speech rate from task style or attention paid to speech. Moreover, it would be interesting to explore
further the effect of age on speech rate, particularly to determine whether the effect of age on the duration of the fricative is due to sociolinguistic or to physiological factors. For this it would be useful to corroborate whether Spanish exhibits the same effects of age on speech rate, where younger speakers have a higher rate than older speakers, as found by Shipp et al. (1992) for English.

Another independent variable that could be analyzed is lexical frequency to determine its effect on the devoicing change. Bybee (2002:271) states that “changes that affect high-frequency words first are the result of the automation of production, while low-frequency words change first when the change makes the words conform to the stronger patterns of the language.” Examining the lexical diffusion of the devoicing of /f/ could shed more light on the underlying causes of this change in BAS as well as provide more information on the possible distribution of the phonetic characteristics of the exemplars that would make up the exemplar clouds of the lexical items containing /f/.

In Chapter 2 (see Section 2.5.2) it was demonstrated that even phonologically voiceless fricatives like /s/ are not actually voiceless all of the time. They exhibit variation in their percentage of voicing depending on phonetic factors. In this study, the variable voicing rates of /s/ served as a benchmark against which to compare the socially conditioned voicing variation in /f/ in order to determine whether the devoicing change had reached completion. It would be illuminating to investigate to what extent this voicing variation is shared by the other phonologically voiceless fricatives in the BAS phonemic inventory: /f, x/. This would allow for a more comprehensive criterion for determining the baseline voicing rates for a voiceless fricative and consequently for determining the end of the /f/ devoicing change in BAS. Furthermore, as mentioned in 2.4.2, percentage voicing is not the only correlate of fricative voicing. Other possible correlates,
such as relative intensity or duration of the fricative and its adjacent segments, could be measured to determine whether they are also affected by social and phonetic factors.

A similar investigation could be extended to the stops as well, in order to determine what type of voicing variation voiceless stops exhibit in BAS. It has been found that in many varieties of Spanish voiceless stops undergo voicing and even spirantization in casual speech at very high rates (Machuca Ayuso 1997, Martínez Celdrán 2009, Torreira and Ernestus to appear, Lewis 2001), such that [VdV] can often surface as a realization of /VTV/. In theory, this could lead to an association of occluded voiced stops with the voiceless stop category in Spanish. However, it was not the case that occluded voiceless stops were consistently misidentified as voiceless stops in the perceptual assimilation task. It would be useful to investigate also the voicing/lenition rates of voiceless stops in BAS Spanish, to determine whether this effect is due to a relative infrequency of voicing of voiceless stops (as was found by Colantoni and Marinescu (2010) for other varieties of Argentinean Spanish) or whether they do present voicing variation but it does not affect the discrimination of the non-native contrast. Another possibility for the lack of confusion of [VdV] for /VTV/ is that the voicing of voiceless stops is associated with casual, connected speech. The stimuli in the discrimination and perceptual assimilation tasks were presented as isolated “words”, in a very controlled context, where the realization of /VTV/ as [VdV] might be incongruous with the situational context of casual speech. A discrimination experiment where the stimuli were presented in a casual carrier phrase might increase the confusability of the non-native /VTV/-/VdV/, thus providing further evidence that categorial perceptual boundaries can be modified depending on external contextual manipulations (Hay et al. 2006, Niedzielski 1999).

In terms of the effect of social variation on cross-language speech perception, it seems that the exposure to a greater range of voicing caused by the presence of social
variation in /z/ creates a wider range of acceptability for the postalveolar fricative category, which explained the differences in discrimination between s/z and J/z. In order to corroborate the social explanation for this effect, it would follow to replicate this study with listeners of other varieties of Spanish where [ʃ] and [ʒ] are not in sociolinguistic variation. If the interpretation of the social variation effects is correct, then these listeners of other varieties should not exhibit different accuracy rates between s/z and J/z. This would be expected as long as the BP voicing pair /ʃ/ is a case of Same Category assimilation, as they are in Buenos Aires Spanish. If the ABX study is extended to speakers of other varieties it is essential that the perceptual assimilation task be administered as well. If the BP /ʃ/ and /ʒ/ map onto two different native categories, /tʃ/ and /tʒ/ for example, then the accuracy results would not be comparable to those of Buenos Aires Spanish. It would, however, provide a very interesting case to study the potential cross-dialectal differences in the categorization of non-native contrasts.

The most obvious and needed extension to this investigation is the development of a subjective reaction experiment using BAS stimuli in order to determine whether and to what extent the speakers in the BAS speech community are aware of the devoicing of /ʒ/. A very useful side-effect of the present investigation is that it provides hundreds of instances of /ʒ/ with varying levels of voicing, which could be employed as is or acoustically modified as stimuli in these subjective reaction tests. This would shed light on whether it is the voicing of /ʒ/ that is the most socially salient feature or whether it is other cues, such as affrication in the CV context, fronting of the place of constriction, or other acoustic or articulatory correlates, relative intensity with the adjacent segments, that are identified more saliently with upper class speakers. Such an extension would provide a more complete picture of the variation and change in the voicing of /ʒ/ in Buenos Aires Spanish.
5.4 Concluding remarks

This dissertation has re-examined one of the classic cases of sociolinguistic variation in Spanish dialectology: the devoicing of /ʒ/ in Buenos Aires Spanish. By approaching it from an experimentally motivated instrumental perspective, this investigation has provided new insight into this phenomenon. Not only has this study provided a current profile of the social distribution of the voicing variation in /ʒ/, it has also proposed a novel method in order to determine the completion of the devoicing change, by comparing the gradient voicing variation in /s/ to the gradient variation of /ʒ/. Understanding the inherent variability in the system provides firmer footing on which to make determinations about the representation of sociophonetic variation and the completion of change in the sibilants of Buenos Aires Spanish. Furthermore, the experimental results in this dissertation have provided evidence suggesting that the sociophonetic variation in the system affects how well listeners are able to distinguish between phonetic differences that in another language may result in separate phonemes. By analyzing the interaction of phonological, phonetic, and social factors on speech production and perception, one can better understand the nature of phonological structure as well as phonological change.
Word List Reading Task

The following is the list of words that was presented to the participants in the Word List Reading Task discussed in Chapter 2. The highlighted sequences indicate the sibilants that were analyzed: /ʒ/ are in green and /s/ are in blue.
detergente
pasaje
Carlos
desembarcaron
queríamos
perfecta
Humberto
manganeso
llénenle
acampa
infinito
depende
campaña
sífilis
palta
guerrilleros
escarpines
señal
delfines
huérfano
golpea
enfermera
blandenguerías
pena
tratarse
sílfides
elfos
álgebra
uranio
alpargatas
departamento
sudcoreano
acarqué
alpacas
arremangarte
Atahualpa
delta
subcampeón
temen
Guillermo
zarzamora
arrancaba
Temperley
| 130 | bajen | 173 | sachet | 216 | irse |
| 131 | subcontinente | 174 | bife | 217 | manjares |
| 132 | virginal | 175 | ánforas | 218 | adyacente |
| 133 | precipitación | 176 | olímpico | 219 | siguieron |
| 134 | zapatillas | 177 | faltaría | 220 | huervanita |
| 135 | Pimpinela | 178 | trabajadores | 221 | interjección |
| 136 | bancaremos | 179 | dentro | 222 | después |
| 137 | adjetivos | 180 | universitario | 223 | flamenco |
| 138 | shopping | 181 | Benjamín | 224 | máltes |
| 139 | signa | 182 | terquedad | 225 | yñanda |
| 140 | penal | 183 | pequeño | 226 | alternativa |
| 141 | círculo | 184 | subte | 227 | damnificados |
| 142 | dijiste | 185 | levále | 228 | panó |
| 143 | Ayacucho | 186 | políticas | 229 | yñarés |
| 144 | observación | 187 | insípido | 230 | fantástico |
| 145 | felicitaciones | 188 | verano | 231 | alfajores |
| 146 | índice | 189 | quirquincho | 232 | dije |
| 147 | subyugado | 190 | pasan | 233 | pensé |
| 148 | partidos | 191 | digital | 234 | lambada |
| 149 | déjense | 192 | quince | 235 | principé |
| 150 | Rafael | 193 | jirafas | 236 | absolutamente |
| 151 | berenjenas | 194 | Tarzán | 237 | tambaleándose |
| 152 | pantalla | 195 | subconsciente | 238 | apartada |
| 153 | genia | 196 | sindicato | 239 | tiraste |
| 154 | guerra | 197 | bajáme | 240 | subjetiva |
| 155 | yirando | 198 | invitación | 241 | gamba |
| 156 | Félix | 199 | palpitaciones | 242 | plata |
| 157 | esfinge | 200 | empequeñecerte | 243 | moñito |
| 158 | centenares | 201 | gingivitis | 244 | fanfarrias |
| 159 | gin tonic | 202 | invitén | 245 | manzana |
| 160 | ínfima | 203 | Melquíades | 246 | salpiques |
| 161 | Pepe | 204 | quizás | 247 | vértebra |
| 162 | presenta | 205 | independencia | 248 | eliminar |
| 163 | algar | 206 | absurdo | 249 | Algeciras |
| 164 | falsificar | 207 | danza | 250 | ganas |
| 165 | pertenecen | 208 | circulo | 251 | carcajadas |
| 166 | Sergio | 209 | inyecta | 252 | raya |
| 167 | riquísimas | 210 | participación | 253 | embalsamadas |
| 168 | criticó | 211 | Tito | 254 | gerente |
| 169 | Japón | 212 | yate | 255 | pasajeros |
| 170 | arquitecto | 213 | Eugenio | 256 | barbá |
| 171 | gire | 214 | alcaparras | 257 | Germán |
| 172 | calzado | 215 | obtener | 258 | púlpito |
Sentence Reading Task

The following is the list of sentences that was presented to the participants in the Sentence Reading Task discussed in Chapter 2. The highlighted sequences indicate the sibilants that were analyzed: /s/ are in green and /s/ are in blue.
1. Me acerqué para ver mejor.
2. Ese comité está encargado de fomentar la participación en los programas extracurriculares.
3. Venden manjares de todo tipo en esa confitería, y de yapa siempre te dan algún caramelito.
4. La tarea de los niños era circular con un marcador de felpa todos los adjetivos de la oración.
5. La palta no me gusta nada.
6. A los príncipes se los recibió con fanfarrias y discursos.
7. Andá rápido a buscar el extinguidor de fuego, que se quema la cocina.
8. El cura en el púlpito habló del arcángel Gabriel en su sermón.
9. Yo siempre les inculqué buenos valores a mis hijos.
10. De tanto arremangarte la camisa tenés las mangas todas arrugadas.
11. No podés reírte de la pulsera de marfil que me compré.
12. Ella va a trabajar en la campaña presidencial el año que viene.
13. Ganas de viajar a Bélgica no me faltan, pero me falta la plata para el pasaje de avión.
14. Quince minutos estuve tocando el timbre y no atendía nadie.
15. Hay ángulos obtusos y agudos.
16. Catamarca es una provincia del noroeste argentino que nunca visité, aunque sí visité Tucumán y Salta que están cerca.
17. Física y álgebra son las materias que siempre me traen problemas.
18. Dame la mayonesa para condimentar la ensalada de alfalfa esa, que es lo más insípido que probé en mi vida.
19. No sé si los cónyuges tienen derechos de visita.
20. La carpa se mojó toda.

(The highlighted sequences indicate the sibilants that were analyzed in Chapter 2, /s/ are in blue.)
21  La bikini de Virginia tiene pintitas amarillas muy lindas.
22  Tito estudió danza en Algeciras al sur de España y ahora está hecho un 
flamenquero profesional.
23  Humberto quedó huérfano de muy joven.
24  El objetivo del juego es hacer un círculo con los bloquecitos.
25  Guille tuvo la gamba enyesada dos meses y todos le firmaban el yeso 
cuando lo iban a visitar.
26  Sos una genia total.
27  Ese puesto conlleva mucha responsabilidad.
28  El uranio es un elemento radioactivo.
29  Gire a la izquierda en la intersección de Ayacucho y Santa Fe.
30  El psicólogo puso mucho énfasis en lo importante del subconsciente en 
la vida de una persona.
31  Señorita enfermera, por favor, póngale una inyección de anestesia y 
después enyésele la pierna.
32  El quirquincho es un armadillo sudamericano.
33  La interfaz del programa es muy mala.
34  El ejército se estableció en el territorio adquirido en el tratado y mantuvo 
a los habitantes en abyecta pobreza.
35  Tarzán cazaba jirafas y yacarés en la selva virginal africana.
36  Los delfines se veían nadando justo bajo la superficie del agua.
37  Bernardo siempre le arrancaba los geranios del jardín a mamá y ella lo 
retaba sin parar.
38  Mi tío tiene una finquita en Salta donde cría llamas y alpacas para 
producir lana.
39  Vayan a prepararle un té de manzanilla a la abuela, por favor.
40  Félix es el perpetuo estudiante, empezó por estudiar física cuántica, 
después le dio por ser filósofo y ahora quiere irse a estudiar una lengua 
indígena del Pacífico sur.
41  Tiraste las alpargatas al piso y me tropecé con ellas esta mañana.
42 Sergio se leyó El Principito cuando tenía once años y hasta hoy sigue leyendo constantemente.

43 A pesar de las objeciones de la defensa, condenaron al acusado por falsificar billetes extranjeros.

44 Queríamos ir al campamento rápido, pero no sabíamos con certeza qué camino seguir.

45 Dentro del museo hay ánforas romanas y momias egipcias embalsamadas hace miles de años.

46 Virginia está obsesionada con el grupo Pimpinela desde que era chica.

47 Mi tía se volvió una compradora compulsiva con la tarjeta de crédito.

48 Lo peor fue la inyección que le pusieron, nunca pensó que esa jeringuita ínfima fuese a doler tanto.

49 Pidieron donantes de sangre y por eso me presenté en el hospital.

50 Su ignorancia se puede adjudicar a su falta de interés.

51 Pidan lo que quieran que nosotros invitamos esta noche.

52 Yolanda le quiere tejer unos escarpines a la pobre huerfanita que vio.

53 Déjense de blandenguerías y hagan los ejercicios bien.

54 Esa señora no quiere que le inculquen falsedades a su hijo.

55 Desde el accidente Juan renguea cuando camina.

56 Vino la azafata con la cena, pero vos estabas dormida.

57 Al criminal le sacaron una foto de perfil y otra de frente.

58 Tenemos que comprar una cámara digital con una pantalla más grande.

59 Manejaba tan rápido que volcó el auto en una zanja profunda al lado de la ruta.

60 Atahualpa Yupanqui fue un importantísimo cantante de folclore argentino del siglo veinte.

61 Bajen la voz.

62 David no paró de reirse a carcajadas con nuestra lambada en el carnaval carioca.

63 La cordillera del Himalaya está al norte del subcontinente indio.
Se realizó un estudio en la subpoblación de pingüinos de la Península de Valdés.

Tantas veces que Marta visitó Buenos Aires y nunca vio la belleza de un jacarandá en flor.

Ponele un moñito al regalo.

Las silílices y los elfos son seres fantásticos.

Sin tonic era lo único que bebía mi tío y llegaba después tambaleándose a su casa.

Los pasajeros del yate desembarcaron por la rampa de atrás.

Le vale estos analgésicos a Virginia que le duele hasta la última vértebra del cuerpo.

“Guerrilleros impiden el paso de delegación extranjera”, anunciaron en la radio.

Gente falsa hay en todos lados, pero gente honesta también.

Banfield y Temperley son localidades adyacentes de la provincia, pero pertenecen a partidos distintos.

Mi prima tiene la nariz muy perfilada y los ojos grandes.

¡Qué manía la tuya de siempre llegar tarde!

En esta casa el lenguaje obsceno está absolutamente prohibido.

Angel tomaba clases de solfeo en el colegio.

El herpes y la sífilis son enfermedades venéreas.

Sarita entró a la cocina sin zapatillas, volcó el detergente y se torció el tobillo derecho.

Guillermo y Rafael, acérquense más acá, que si no la foto queda mal.

El maltés siempre persigue un gato que vive al lado.

Quizás el intendente logre embellecer la ciudad con su nueva campaña de parquización urbana.

Ir al Tigre y pasear en barco por el delta es una buena alternativa para un fin de semana.

Pasan los años y mi tía sigue con la terquedad de siempre.
| 85  | La estirpe de los Wilson lleva aquí varias generaciones. |
| 86  | Yirando hasta la madrugada y durmiendo hasta el mediodía no vas a progresar en la vida. |
| 87  | En un rectángulo cada lado es perpendicular al otro. |
| 88  | Saquen a ese perro callejero de la casa, con lo enclenque que está, seguro que está enfermo con sarna o algo. |
| 89  | Tu tía dijo que tenía palpitaciones porque sentía como alfileres y golpecitos en el pecho. |
| 90  | Desde chiquitito Roberto siempre quiso ser carpintero o payaso de circo. |
| 91  | Quéjense con el gerente si tienen algún problema con los vendedores de la tienda. |
| 92  | Comprá un sachet de leche. |
| 93  | Mi familia siempre acampa aquí, pero de noche el bosque se vuelve laberíntico y no sé por dónde agarrar. |
| 94  | Me metí el llavero en el bolsillo cuando salí de casa. |
| 95  | Con el telescopio del observatorio se veían galaxias lejanas. |
| 96  | Ganó Boca Juniors el torneo clausura, así que mi primo va a estar insoportable. Ahora nos bancaremos todas sus cargadas. |
| 97  | Pepe me trajo de la Patagonia una mermelada de zarzamora riquísima. |
| 98  | Después de la última inquisición intimidante del fiscal Sánchez, el acusado confesó todo. |
| 99  | La campana da la señal cada hora. |
| 100 | Virgilio alquilaba un departamento por poca plata, pero el subte le quedaba lejos. |
| 101 | Decidió irse del ámbito universitario para trabajar de arquitecto en una compañía privada. |
| 102 | No lo vas a encontrar jamás en esta multitud de gente. |
| 103 | Había un miedo subyacente en su tono de voz. |
| 104 | “Yira, yira” es el tango preferido de mi mamá. |
| 105 | Salta y Catamarca son provincias argentinas del noroeste, Tucumán y... |
| 106 | Dijiste que tenías plata, pero con eso no te alcanza para nada. |
| 107 | Germán Reyes se presentó al comité olímpico para recibir el premio por obtener una puntuación perfecta en gimnasia. |
| 108 | Benjamin dice que es absurdo ver una película extranjera sin subtítulos si no hablás el idioma. |
| 109 | Fantástico, entonces agarrá vos los alfajores mientras Martín lleva la rosca de rye afuera. |
| 110 | El maltrato conyugal es un delito. |
| 111 | La entrada a la quinta está tan apartada de la ruta, que me pasé de largo. |
| 112 | La mermelada está bienisima para ponérsela a los panqueques o a las tostadas. |
| 113 | La planta que había crecido tan bien empezó a empequeñecerse después de la sequía. |
| 114 | Ella no quiere ir al dentista a tratarse la gingivitis que tiene. |
| 115 | Temen que la propuesta del diputado termine por engendrar enemistad entre los representantes. |
| 116 | Me da pena que el verano se esté terminando. |
| 117 | Temen que los damnificados por las inundaciones no puedan volver por más de dos meses a sus hogares. |
| 118 | La cartelera marplatense presenta un sinfín de espectáculos de revista. |
| 119 | Marta se compró un perrito maltés y le puso Melquíades de nombre. |
| 120 | Falta invitar a Elsa y a Rafael a la fiesta, les mando la invitación mañana. |
| 121 | No inviten a Raquel a la fiesta, que mi tía no se la banco para nada. |
| 122 | El calzado femenino está en el cuarto piso. |
| 123 | Tegucigalpa es la capital de Honduras. |
| 124 | Vandalizar un símbolo nacional es ofensivo. |
| 125 | Yo no les estoy enseñando a dibujar para que después calquen los dibujos del libro. |
126 No vayan a zarparse con la música que después se quejan los vecinos.
127 Parece que el Príncipe de España criticó a los que reclaman independencia en el País Vasco.
128 Carlos es un perfeccionista empedernido.
129 Su observación subjetiva demostró lo tonto que era.
130 El dibujo se calca primero y luego se traslada a una cartulina.
131 El presidente sudcoreano fue a Egipto a visitar las pirámides y la esfinge de Giza.
132 Felicitaciones al capitán Domínguez por su desempeño ejemplar como jefe de la policía provincial y por siempre estar a la defensa de la seguridad del pueblo.
133 Gilberto Tejada, por favor acérquese a la ventanilla para ser atendido.
134 ¡Qué idea tan genial tuviste!
135 El profesor Peña es el fantárrón más grande que tiene el departamento de física de la universidad.
136 Él es diabético y se inyecta insulina todos los días.
137 Vimos un monito en el zoológico.
138 Faltaría llamar a Martín para completar la comparsa antes del carnaval.
139 El manganeso es un elemento químico de la tabla periódica.
140 En esa plaza convergen todas las calles del pueblo.
141 Perdón señora, pero su auto está goteando un líquido negro por debajo.
142 Alguien golpea a la puerta.
143 Llenele el tanque al auto en la estación de servicio Shell de la esquina.
144 Pense que nunca me iba a curar del dengue que me dio en el Caribe.
145 Mi tío es un huraño que no sale ni a la esquina.
146 Jorge se pone nostálgico cuando cuenta que su equipo salió subcampeón por un penal robado.
147 Ayer me desperté con el cantar de los pájaros en el jardín.
148 Es difícil imaginarse el abismo infinito del universo y los miles de centenares de estrellas que hay.
Dije que no podías ir de jeans y zapatillas al casamiento de tu hermana, ¿qué te pusiste?

Trató de defenderse como pudo, pero le llevaron la billetera con todos los documentos.

Me desperté muy tarde.

El conjunto de short y camisa resaltaba la esbeltez de su figura.

Vayan al super a buscar berenjenas para el escabeche y compren un paquete pequeño de alcaparras también.

Bife con ensalada para mí y milanesa con puré para el nene.

Japón es un país que me encantaría visitar cuando tenga la plata suficiente.

Yo me las sigo ingeniando para salir adelante.

“Jaque” es la interjección que se dice en ajedrez cuando se amenaza al rey.

Marcá los cinco dígitos de tu contraseña rápido.

Ya Mercedes terminó el doctorado en ciencias políticas, pero no va a ejercer su profesión todavía.

No salpiques agua afuera de la pileta.

No te van a permitir llevar objetos filosos en el avión.

Los esclavos subyugados se rebelaron contra el ejército invasor.

Bajame de allá la ginebra para hacerle un gin tonic a tu tío que le encanta beber eso.

Los chicos miraron el show de títeres sin parpadear ni un segundo.

Sigan caminando descalzas con este frío, que ín verán como se enferman mañana.

Depende de lo que negocien el sindicato y el gobierno, pero es posible que los trabajadores hagan huelga mañana.

Tú malos y deciles que traigan tarta de manzana con helado.

Eugenio compró témpera verde.

El índice de precipitación es lo que determina si un lugar es un desierto o
<table>
<thead>
<tr>
<th>no.</th>
<th>Fingí que te gustó el regalo y sonréí para que tu tía no se vaya a sentir mal.</th>
</tr>
</thead>
<tbody>
<tr>
<td>170</td>
<td>Guerra es un apellido bastante frecuente en castellano.</td>
</tr>
<tr>
<td>171</td>
<td>&quot;Beta&quot; es la segunda letra del alfabeto griego.</td>
</tr>
<tr>
<td>172</td>
<td>Salgan al jardín, chicas, llevense las siestas para que jueguen con las muñecas nuevas.</td>
</tr>
<tr>
<td>173</td>
<td>En el jardín hay un aljibe antiguo y un rosal con un perfume riquísimo.</td>
</tr>
</tbody>
</table>
AX Discrimination Task Instructions

The following are the instructions that were presented to the participants in the AX discrimination experiment discussed in Chapter 3. Each page was presented in print to the participant before the appropriate block. The bullet points after Práctica 1 and Práctica 2 were presented as feedback after each of practice sessions. The order of the phrase-initial and intervocalic blocks was counterbalanced.
El experimento tendrá cuatro partes: Práctica 1, Parte 1, Práctica 2, Parte 2
Durará aproximadamente 20 minutos.

Práctica 1

Escuchará varios pares de palabras inventadas. Para cada par de palabras tendrá que decidir si las dos palabras que escuchó son IGUALES (dos ejemplares de la misma palabra) o DIFERENTES (dos palabras distintas).

Si considera que el par de palabras escuchadas se trata de dos ejemplares de la misma palabra haga clic en IGUALES y si considera que el par de palabras escuchadas se trata de dos palabras distintas haga clic en DIFERENTES.

Como esta práctica es sólo para familiarizarse con el formato del experimento, la respuesta correcta aparecerá en la parte superior de la pantalla mientras escucha el par de palabras. Después de que haya hecho clic en IGUALES o DIFERENTES escuchará el par de palabras siguientes.

Cuando haya terminado la práctica podrá hacerle cualquier pregunta al investigador.
• Notó que si las palabras eran diferentes, se diferenciaban siempre en el primer sonido.

• Las palabras de la Parte 1 seguirán este mismo patrón.
Parte 1

Escuchará varios pares de palabras en un idioma extranjero. Para cada par de palabras tendrá que decidir si las dos palabras que escuchó son IGUALES (dos ejemplares de la misma palabra) o DIFERENTES (dos palabras distintas).

Si considera que el par de palabras extranjeras escuchadas se trata de dos ejemplares de la misma palabra haga clic en IGUALES y si considera que el par de palabras extranjeras escuchadas se trata de dos palabras distintas haga clic en DIFERENTES.

El formato de la Parte 1 será como el de la Práctica 1, sólo que esta vez no aparecerán las respuestas correctas en la pantalla.
Práctica 2

La práctica 2 es muy similar a la Práctica 1

Escuchará varios pares de palabras inventadas. Para cada par de palabras tendrá que decidir si las dos palabras que escuchó son IGUALES (dos ejemplares de la misma palabra) o DIFERENTES (dos palabras distintas).

Si considera que el par de palabras escuchadas se trata de dos ejemplares de la misma palabra haga clic en IGUALES y si considera que el par de palabras escuchadas se trata de dos palabras distintas haga clic en DIFERENTES.

Al igual que en la Parte 1, la respuesta correcta aparecerá en la parte superior de la pantalla mientras escucha el par de palabras. Después de que haya hecho clic en IGUALES o DIFERENTES escuchará el par de palabras siguientes. Cuando haya terminado la práctica podrá hacerle cualquier pregunta al investigador.
• Notó que:
  – Si las palabras eran diferentes, se diferenciaban siempre en el sonido entre las dos vocales.
• Las palabras de la Parte 2 seguirán este mismo patrón.
**Parte 2**

Escuchará varios pares de palabras en un idioma extranjero. Para cada par de palabras tendrá que decidir si las dos palabras que escuchó son IGUALES (dos ejemplares de la misma palabra) o DIFERENTES (dos palabras distintas).

Si considera que el par de palabras extranjeras escuchadas se trata de dos ejemplares de la misma palabra haga clic en IGUALES y si considera que el par de palabras extranjeras escuchadas se trata de dos palabras distintas haga clic en DIFERENTES.

El formato de la Parte 2 será como el de la Práctica 2, sólo que esta vez no aparecerán las respuestas correctas en la pantalla.
ABX Discrimination Task Instructions

The following are the instructions that were presented to the participants in the ABX discrimination experiment discussed in Chapter 3. The instructions were presented as part of the experiment in DMDX (Forster and Forster 2003). The order of the phrase-initial and intervocalic blocks was counterbalanced.
PRÁCTICA 1

Escuchará varios tríos de palabras inventadas.
La primera y la segunda palabra siempre serán diferentes entre sí.
La última palabra que escuche será otro ejemplo
o de la primera palabra o de la segunda.

Si considera que la última palabra corresponde a la primera palabra,
presione la tecla marcada ABA.
Si considera que la última palabra corresponde a la segunda palabra,
presione la tecla marcada ABB.

[Presione el espacio para continuar con las instrucciones]

ABA o ABB son las únicas opciones,
para cada trio tendrá que elegir una de estas dos opciones.
Escuche atentamente, ya que no podrá volver a escuchar ningún trio.
Si se equivoca no podrá volver para atrás,
simplemente continúe con el experimento.

Conteste lo más rápido que pueda. Si tarda más de 4 segundos,
perderá su oportunidad y pasará al próximo trio automáticamente.
Después de que haya presionado la tecla ABA o ABB,
escuchará el trio de palabras siguientes.

Si tiene alguna pregunta llame al investigador ahora.
Si no tiene preguntas presione el espacio para comenzar la Práctica 1.
Fin de la Práctica 1.

Habrá notado que las diferencias en entonación no eran importantes.

La diferencia crucial entre las palabras estaba en el sonido de la primera consonante.

La Parte 1 seguirá un formato similar.

Si tiene alguna pregunta llame al investigador ahora.
Si no tiene preguntas presione el espacio para leer las instrucciones de la Parte 1.
PARTE 1

Escuchará varios tríos de palabras en un IDIOMA EXTRANJERO. La primera y la segunda palabra siempre serán diferentes entre sí. La última palabra que escuche será otro ejemplo o de la primera palabra o de la segunda.

Si considera que la última palabra corresponde a la primera palabra, presione la tecla marcada ABA.
Si considera que la última palabra corresponde a la segunda palabra, presione la tecla marcada ABB.

[Presione el espacio para continuar con las instrucciones]

ABA o ABB son las únicas opciones, para cada trió tendrá que elegir una de estas dos opciones. Escuche atentamente, ya que no podrá volver a escuchar ningún trió. Si se equivoca no podrá volver para atrás, simplemente continúe con el experimento.

Conteste lo más rápido que pueda. Si tarda más de 4 segundos, perderá su oportunidad y pasará al próximo trió automáticamente. Después de que haya presionado la tecla ABA o ABB, escuchará el trió de palabras siguientes.

Si tiene alguna pregunta llame al investigador ahora. Si no tiene preguntas presione el espacio para comenzar la Parte 1.
Fin de la Parte 1.

Si quiere puede tomar un pequeño descanso.

Presione el espacio para continuar con las instrucciones de la Práctica 2.
PRÁCTICA 2

Escuchará varios tríos de palabras inventadas.
La primera y la segunda palabra siempre serán diferentes entre sí.
La última palabra que escuche será otro ejemplo
o de la primera palabra o de la segunda.

Si considera que la última palabra corresponde a la primera palabra,
presione la tecla marcada ABA.
Si considera que la última palabra corresponde a la segunda palabra,
presione la tecla marcada ABB.

[Presione el espacio para continuar con las instrucciones]

ABA o ABB son las únicas opciones,
para cada trío tendrá que elegir una de estas dos opciones.
Escuche atentamente, ya que no podrá volver a escuchar ningún trío.
Si se equivoca no podrá volver para atrás,
simplemente continúe con el experimento.

Conteste lo más rápido que pueda. Si tarda más de 4 segundos,
perderá su oportunidad y pasará al próximo trío automáticamente.
Después de que haya presionado la tecla ABA o ABB,
escuchará el trío de palabras siguientes.

Si tiene alguna pregunta llame al investigador ahora.
Si no tiene preguntas presione el espacio para comenzar la Práctica 2.
Fin de la Práctica 2.

De nuevo, las diferencias en entonación no eran importantes.

Habrá notado que esta vez la diferencia crucial estaba en el sonido de la consonante en el medio de la palabra.

La Parte 2 seguirá un formato similar.

Si tiene alguna pregunta llame al investigador ahora.
Si no tiene preguntas presione el espacio para leer las instrucciones de la Parte 2.
PARTE 2

Escuchará varios tríos de palabras en un IDIOMA EXTRANJERO.
La primera y la segunda palabra siempre serán diferentes entre sí.
La última palabra que escuche será otro ejemplo o de la primera palabra o de la segunda.

Si considera que la última palabra corresponde a la primera palabra, presione la tecla marcada ABA.
Si considera que la última palabra corresponde a la segunda palabra, presione la tecla marcada ABB.

[Presione el espacio para continuar con las instrucciones]

ABA o ABB son las únicas opciones, para cada trío tendrá que elegir una de estas dos opciones.
Escuche atentamente, ya que no podrá volver a escuchar ningún trío.
Si se equivoca no podrá volver para atrás, simplemente continúe con el experimento.

Conteste lo más rápido que pueda. Si tarda más de 4 segundos, perderá su oportunidad y pasará al próximo trío automáticamente.
Después de que haya presionado la tecla ABA o ABB, escuchará el trío de palabras siguientes.

Si tiene alguna pregunta llame al investigador ahora.
Si no tiene preguntas presione el espacio para comenzar la Parte 2.
Fin del experimento.

Por favor llame al investigador.
Perceptual Assimilation and Rating Task

Instructions and Worksheet

The following are the instructions and worksheets for the Perceptual Assimilation and Rating task discussed in Chapter 4. The order of the phrase-initial and intervocalic blocks was counterbalanced.
Transcripción y Puntuación

A continuación oirá algunas de las palabras que escuchó en el experimento. Transcriba cada palabra como la escribiría en castellano. Algunas letras ya están escritas, transcriba en el espacio en blanco el sonido que falté. Si quiere hacer algún comentario, escribalo al lado. Después de transcribir el sonido, mire la letra (o letras) que transcribió, y piense cómo se pronuncia(n) habitualmente en castellano. El sonido que escuchó, ¿es un buen ejemplo de esa pronunciación habitual? Dele una puntuación del 1 al 5 de acuerdo a cuán buen ejemplo sea de esa pronunciación habitual. (1 = muy mal ejemplo, 5 = muy buen ejemplo).

Bloque 1 – posición inicial

<table>
<thead>
<tr>
<th>Orden</th>
<th>Consonante faltante</th>
<th>Resto de la palabra</th>
<th>Comentario</th>
<th>Puntuación</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ejemplo</td>
<td>r ele</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>Ejemplo</td>
<td>r ele</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>1</td>
<td>ili</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>2</td>
<td>ili</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>3</td>
<td>ala</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>4</td>
<td>ili</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>5</td>
<td>ili</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>6</td>
<td>ala</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>7</td>
<td>ili</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>8</td>
<td>ala</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>9</td>
<td>ala</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>10</td>
<td>ili</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>11</td>
<td>ili</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>12</td>
<td>ala</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>13</td>
<td>ala</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>14</td>
<td>ili</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>15</td>
<td>ili</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>16</td>
<td>ala</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>17</td>
<td>ala</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>18</td>
<td>ili</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>19</td>
<td>ala</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>20</td>
<td>ala</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>21</td>
<td>ili</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>22</td>
<td>ili</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>23</td>
<td>ala</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>24</td>
<td>ala</td>
<td></td>
<td></td>
<td>1 2 3 4 5</td>
</tr>
</tbody>
</table>
Bloque 2 – posición intervocálica

<table>
<thead>
<tr>
<th>Orden</th>
<th>Comienzo de la palabra</th>
<th>Consonante faltante</th>
<th>Resto de la palabra</th>
<th>Comentario</th>
<th>Puntuación</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ejemplo</td>
<td>le rr u</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ejemplo</td>
<td>le rr u</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>li i</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>la a</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>li i</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>la a</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>li i</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>la a</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>la a</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>li i</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>la a</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>li i</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>li i</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>la a</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>li i</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>la a</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>li i</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>la a</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>li i</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>li i</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>la a</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>la a</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>li i</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>li i</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>la a</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>la a</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Bibliography


Davidson, Lisa, and Jason Shaw. submitted. Sources of illusion in consonant cluster perception. *Journal of Phonetics*.


219


